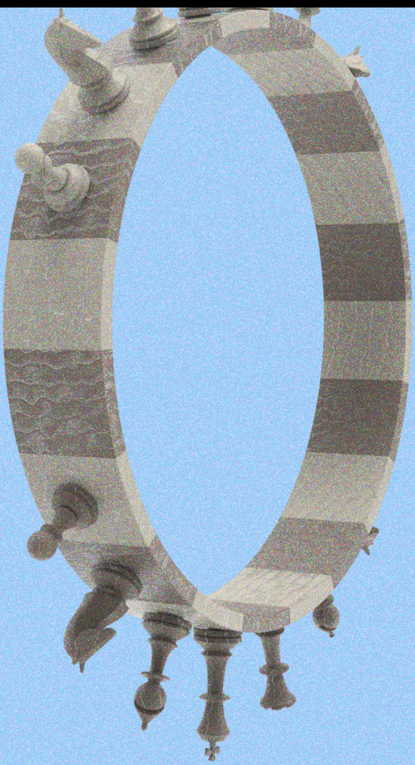
Controlling the Unknown: A Game Theoretic Perspective

Wouter Jongeneel

Master of Science Thesis



Delft University of Technology

Delft Center for Systems and Control

Controlling the Unknown: A Game Theoretic Perspective

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

Wouter Jongeneel

November 17, 2019

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

Delft University of Technology
Department of
Delft Center for Systems and Control (dcsc)

The undersigned hereby certify that they have read and recommend to the Faculty of Mechanical, Maritime and Materials Engineering (3mE) for acceptance a thesis entitled

Controlling the Unknown: A Game Theoretic Perspective

by

Wouter Jongeneel

in partial fulfillment of the requirements for the degree of

Master of Science Systems and Control

Dated: November 17, 2019

Supervisor(s): ______________________
dr. P. Mohajerin Esfahani

Reader(s): ______________________
prof.dr.ir. M. Verhaegen

____
dr. G. Giordano

____
dr. M.T.J. Spaan

Abstract

We consider the problem of safely controlling an unknown stochastic linear dynamical system subject to an infinite-horizon discounted quadratic cost. Many of the existing model-based approaches for handling the corresponding robust optimal control problem resort to game theoretic formulations of the uncertainty, either explicit- or implicitly. It is widely known that in practice the corresponding control laws can be rather conservative. In this work, we give further theoretical evidence that this is an inherent property of the underlying game theoretic formulation. We show that the most common uncertainty sets, for example resulting from linear least-squares identification, are almost surely different from the geometry a game theoretic adversary samples from. Nevertheless, we provide theoretical- and empirical evidence that a game theoretic control law has favourable properties over the nominal control law when the estimated model is obtained using regularized linear least-squares.

Contents

Before the Beginning

This thesis marks the end of two years at the Delft Center for Systems and Control (DCSC). During that period my interest in the field has only grown and this project was no exception.

I became especially excited about the rapidly developing paradigm of *finite-sample* properties of dynamical control systems. Due to the vast literature, most notably by Ben Recht and his group, we could identify a direction to work on: *Exact and tractable reformulations of robust linear quadratic regulator problems.* It turned out to be an interesting direction, with in my opinion as its main outcome a lot of new open problems.

This work was not possible without the support of the following people. Most and for all I would like to thank my supervisor dr. Peyman Mohajerin Esfahani. His excitement has been exceptional, I truly wonder, when, and if you sleep. The help from overseas is also highly appreciated; dr. Tyler Summers for introducing us to the possibilities of game theory and Benjamin Gravell for his notes and remarks on policy iteration. Moreover, I would like to thank all committee members, the TU Delft, in particular DCSC, 3ME and EWI, but also the library, if it was not for their enormous collection of non-digitalized books I would not have found the, for me highly influential, thesis by Jan Willem Polderman. Even more so, I would like thank all these individuals who put their notes, presentations, books and other material online. I hope to return the favour one day.

At last, besides the abundance of technical support I am grateful to my (feathery) friends and family.

I no longer view Linear Control as merely a prerequisite to learn Non-linear Control, but rather fully appreciate the wealth of structure it displays, *the window remains open.*

–Wout

“Mathematics is the part of physics where experiments are cheap.”
— *V.I. Arnold*

Chapter 1

Introduction

1-1 The Linear Quadratic Regulator, Still Relevant

It can be argued[1] that back in 1696, Bernoulli started the field of *Optimal Control* in Groningen[SW97]. He was interested in how to slide from point A to B, under gravity, the fastest, *i.e.,* the *brachistochrone* problem, which is a minimal-time problem. Together with the fact that Bernoulli was a mentor of Euler, thereby indirectly, Lagrange, the field of the calculus of variations was born. It took a few centuries before the field of Optimal Control was actually established. Around the 1950s the drawbacks of the frequency domain approach within control were recognized and as Lewis [Lew92] puts it nicely *"On the failure of any paradigm, a return to historical and natural first principles is required."* Think about Bernoulli, linking optimality and natural first principles. Then just before the end of the 1950s, Bellman and Pontryagin brought to life the two pillars of Optimal Control theory: *Dynamic Programming* and *the Maximum Principle*, respectively. This was an extraordinary fruitful period, since in 1960, Kalman, Bertram and Bucy introduced the power of Lyapunov theory, the Linear Quadratic Regulator (LQR) and the Kalman filter. Even the dual relationship was immediately acknowledged. This work hinges on the elegant stability theory by Lyapunov from 1892 but also on the work of Jacobo Riccati, who in the period of 1719-1724 started the investigation in the family of equations now bearing his name. Interestingly, the first correspondence Riccati had about this equation was with a Bernoulli (Nicolaus II) [BLW91, ch. 1]. After the legendary papers by Kalman and coworkers, LQR theory was born and further developed by people like Brockett, Willems and Wonham, to name a few. The main reason for further development and study was of course the remarkable fact that we have, up to algebraic equations, closed-form solutions.

But why do people *still* care about this LQR problem?

Although most, if not all, real world dynamical system are non-linear, Linear Quadratic (LQ) problems still appear in for example economics [Pri10] and medicine [CDS+09], but most and for all, the dual problem, the Kalman filter is still heavily used. For example, this filter got us to the moon [GA10]! This is however not our reason of interest.

[1] I have the same bias as those authors. He was Swiss, but working in Groningen.

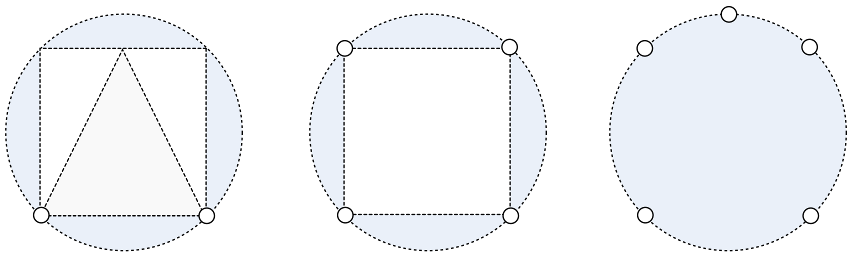
1-1-1 Finite-Samples, a Paradigm Shift

Towards safe and practical control algorithms we can observe a paradigm shift in the community. The general approach has always been to use some statistical algorithm to estimate model parameters based on experimental data and design a controller solely based on this model. As elegant as the theory is, classical system identification theorems are asymptotic, *i.e.,* they hinge on *infinite* sample lengths (*cf.* [VV07, ch. 9]). In practice we do not have that much data, instead, one is interested in what can be said about controller performance, on a yet partially unknown system, based on *finite* data.

A first step in this direction is to at least identify *online*, *i.e.,* be *adaptive* [AW73, CK98], which nowadays one might refer to as *reinforcement learning.* This simply means that at each step you update your model based on incoming information. The main paradigm shift we currently see (in control theory) was first described by Claude-Nicolas Fiechter. In his work he introduced *probabilistic performance bounds* for the discounted discrete-time LQR problem [Fie97], which was recently formalized to finite-sample and regret-bounds for averaged cost LQR [DMM+17, DMM+18]. Initial work on the sample complexity in Reinforcement Learning has been done before in [Kak03]. However, there the focus was on finite state and action spaces, while we consider them to be uncountable.

All of these approaches aim for "*end-to-end*" (from data to policy) *finite-sample* (sometimes called "non-asymptotic") probabilistic performance bounds. For example, under certain assumptions, this allows one to say that after having obtained N data-points, a Least-Squares identification approach combined with LQR design will stabilize the *real* system with a probability of at least $1-\delta$. The most straight-forward path to these bounds is (1) obtain data, (2) identify a model, (3) compute probabilistic error bounds on the parameters of the model and (4) design a controller which can handle these errors in the model. These kind of controllers are referred to as ***robust controllers*** with probabilistic performance bounds.

For example, if we take the example from Figure 1-1 (a), then after just two sample points we cannot distinguish the three shapes. However, in (b), after four points we know it cannot be a triangle and in (c), after just five points, we *can* be sure it is a circle, although the shape is made up out of a continuum of points.



(a) (b) (c)

Figure 1-1: After five samples we can distinguish a circle from a square and a triangle.

The crux is, under certain conditions, one can be rather sure about the parametrization of a linear dynamical system, even after having just obtained a finite set of data. This would allow for safely introducing our favourite control algorithms into the real world, without perfect models.

Now, the incentive to study the discrete-time LQR problem within the aformentioned model-based learning framework follows from the problem being an ideal benchmark. The reason is four-fold; (i), the Linear Quadratic Regulator problem is well-studied (*cf.* [Kal60, BLW91, LR95, Ber05]) and has a closed-form solution which makes it already a good benchmark problem, especially since the action and state space are uncountable; (ii), it is notorious for being a linear optimal control problem for which self-tuning (see [AW73]) does not hold [Pol87, vS94]; (iii), in contrast to the continuous-time LQR, the LQG [Doy78] as well as the discrete-time LQR (*cf.* [KS72]) have generally no, or worse stability margins [Sha86]; (iv), our understanding of the corresponding perturbating theory is limited, especially in the discrete-time case, and often hinges on assuming A to be invertible [LR95, KPC93, Sun98] such that the robust LQR (RLQR) problem is not merely a corollary to LQR. In other words, the LQR problem is non-trivial when the dynamical system is partially unknown, but if it is known, we understand it very well.

As mentioned before, this framework can be largely splitted into two parts; (1), *the identification* and (2), *the design of a robust control law*. These are both interesting and still challenging directions. However, motivated by the involved and truncated optimization programs in [DMM$^+$17, DMM$^+$18] we set out to investigate if we can retain exactness of the RLQR solution while not sacrificing tractability of the algorithm. The motivation for tractability is obvious, the motivation for exactness follows from the desire to overcome, or at least, better understand conservatism within robust control.

The desire to solve RLQR problems is not new, however, neatly linking identification and control is still largely an open problem is, *e.g.,* see the 2005 survey paper [Gev05]. A classical μ-synthesis approach is indeed generally intractable [PR93, BYDM93] while as explained in Chapter 2, a tractable Linear Matrix Inequality (LMI) approach like proposed in [dOBG99] may be conservative, plus, to paraphrase Gevers [Gev05, p.9], in many cases is the uncertainty set *God-given*, instead of directly related to statistics. This work investigates to what extend dynamic game theory can be a middle-ground. The motivation for game theory is the close (multiplier) relation to robust control, while having readily available computational tools (chapter 3 elaborates on this remark).

1-2 Contribution

Towards a better understanding of tractable and exact design of robust linear quadratic regulators, our contributions are:

(i) New set theoretic interpretation of Linear-Quadratic (LQ) dynamic games as robust LQ regulators, see Definition 3-1.1, Proposition 3-2.1 and Corollary 4-3-1 plus section 3-2-1 for an extension, with respect to [JSM19], to uncertainty in A *and* B.

(ii) An exact solution to a robust LQR problem, including a method to, numerically and analytically, obtain worst-case models given any stabilizing linear controller, see Proposition 3-2.3, Theorem 3-2.4 and Lemma 3-4.1. In section 3-2-3 these results are linked to standard norm-balls.

(iii) Moreover, we provide insights in *qualitative* features of the worst-case model in sections 3-3 and 3-4-2-3-4-4. Specifically, in Lemma 3-3.3.(iii) and section 3-4-2-5 we show that our framework opens up the door for robust optimization over topologically equivalent drift terms. Furthermore, by using the structure from Definition 3-1.1, we show in section 3-3-1 that ellipsoidal confidence sets, as often seen in practice, are almost surely conservative within a game theoretic context. In addition, we show in Lemma 3-3.10 that generic behaviour resulting from additive perturbations to (A, B) can almost surely not be generated from a dynamic game. Even more so, Lemma 3-3.3.(iv) shows that the our worst-case model is larger, in Frobenius-norm, than the nominal model , which turns out to be in conflict with standard unbiased estimators, *e.g.,* see section 3-4-3-1. On the other hand, this lemma suggests that our game theoretic controller might be a more natural choice than the nominal LQ regulator once ℓ_2-regularization is used during identification. Section 3-4-4 and secion 3-5-2 present affirmative empirical evidence that our framework *is* capable of outperforming nominal control laws on the Laplacian example from [DMM$^+$17, DMM$^+$18], indeed, especially after introducing ℓ_2-regularization during identification.

These results imply that a game theoretic formulation of robust LQ regulators is inherently conservative, but nevertheless has a rich structure, which the author believes, can, and should be, further explored.

In addition, chapter 4 shows initial work on gradient algorithms within dynamic game theory, which in this thesis neatly functions as a ridiculously long (but otherwise missing) proof of path-connectedness of our uncertainty set.

Part of the initial work resulted in the short conference paper [JSM19]; this thesis work is a significant elaboration. However, there is indeed substantial overlap with introductory parts of chapter 2 and chapter 3.

1-3 Outline, Structure of the Thesis

In section 1-1 we have just outlined the motivation behind this thesis. Then, in chapter 2 we formally introduce the problem and motivate the viewpoint taken in the next chapters. Chapter 3 presents the main results, starting with the introduction of a new uncertainty set and corresponding RLQR solution, followed by a structural investigation and ample numerical experiments. In chapter 4 we take a brief detour and discuss gradient-based policy iteration in the context of dynamic games. Finally, chapter 5 contains open-problems and concludes the work.

Two appendices accommodate auxiliary tools (A) and further background information (B).

All simulations are carried out using MATLAB R2018b. Other figures are made using Inkscape, with textext[2]. In addition, the cover is made using SolidWorks 2018, the Gimp 2.10 and online chess models [3].

Notation We use standard notation, but to be clear. Let $\mathbb{R}_{\geq 0}$ denote the set of non-negative real numbers and $\mathbb{N}$ the set of non-negative integers. If $z \in \{1, 2, \ldots, Z\}$, we write $z \in \overline{Z}$.

[2] https://textext.github.io/textext/
[3] https://grabcad.com/library/chess-board-and-pieces-3

The $n \times m$ matrix of all ones is denoted by $1_{n\times m}$ whereas I_n is the identity element of $\mathbb{R}^{n\times n}$ and δ_{ij} is the Kronecker delta. The general linear group is denoted by $\mathsf{GL}(n,\mathbb{R})$, where a superscript indicates the sign of the determinant, *e.g.* $T \in \mathsf{GL}^{+}(n,\mathbb{R}) \iff \det(T) > 0$. The operation $\mathrm{vec}(A)$ vectorizes the matrix A by means of column stacking and $A \otimes B$ denotes the Kronecker product between A and B. 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 $\|A\|_2$ whereas $\lambda(A)$ denotes the entire spectrum of A, $\rho(A) = \max_{i=1,\dots,n} |\lambda_i|$ is the spectral radius of $A \in \mathbb{R}^{n\times n}$ and $\kappa(A) = \sigma_{\max}(A)/\sigma_{\min}(A)$ is the condition number of A. For $\mathrm{Tr}(\cdot)$ being the trace operator, 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. Then $\|X\|^2_{F,Q}$ is used to denote $\mathrm{Tr}(X^\top Q X)$ for $Q \succ 0$. Furthermore, A is said to be *exponentially stable* if its spectrum is fully contained in the open unit disk, denoted $\mathbb{D}_r$, $r = 1$. We speak of *stability* when the spectrum can only be contained in $\mathrm{cl}(\mathbb{D}_1)$, which is the closed unit disk. The expectation operator is given by $\mathbb{E}[\cdot]$, whereas $\mathbb{P}\{\mathcal{E}\}$ denotes the probability of some event $\mathcal{E}$. A random variable with mean μ and covariance Σ for a distribution $\mathcal{P}$ is expressed by $X \sim \mathcal{P}(\mu, \Sigma)$. If a distribution $\mathcal{P}$ is supported on $\Xi \in \mathbb{R}^{\xi}$, it is an element of $\mathcal{P}(\Xi)$. 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*, while in a control context the subscript $c\ell$ stands for *closed-loop*. The limit of a function $f(x)$ for $x \to y$ from below, is denoted $\lim_{x\uparrow y} f(x)$. At last, to emphasize when we use scalar systems, the notation $x_{k+1} = a x_k$ is used instead of the generic capitalized $x_{k+1} = A x_k$.

Acronyms and Terminology We will use just a few acronyms, which are standard anyway. Most notably: Linear Quadratic (LQ), LQ regulator (LQR), Robust LQR (RLQR), Linear Matrix Inequality (LMI), Semi-Definite Program (SDP), Quadratic Program (QP) and Linear Program (LP).

Moreover, we barely use terminology, except for one word: ***nominal***. Here we take the definition of the word as used in the control literature [VF95]. To paraphrase, "*To design a robust controller, we need a nominal model and a description of an additive or multiplicative uncertainty from which the controller will be robust.*" For example, given a (partially) unknown dynamical system $\Sigma : \{x_{k+1} = Ax_k + Bu_k$, let $(\widehat{A}, \widehat{B})$ be estimators for (A, B), where it is known that $A = \widehat{A} + \Delta_A$, $\Delta_A \in \Delta_{\mathcal{A}}$, $B = \widehat{B} + \Delta_B$, $\Delta_B \in \Delta_{\mathcal{B}}$. Then we call $\widehat{\Sigma} : \{x_{k+1} = \widehat{A}x_k + \widehat{B}u_k$ the *nominal model*, the uncertainties extend from there. Moreover, any optimal control law for $(\widehat{A}, \widehat{B})$ will be a *nominal controller*. Indeed, the most frequently used models and controllers in practice are the nominal ones.

Chapter 2

The Robust Linear Quadratic Regulator Problem

In chapter 1 we saw that towards the implementation of safe control algorithms we like to solve some robust control problem. Then, in this chapter we will formally introduce the robust control problem at hand: *the Robust Linear Quadratic Regulator* (RLQR) problem. The question remains, robust with respect to *what*? In section 2-2 we address precisely this question via a discussion on theoretically interesting and practically relevant uncertainty sets.

We will not discuss the standard LQR problem. The reader who wants to learn more about classical linear (discrete-time) control and the corresponding LQ regulator, is referred to [LR95, Cai18].

2-1 Introduction to The Discrete-Time Discounted RLQR

In this section we formally introduce the robust control problem studied in this work. Consider the stochastic linear discrete-time dynamical system

$$\Sigma : \begin{cases} x_{k+1} = Ax_k + Bu_k + v_k, \\ v_k \overset{i.i.d.}{\sim} \mathcal{P}(0, \Sigma_v), \quad x_0 \sim \mathcal{P}(0, \Sigma_0), \end{cases} \tag{2-1.1}$$

where $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}$, $\Sigma_0, \Sigma_v \in \mathcal{S}^n_{++}$, and $\{v_k\}_{k \in \mathbb{N}}$ is a 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 $\mathbb{E}[v_i v_j^\top] = \delta_{ij}\Sigma_v$ for all $i, j \in \mathbb{N}$. In our setting, we assume the system matrices (A, B) not to be known precisely. Our prior estimate of (A, B) is denoted by $(\widehat{A}, \widehat{B})$ such that

$$A = \widehat{A} + \Delta_A, \quad B = \widehat{B} + \Delta_B$$

where (Δ_A, Δ_B) represents our prior estimation error. A particular example of such a setting naturally emerges in statistics or identification problems where $(\widehat{A}, \widehat{B})$ is the current estimate of (A, B), and $\mathbb{A}$ represents an uncertainty set containing (Δ_A, Δ_B) with high probability, *i.e.*, $\mathbb{P}\{(\Delta_A, \Delta_B) \in \mathbb{A}\} \geq 1 - \delta$ for a small prespecified confidence level δ.

Given the matrices $Q \in \mathcal{S}^n_+$, $R \in \mathcal{S}^m_{++}$, discount factor $\alpha \in (0, 1)$ and the tuple $(\widehat{A}, \widehat{B}, \mathbb{A}, \Sigma_v, \Sigma_0)$ we seek an optimal policy $\pi^\star = \{\mu_0^\star, \mu_1^\star, \dots\}$ that solves the discounted ***Robust Linear-Quadratic Regulator*** (RLQR) problem[1]:

$$\begin{aligned}
\inf_{\{\mu_k\}_{k\in\mathbb{N}}} \sup_{(\Delta_A,\Delta_B)} \quad & \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^{\infty} \alpha^k \left(x_k^\top Q x_k + u_k^\top R u_k\right)\right], \\
\text{s.t.} \quad & x_{k+1} = (\widehat{A} + \Delta_A)x_k + (\widehat{B} + \Delta_B)u_k + v_k, \\
& v_k \overset{i.i.d.}{\sim} \mathcal{P}(0, \Sigma_v), \quad x_0 \sim \mathcal{P}(0, \Sigma_0), \\
& u_k = \mu_k(x_k), \quad (\Delta_A, \Delta_B) \in \mathbb{A}.
\end{aligned} \tag{2-1.2}$$

The main objective of this work is to study the full uncertainty set $\mathbb{A}$ coming from game theory in relation to common uncertainty sets. To this end, we first proceed with an assumption and several important definitions that clarify and facilitate the presentation of the results in the later stages.

Assumption 2-1.1 (Linear time-invariant policy)**.** *In the problem* (2-1.2)*, we restrict the class of control policies μ_k to linear time-invariant controllers $\mu_k(x) = Kx$ where $K \in \mathbb{R}^{m\times n}$ is a constant matrix for all $k \in \mathbb{N}$, referred to as a* "feedback"*- or* "control" *gain.*

To continue, define a shorthand notation for the discounted LQR cost function as follows.

Definition 2-1.2 (Discounted LQ cost)**.** *Consider the dynamical system $x_{k+1} = Ax_k + v_k$ where the noise process and the initial 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 introduce the linear quadratic (LQ) cost function $\mathcal{J} : \mathbb{R}^{n\times n} \times \mathcal{S}^n_+ \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ as*

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

Since we consider a *discounted* LQ cost, it is helpful to also introduce a respective notion of stability. In fact, conventional stability is not necessarily required for the LQ cost to take finite values.

Definition 2-1.3 ($\sqrt{\alpha}$-stability (*cf.* Def. 7.6.1 [HS07]))**.** *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. So exponential stability $\implies$ $\sqrt{\alpha}$-stability, but not the other way around.*

2-2 Meaningful Uncertainty Sets

In this section we discuss what this $\mathbb{A}$ could actually be. With our finite-sample framework in mind, we deliberately chose for a time-domain, and especially *system-theoretic*, viewpoint of Robust Control.

[1] Without loss of generality, we omit cross terms in the cost, *e.g.*, of the form $x_k^\top S u_k$, to keep notation simple.

But first, it is the authors opinion that the community has a slight, yet increasing, problem with using pop-terminology. Linear regression is all of a sudden "*supervised learning*" and PID control would probably qualify as an "*A.I.*" to some.

Unfortunately, the abundance of buzzwords has also manifested itself in the exorbitant use of the words "*robust*" and "*optimal*". The problem is that classifying something as either robust or optimal is completely meaningless if it is not specified *with respect to what* we have this idealized performance. The statement that controller K' is "more robust" than controller K is therefore absolute nonsense, for example, any linear controller is robust against a well-defined set of model perturbations, in our case: $\{(\Delta_A, \Delta_B) \ : \ \sqrt{\alpha}\rho(A + \Delta_A + (B + \Delta_B)K) < 1\}$. However, in practice one is usually concerned with a smaller (and closed) set of uncertainties. In the next section we shed some light on what these errors might be.

2-2-1 Existence and Uniqueness of Solutions

As the first example will show, Assumption 2-1.1 restricts possible $\mathbb{\Delta}$. There is no time-invariant K which can stabilize all stabilizable pairs (A, B):

Example 2-2.1 (The lack of a universal linear control law). *Consider for some finite scalar c and $d \in (-1, 1)$ the matrices*

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

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

Example 2-2.1 can be interpreted in the spirit of *switching control*, *i.e.*, once (A_1, B) switches to system (A_2, B) your linear control law should switch as well. To aid the discussion we use the next Lemma:

Lemma 2-2.2 (Discrete-time version of Lemma 3.1 from [FB86]). *Let $\Sigma_{n,m}^{(\min)}$ be the set of minimal realizations (A, B) with $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$. We consider the intersection of some compact subset $\mathcal{K} \subset \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$ with $\Sigma_{n,m}^{(\min)}$, simply denoted $\Sigma^\star$. Let $\beta \in (0, 1)$ be some fixed decay rate. Then,*

(i) there exist compact subsets Σ_i, such that for a finite f we have $\bigcup_{i=1}^{f} \Sigma_i = \Sigma^\star$;

(ii) moreover, there exists a scalar $c \in (0, 1 - \beta)$ such that $\forall i \in \{1, \ldots, f\}$ we have $\rho(A + BK_i) \leq (\beta + c)$ for $(A, B) \in \Sigma_i$ and a corresponding set of control gains $\{K_i\}_{i=1}^{f}$.

Proof. Given some $\sigma \in \Sigma^\star$ and sufficiently small $\epsilon > 0$ and let $K_\sigma : \rho(A + BK_\sigma) < \beta + \epsilon < 1$. By continuity of the eigenvalues in (A, B) there is some neighbourhood U_σ around σ such that all systems in U_σ stabilized by K_σ have their eigenvalues strictly bounded by $\beta + \epsilon$. Thus we can find an open-cover of $\Sigma^\star$, *i.e.*, $\Sigma^\star \subset \bigcup_\sigma U_\sigma$. However, since $\Sigma^\star$ is compact we can find a finite sub-cover, *i.e.*, for some f we have $\Sigma^\star \subset \bigcup_{i=1}^{f} U_{\sigma_i}$. Then the corresponding set of control gains K_i is the finite set from part (ii), whereas $\Sigma_i = \overline{U}_{\sigma_i}$ are the corresponding compact sets from part (i). Letting $c = \text{argmax}_i\{\beta + \epsilon_i\}$ concludes the proof. $\square$

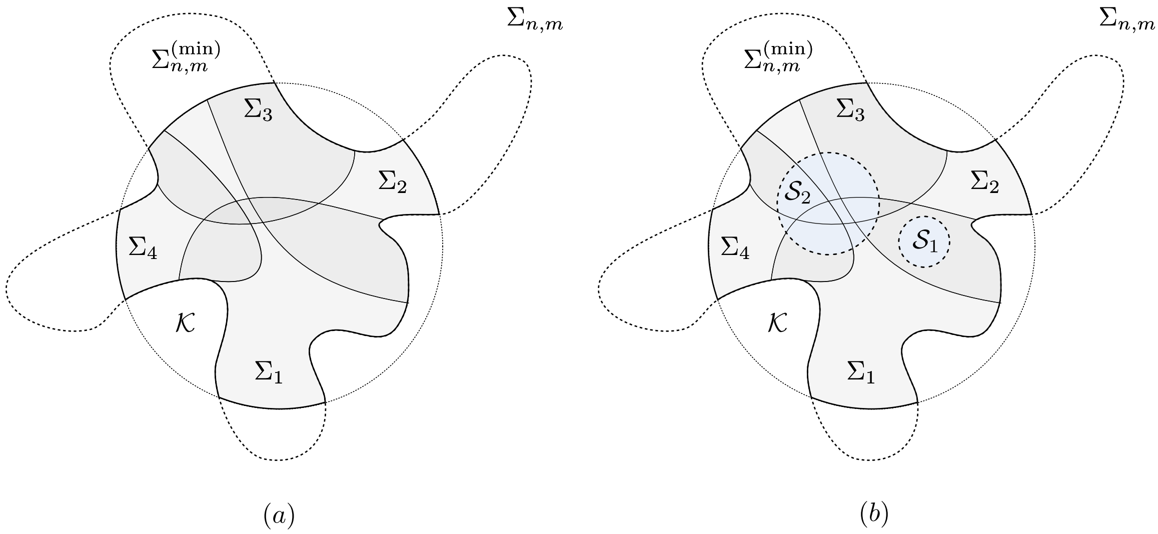


Figure 2-1: *(a) Schematic overview of Lemma 2-2.2. (b) for $\mathcal{S}_1$, both K_1 and K_2 can stabilize the set, while $\mathcal{S}_2$ cannot be stabilized completely by any of the four controllers.*

This segmentation of $\Sigma^\star$ is visualized in part (a) of Figure 2-1.

As indicated by Lemma 2-2.2, given a compact subset $\mathcal{K}$ intersected with the set of stabilizable pairs (A, B) one can indeed introduce a finite covering where all the elements of each segment can be stabilized via a common feedback gain, *e.g.,* (A_1, B) and (A_2, B) from Example 2-2.1 are never members of the same segment while for example $\|A_1\|_2 = \|A_2\|_2$. This simple observation indicates that the ***existence*** of a solution to (2-1.2) is not immediately obvious, even for simple norm-balls. Think of the set $\mathcal{S}_2$ in Figure 2-1 (b), controllers K_2, K_3, K_4 can all stabilize a part of $\mathcal{S}_2$, but not the entire set. In other words, (2-1.2) with $\mathbb{\Delta} = \mathcal{S}_2$ would be infeasible[2]. Thus, Lemma 2-2.2 tells us that our nominal system plus our uncertainty set should be contained in some Σ_i for the problem to be well defined. It turns out that this characterization, *i.e.,* existence conditions for robust controllers, are pretty much an open problem to the point that Ackermann even argues that "*the design of a robust controller is more of an art than a science [Ack02, p.75].*"

Regarding ***uniqueness***, think of the set $\mathcal{S}_1$ in Figure 2-1, both K_1 and K_2 can stabilize the set of models. To order the set of controllers, a cost function can be introduced, after which the *least expensive* control law is implemented. However, usually, we have no knowledge about stability, but merely about the induced cost. Is optimality enough? Consider a scalar example; we would like to find $\operatorname{argmin}_{\{u_k\}_{k=0}^{\infty}} \sum_{k=0}^{\infty} u_k^2$ constrained by $x_{k+1} = a x_k + u_k$, $|a| > 1$. This is a classic problem, minimizing the energy, which sounds like it must be related to stability. The solution is obviously $u_k^\star = 0 \; \forall k$, but the resulting closed-loop trajectory is unstable! So yes, you can optimize almost everything, but the corresponding interpretation is not included. One of the nice properties of the family of LQ control problems is that the cost, and qualitative properties like stability, *can* be linked.

[2]Strictly speaking, there might be another cover with some K_i which *does* stabilize the full set $\mathcal{S}_2$. Figure 2-1.(b) merely visualizes what we mean with $\mathbb{\Delta}$ being *too* big, think of Example 2-2.1.

2-2-2 Convexity in Robust Control

Towards exact and tractably solving (2-1.2) one might hope that $\mathbb{A}$ is convex? So, as a seemingly natural next question, one may wonder how these individual segments look like, and in particular, whether a set of stabilizable pairs with a common stabilizing feedback is necessarily convex in $\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times m}$. The following example provides a negative answer to this question.

Example 2-2.3 (Non-convex segment). *Consider for $a = 2$ and $d = 0.5$ the matrices*

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

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

These quick examples indicate that convex uncertainty sets for (A, B) in $\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times m}$ can be a restrictive point of view indeed and may be potentially ***conservative***. To be clear on what is meant, consider Figure 2-2. There are two viewpoints; (a) *convexifying*: the uncertainty set $\mathbb{A}$ is like X_1 non-convex, which is natural from a system-theoretic point of view, while the robust framework only allows for convex sets, hence, an outer approximation (C_o) comes into place. Secondly, (b) *non-convex domain of LQR*: say the uncertainty set $\mathbb{A}$ originates from statistics and is actually convex like C_i in (b). Now, feasibility might be only asserted for small – inscribed – sets like C_i instead of for example X_2. Of course, in (a), this larger set C_0 must itself be feasible again, which is indeed not easily satisfied due to the non-convex nature of the LQR domain in (A, B), *i.e.,* C_o must be again some inscribed set, like C_i, in some possibly non-convex set X_2.

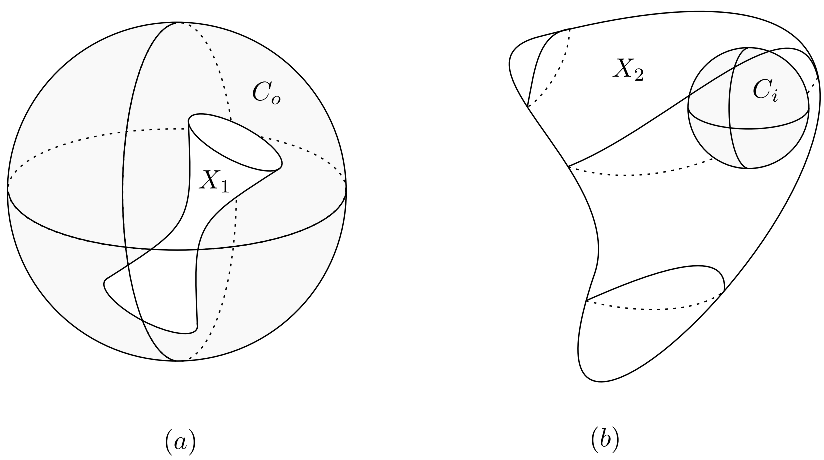


Figure 2-2: Convex out- and inscribed uncertainty sets.

See section 3-3-1 for a further discussion on the conservativeness of ellipsoids, specifically, or see Figure 3.7 in [BPT12] for a full and explicit example. Now, note that we do not claim that non-convexity is desirable, but merely observe that it should not be ruled out for element-wise uncertainties in the pair (A, B). This last part is important, since using functional multiplier theory [Lue69], the LQR problem can be posed as a convex optimization problem, yet in ℓ_2^{∞}, not $\mathbb{R}^{m\times n}$. One can even formulate the LQR problem as an infinite-dimensional LP [HLL96, Bar02]. Of course, the pitfall is that although the *lifting* made the problem convex, we will lose exactness, and usually tractability, when trying to solve the problem in these settings.

Therefore, we keep working in the more computationally friendly domain $\mathbb{R}^{n\times n}\times\mathbb{R}^{n\times m}$. Even there, one might say that the $\|\cdot\|_2$-norm is related to stability, and indeed is a convex function. This is true, but rather a crude approximation since $\rho(A) \leq \|A\|_2 \; \forall A \in \mathbb{R}^{n\times n}$. For example, take

$$A + BK = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}, \quad a \in (-1, 1), \quad b \in \mathbb{R}.$$

Then we always have $\rho(A+BK) < 1$ while to make $\|A+BK\|_2 < 1$ we need $a < \sqrt{1-b^2}$. This implies that the operator-norm approach to stability can be rather conservative. Nevertheless, one might need to consider a trade-off here since these norm-balls are the prevelant uncertainty set in non-asymptotic system identification, see [MT19] for a review.

Similarly, for some given D, E, we often see uncertainty sets of the form: $\{\Delta_A : \Delta_A = DFE\}$ for $\{F : F^\top F \preceq I\}$, or $\{F : \|F\| \leq 1\}$, which are both convex[3]. These kind of sets usually appear out of academic interest, or as put in [Gev05] , they are *God-given*, being “easy” to solve over (see chapter 3).

Another form one often encounters, which is especially simple in continuous-time, relates to Lyapunov equations (in linear control). If we can find a $P \succ 0$ such that $A_i^\top P + A_i P \prec 0 \; \forall i \in \mathcal{I}$ then any A in the convex polytope with vertices $\{A_i\}_{i\in\mathcal{I}}$ is Hurwitz. However, existence of such a P, given some set of vertices is not straight-forward, in part due to the non-convexity of the set of stable (Hurwitz) matrices, *i.e.,* the polytope might be too big.

In conclusion, convexity does not naturally appear in our system-theoretic setting, it is rather imposed towards a *tractable* algorithm. So, should we consider non-convex optimization instead? To answer this question, let us simply quote Maryam Fazel[4]: “*A lot of problems that arise in Machine Learning are not convex. So people nowadays are talking a lot about non-convex optimization but the problem is, saying something is non convex, is not saying anything about it*”. Unfortunately, her remark might be underappreciated. We will adhere to her comment and indeed look for more structure.

[3] The norm-case is easy to see, for first case, write $F = \sum_i x_i F_i$, for F_i being a basis element of F. Then, using Schur complements, $F^\top F \preceq I$ is equivalent to

$$\begin{pmatrix} I & \sum_i x_i F_i^\top \\ \sum_i x_i F_i & I \end{pmatrix} \succeq 0.$$

This LMI shows convexity of the uncertainty set (*cf.* [BGFB94]).

[4] https://www.youtube.com/watch?v=uyZOcUDhIbY

2-2-3 Semi-Algebraic Sets in Robust Control

In (2-1.2) we let (A, B) be unknown, but towards a practically relevant optimization problem we do make a few structural assumptions. We assume that we know the dimensions n and m, but also that at least there exists a linear controller which can $\sqrt{\alpha}$-stabilize the pair (A, B). In other words, we assume that $(\sqrt{\alpha}A, B)$ is ***stabilizable***. Looking back at (2-1.2), when $0 \in \mathbb{\Delta}$, which is usually the case, then a natural implication is that $(\sqrt{\alpha}\widehat{A}, \widehat{B})$ must be stabilizable as well. Therefore, it is informative to study the space of stabilizable pairs (A, B), or subsets of it, since we might optimize over this space and not simply some product of $\mathbb{R}$.

The space of controllable pairs (A, B) is, under the standard topology, dense in $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$. To someone not familiar with control this may seem as if controllability is almost surely satisfied. Of course, the catch is that, say you have no input, then yes, you are arbitrary close to a controllable model, but you remain stuck without any input. It does not magically (nor statistically) appear. Thus, this denseness remark is not the most informative. Instead we state a classical result from [HM77], which hinges on the polynomial nature of the determinant.

Lemma 2-2.4 (Theorem 4.4 [HM77]). *The set $\mathcal{C} = \{(A, B) \ : \ A$ is not cyclic or (A, B) is not controllable$\}$ is **Zariski closed**.*

Without going into too much details, let $p_j(x)$ be a polynomial, then a set X is Zariski closed when it can be written as $X = \bigcup_{i=1}^{n} S_i$, $S_i = \{x \in \mathbb{C}^n \ : \ p_j(x) = 0, \ j \in J\}$, *i.e.*, if it is a finite union of algebraic sets.

Now, consider a linear discrete-time dynamical system $x_{k+1} = (\widehat{A} + \Delta_A)x_k + (\widehat{B} + \Delta_B)u_k$, where both Δ_A and Δ_B are unknown. The state of the art identification schemes can give probabilistic norm bounds of the form $\mathbb{P}\{\|\Delta_A\|_p \leq \varepsilon\} \geq 1 - \delta$. To control our unknown dynamical system we would like to find a controller which can handle all of these potential uncertainties, *i.e.*, find a robust controller. There is of course a constraint, such a robust control law exists only if $(\widehat{A} + \Delta_A, \widehat{B} + \Delta_B)$ is stabilizable for all (Δ_A, Δ_B). However, recall, that even then, the set of uncertainties might be too big for 1 controller to stabilize all the perturbations. Thus, we need stabilizability and a sufficiently small *ball*. Let $\mathcal{B}$ denote some ball around $(\widehat{A}, \widehat{B})$, then an object of interest is $\Sigma_{n,m}^{(\min)} \bigcap \mathcal{B}$, which is a ***semi-algebraic set***. Identification schemes give us the balls and we keep what is relevant.

Going from non-convex optimization to optimization over semi-algebraic sets might not seem like any improvement, since the field is still broad, but at least we have some structure and thereby a clear set of tools. This is especially interesting since this toolbox is only becoming larger [Las09, BPT12], also with respect to our purposes [HLS09].

2-2-4 Structure Preserving Worst-case Models

Most, if not all, robust optimal control laws are derived from optimization programs of the form

$$(u^\star, m^\star) = \arg \inf_{u\in\mathcal{U}} \sup_{m\in\mathcal{M}} c(u, m) \tag{2-2.1}$$

where $u^\star$ is the optimal input, $m^\star$ the worst-case model and c some real-valued cost function. When is the solution to (2-2.1) meaningful? Of course, when m should represent some model, say for a walking robot, then $m^\star$ should not be a model for an airplane. Differently put, the previous sections discussed $\mathbb{\Delta}$ mostly from an abstract point of view, but it should not be forgotten that in practice, this set is coming from somewhere.

One way to make $\mathcal{M}$ somewhat realistic is to constrain parameters to their physically possible limits, *e.g.,* strictly positive inertias in mechanical systems like in (3-4.24). For a lot of systems we do however have not such a clear parametric description. Then, one approach to constrain $\mathcal{M}$, which is borderline philosophical, is to let $\mathcal{M}$ only contain models being ***topologically equivalent***[5] to the real system. This approach puts more emphasis on the behaviour instead of the parametrization. We have to see if this is however computationally attractive as well. To give a short example, parametrize a scalar system by $a \in \mathbb{R}$, denoted $\sigma(a) : \{x_{k+1} = ax_k$. Now define two seemingly similar sets: $\sigma^+ := \{\sigma(a) \ : \ a \in (0,1)\}$, $\sigma^- := \{\sigma(a) \ : \ a \in (-1,0)\}$. Pick any $\sigma_1 \in \sigma^+$, $\sigma_2 \in \sigma^-$, then σ_1 and σ_2 are *not* topologically equivalent, one can think of a damper and a spring. The latter keeps flipping the state from one side of 0 to the other (see Figure 2-3). Since homeomorphims φ of the real-line are (strictly) monotone, we can not even

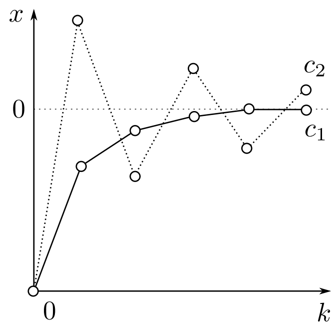


Figure 2-3: Given the systems σ_1 and σ_2 we show typical behaviour decipted by the curves c_1 and c_2 respectively.

find a homeomorphim between any of the systems trajectories of σ_1 and σ_2. The point is that these systems are therefore inherently different, it is not a matter of our arbitrary choice of coordinates. However, all systems within either σ^+ or σ^- are topologically equivalent, they capture the same *qualitative behaviour*.

2-2-5 In Conclusion, A Trade-off

In conclusion, we see that for problem 2-1.2 uniqueness and especially existence of a robust controller is largely an open problem. Moreover, we know that convex $\mathbb{\Delta}$ are unlikely to naturally appear in a system theoretic context. However, they do appear often in a statistical context, allowing for immediate finite-sample guarantees, which as mentioned before, will be inherently conservative. At last, we like to remark that the nominal-, real- and worst-case model should have some structural properties in common for the framework to make sense, which constrains potential $\mathbb{\Delta}$.

In the next chapter we set out to see if there is structurally nice semi-algebraic uncertainty set $\mathbb{\Delta}$ to exploit in order to alleviate predominant conservatism in robust control, possibly at the cost of more involved concentration inequalities.

[5]One can think of a (possibly non-linear) generalization of a similarity transform, see section B-2 for an introduction and section 3-3 for a further discussion.

Chapter 3

Game Theoretic Robust Control

The previous chapter introduced our search for structurally nice (semi-algebraic) uncertainty sets. It is hypothesized that since dynamic LQ games can be efficiently solved plus interpreted as robust control problems they might embody precisely what we are looking for.

Hence, this chapter is centered around quantifying the robustness resulting from a dynamic game with quadratic cost[1]. Early accounts of this idea can be found in for example the monograph by Whittle [Whi90, p.90], where the remark is made, in his notation, that extremizing a *risk-sensitive* optimal control cost function of the form $\mathbb{S} = \mathbb{C} + \theta^{-1}\mathbb{D}$ can be interpreted as extremizing $\mathbb{C}$ subject to $\mathbb{D} \leq d$ for some d. Which then can be understood as another constrained optimal control problem indeed.

There is a large body of work in this direction, especially the contribution by one of its pioneers, Ian Petersen, is very close to our approach. It can be argued that this line of work started in the continuous-time with [PH86, Pet87], where a constructive, yet heuristic algorithm was proposed to find a stabilizing controller when one has an uncertain system matrix of the form $(A + DF(t)E)$, $F(t)^{\top}F(t) \preceq I$.

A few years later the celebrated paper [KPZ90] gave necessary and sufficient conditions for the continuous-time system $\dot{x}(t) = (A+\Delta_A(t))x(t)+(B+\Delta_B(t))u(t)$, $(\Delta_A \;\; \Delta_B) = DF(t)(E_1 \;\; E_2)$, $\|F(t)\| \leq 1$ to be quadratically stabilizable, plus they further clarified the link with $\mathcal{H}_\infty$-control. This result was later generalized to the discrete-time case in [GBA94]. 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, although currently driven by measure concentration results, *e.g.,* see [AL18, SMT$^+$18].

In the stochastic case, distributional uncertainties in the form of *relative entropy* constraints are considered, see the monographs [HS07, RPUS00].

Although these problems are well understood, the catch is that usually the set defined by $\mathbb{D} \leq d$ (in the notation of Whittle) depends on the the extremizing parameters. Therefore it

[1] Or equivalently a linear Gaussian optimal control problem with quadratic cost, identity covariance and exponential utility function, *i.e.* LEQR [Jac73]

is not clear *a priori*, over which set the robust control problem is solved, this is effectively only known *a posteriori*. Moreover, again in the notation of Whittle, most results are not constructive, only showing existence of pairs (θ, d) or rely on heuristics. Furthermore, most authors do not consider the full set defined by $\mathbb{D} \leq d$, but generally some "*inscribed ball*". See for example [RPUS00, ch.10], which gives a simple example of how to go about fitting these kind of ellipsoids to data. What is not clear, is why, from a system theoretic point of view, this should be an ellipsoid in the first place?

Motivated by renewed interest in tractable reformulations of robust LQR problems (*cf.* [AYS11, DMM⁺17, US18, CKM19, Tu19]), we investigate which lessons can be drawn from the readily available dynamic game theory, with emphasis on structural properties like in the recent work [FGKM18, MR18, MRJS19].

We start by introducing a new uncertainty set.

3-1 Introduction of a New Uncertainty Set

Recall the uncertainties as specified in (2-1.2) and assume B to be known. For now, think of $\widehat{A}$ as some nominal closed-loop system, then the key definition of this work is the following set:

Definition 3-1.1 (A new confidence and uncertainty set). *Given a tuple $(\widehat{A}, D, \Sigma_0, \Sigma_v, \alpha)$ and some $\gamma \in \mathbb{R}_{\geq 0}$ define a set of models around $\widehat{A}$ by the* ***confidence set****:*

$$\mathcal{A}_\gamma(\widehat{A}) := \left\{ A \in \mathbb{R}^{n\times n} \;:\; \begin{array}{l} A = \widehat{A} + D\Delta_A \\ \Sigma_x = \alpha A \Sigma_x A^\top + \Sigma_0 + \alpha(1-\alpha)^{-1}\Sigma_v, \quad \Sigma_x \succ 0 \\ \left\langle \Delta_A^\top \Delta_A, \Sigma_x \right\rangle \leq \gamma \end{array} \right\}. \tag{3-1.1}$$

Key in the definition of $\mathcal{A}_\gamma(\widehat{A})$ are the uncertainties Δ_A. Since this ***uncertainty set*** *will appear again later, it has its own notation: $\mathbb{\Delta}_\gamma$, defined by $\mathcal{A}_\gamma(\widehat{A}) = \widehat{A} + D\mathbb{\Delta}_\gamma(\widehat{A})$[2].*

Some readers might recognize a set reminiscent of the uncertainty in [AYS11], more on that in Section 3-3-1. In the remainder of the text we will occasionally take $W := \Sigma_0 + \alpha(1-\alpha)^{-1}\Sigma_v \succ 0$. We could relax this and only demand $\Sigma_x \succeq 0$, but then part of our uncertainty set can be unbounded, corresponding to the part of state-space which is never excited. This is of course not insightful at all. Therefore, we avoid these pathological examples[3]. Moreover, it is important to realize that Σ_x in (3-1.1) is not a constant, indeed some authors would rather write $\Sigma_x = \mathsf{dlyap}(\sqrt{\alpha}A, W)$ to highlight the explicit dependence.

Remark 3-1.2 (Ball structure). *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, but depends on the center $\widehat{A}$ as visualized in Figure 3-1.*

[2] Where $+$ is now overloaded and refers to the Minkowski sum: $A + B = \{a + b \;:\; a \in A, b \in B\}$.

[3] To illustrate this remark, take $\alpha \in (0,1)$, $A = 0_{2\times 2}$, $\Sigma_0 = \Sigma_v = DD^\top$ for $D^\top = \begin{pmatrix} 1 & 0 \end{pmatrix}$ such that $\Delta_A = \begin{pmatrix} a & b \end{pmatrix}$ for some scalars a and b. The parameter b is clearly unbounded, but totally useless since the second state is identically 0.

Moreover, let $\Sigma_0 + \alpha(1-\alpha)^{-1}\Sigma_v \succ 0$*. Then for* Δ_A *to be in* $\mathbb{\Delta}_\gamma(\widehat{A})$ *is the same as being part of the metric ball* $\{\Delta_A \in \mathbb{R}^{d\times n} : \|\Delta_A^\top\|^2_{F,\Sigma_x} \leq \gamma\}$ *for* Σ_x *as in (3-1.1). This further explains why* γ *is referred to as a "radius".*

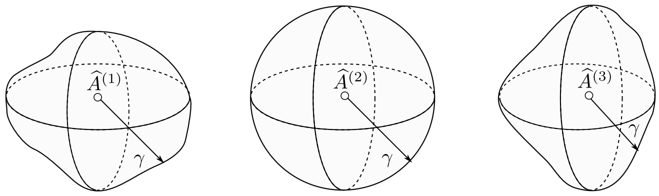


Figure 3-1: The set (3-1.1) can be interpreted as some ball around $\widehat{A}^{(i)}$. However, for a fixed γ the shape of $\mathcal{A}_\gamma(\widehat{A}^{(i)})$ depends on its center $\widehat{A}^{(i)}$.

Remark 3-1.3 (Structural information)**.** *The matrix* D *in Definition 3-1.1 may be used to incorporate a special form of prior structural information into the uncertainty set. For instance, when it is known that all entries of a particular column of* A*, say the* j^{th} *column* $[A]_j$*, are subject to the same level of uncertainty, one can choose* $[D]_j = \mathbb{1}_{n\times 1}$*. On the other hand, without any prior structural information, one should choose* $D = I_n$*.*

To gain a better visual understanding of our uncertainty set, we do a quick example:

Example 3-1.4 (2D $\mathcal{A}_\gamma$)**.** *Let the problem parameters be given by*

$$\widehat{A} = \begin{pmatrix} 0.8 & 0.5 \\ 0 & 0.8 \end{pmatrix}, \quad D = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \Sigma_0 + \alpha(1-\alpha)^{-1}\Sigma_v = I_2$$

and $\alpha = 0.9$*. In Figure (3-2a) we show* $\mathbb{\Delta}_\gamma(\widehat{A})$ *for several* $\gamma \in \mathbb{R}_{>0}$*. It is observed that we have* $\mathbb{\Delta}_\gamma \subseteq \{\Delta \in \mathbb{R}^{1\times 2} : \|\Delta^\top\|^2_{F,W} \leq \gamma\}$ *since* $\Sigma_x \succeq W$*. Then if* γ *is finite,* $W \succ 0 \implies \mathbb{\Delta}_\gamma$ *is contained in a compact set. For* $\gamma = 1$ *such a ball is also included in Figure (3-2a). Furthermore, we can plot for* $Q = I_2$ *the cost* $\mathcal{J}(\widehat{A} + D\Delta_A, Q)$ $\forall \Delta_A \in \mathbb{\Delta}_\gamma(\widehat{A})$ *as is done in Figure (3-2b).*

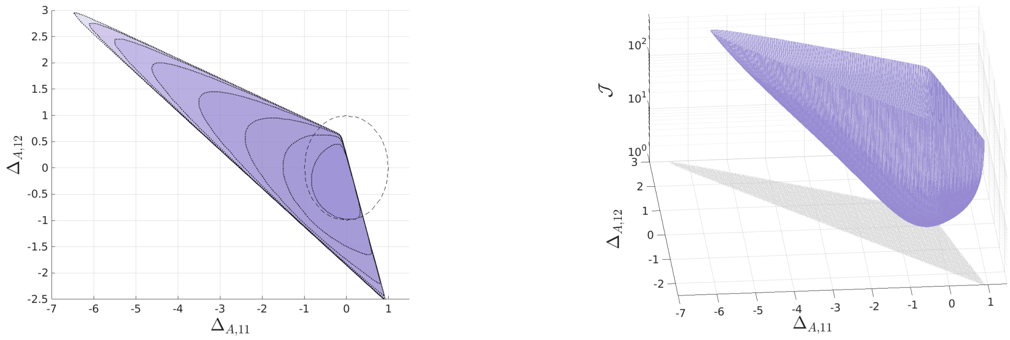
Given the family of confidence sets as in Definition 3-1.1. Then, under Assumption 2-1.1 one can deploy the shorthand notation of Definition 2-1.2 and equivalently describe a RLQR problem (2-1.2) with an uncertain A, yet known B, through the static minimax optimization program

$$\inf_{K\in\mathbb{R}^{n\times m}} \sup_{A_{c\ell}\in\mathcal{A}_\gamma(\widehat{A}+BK)} \mathcal{J}(A_{c\ell}, Q + K^\top RK), \tag{3-1.2}$$

which is our starting point. It is worth noting that the inner maximization step depends on K and that problem (3-1.2) is well-defined for all $\gamma \in \mathbb{R}_{\geq 0}$. In section 3-2-1 we elaborate on how the framework can be extended to deal with uncertainties in B as well.

3-2 Solving a Robust LQR Problem

The main objective of this section is to provide a closed-form solution to the RLQR problem (3-1.2) and study its implications. In particular, we show that (3-1.2) can be interpreted



(a) Uncertainty set from Example 3-1.4 **(b)** LQ cost $\mathcal{J}$ over $\mathbb{A}_\gamma(\widehat{A}_{c\ell})$

Figure 3-2: Given the parameters from Example (3-1.4) in (a) we show $\mathbb{A}_\gamma(\widehat{A})$ for several $\gamma \in [2^0, 2^{14}]$. See that we converge to this particular triangular shape. The plot also includes the outerball $\{\Delta_A : \|\Delta_A^\top\|_{F,W}^2 \leq \gamma = 1\}$ as visualized by the dashed line. Then in (b) we show the quadratic cost over $\mathbb{A}_\gamma(\widehat{A})$ for $\gamma = 2^{14}$. Observe that the cost grows sharply towards the boundary of $\mathbb{A}_\gamma(\widehat{A})$.

as the constrained optimal control version of a dynamic game. To reach to this end, a first step entails evaluating the worst-case LQ cost, *i.e.*, the inner maximization in (3-1.2) for a given controller K. Before targeting this objective, inspired by Lemma 2 from [FGKM18], we provide some insights about the uncertainty set $\mathbb{A}_\gamma$, which are especially interesting from an optimization point of view

Proposition 3-2.1. *The set $\mathbb{A}_\gamma(\widehat{A})$ as defined in Definition 3-1.1 has the following properties:*

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

(ii) For $\gamma > 0$, the set $\mathbb{A}_\gamma(\widehat{A})$ is semi-algebraic, thereby a disjoint union of a finite number of connected semi-algebraic sets[4].

The fact that our uncertainty set is semi-algebraic and does not rule out the lack of convexity is nice from a control theoretic point of view (see section 2-2-3). Moreover, the semi-algebraic nature of the set implies that $\partial\mathbb{A}_\gamma$ is smooth almost everywhere.

The proof of Proposition 3-2.1 is split up in two parts.

Proof of Proposition 3-2.1 (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

[4]After extending the tools from [FGKM18] to the game theoretic regime, we can even show that the set is path-connected, which is done in chapter 4, Corollary 4-3-1.

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 admissable uncertainties in A might live in a non-convex set. □

The set (3-1.1) has another interesting property indeed

Proof of Proposition 3-2.1 (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 $X=\mathrm{mat}(\mathrm{vec}(X))$. Then for $Y:=\mathrm{mat}(Z^{-1}\mathrm{vec}(W))$ the set $\mathbb{A}_\gamma(\widehat{A}_{c\ell})\subset\mathbb{R}^{n\times n}$ can be written as

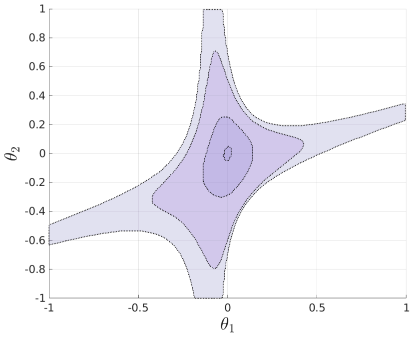
$$\left\{\Delta_{A_{c\ell}} : \begin{matrix} 0\le \mathrm{vec}^\top(I_n)(\Delta_{A_{c\ell}}\otimes\Delta_{A_{c\ell}})Z^{-1}\mathrm{vec}(W)\le\gamma \\ 0<\det(Y_i), \quad i=1,\dots,n\end{matrix}\right\} \tag{3-2.1}$$

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 $A_{c\ell}$ by enforcing $Z\succ 0$, see *e.g.* Theorem 7.2.5 in [HJ90]. 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 (3-2.1) is indeed semi-algebraic for $\gamma>0$, thus a set of polynomial inequalities in the elements of $\Delta_{A_{c\ell}}$ of the form $\mathcal{S}$:

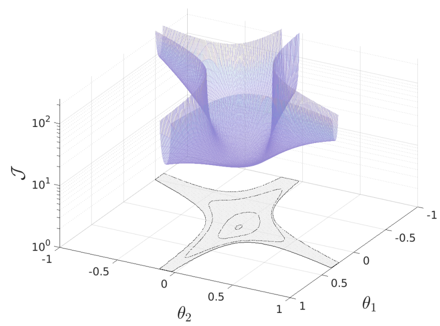
$$\mathcal{S}=\left\{\Delta_{A_{c\ell}}\in\mathbb{R}^{d\times n} \;:\; 0\le p_1(\Delta_A),\ 0\le\gamma p_2(\Delta_A)-p_1(\Delta_A),\ 0<p_i(\Delta_A),\quad i=3,\dots,3+n\right\}.$$

This result is of course closely related to the prominent role played by polynomials in linear control theory. The second part follows directly from the fact that $\mathcal{A}_\gamma$ is semi-algebraic and Theorem 5.19 in [BPR16]. □

The potential lack of convexity can be exemplified. The idea behind this next example is the fact that the line connecting two $\sqrt{\alpha}$-stable matrices need not be necessarily $\sqrt{\alpha}$-stable (*cf.* Definition 2-1.3). So although Example 3-1.4 displays a convex set, this is not generally the case.



(a) Uncertainty set $\mathbb{A}_\gamma$ as in Definition 3-1.1.



(b) LQ cost $\mathcal{J}$ as in Definition 2-1.2.

Figure 3-3: Given the parameters from Example 3-2.2 we show the uncertainty sets and LQ function identified by (θ_1, θ_2) for different levels $\gamma \in \{2^{-4}, 2^1, 2^4, 2^{14}\}$ from darker to lighter gray.

Example 3-2.2 (Non-convexity, Proposition 3-2.1 (i))**.** *We consider the case where the confidence set $\mathcal{A}_\gamma \subset \mathbb{R}^{3\times 3}$ from Definition 3-1.1 is constructed using the parameters $\alpha = 0.95$, $D = Q_{c\ell} = \Sigma_0 = I_3$, and $\Sigma_v = 0.01 I_3$ where $Q_{c\ell} \triangleq Q + K^\top RK$ for some (Q, R, K). Here we will consider several "levels sets" of $\mathbb{A}_\gamma$. Since the set $\mathbb{A}_\gamma$ is essentially a 9-dimensional object, for the sake of illustration we ought to restrict our attention to a 2-dimensional subset of it. For this purpose, we consider the closed loop matrix $A_{c\ell}$, and especially all Δ_A, to be parametrized by*

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

where $\Delta_{A13} = 4.98\theta_1 - 0.25\theta_2$, $\Delta_{A31} = 0.45\theta_2 - 1.08\theta_1$, and the parameters[5] (θ_1, θ_2) belong to the interval $[-1, 1]^2$. Figure 3-3a depicts the 2-dimensional slice of $\mathbb{A}_\gamma$ by means of (θ_1, θ_2) for the levels: $\gamma \in \{2^{-4}, 2^1, 2^4, 2^{14}\}$. Interestingly enough, it is non-convex for large values of γ. Figure 3-3b also illustrates the LQ cost $\mathcal{J}$ in Definition 2-1.2 when $(\theta_1, \theta_2) \in [-1, 1]^2$. It is worth noting that the function $\mathcal{J}(A_{c\ell}, Q_{c\ell})$ takes $+\infty$ when $\Delta_A \notin \bigcup_{\gamma\in\mathbb{R}_{\geq 0}} \mathbb{A}_\gamma$.

Since the proof of Proposition 3-2.1 (ii) is constructive we *could* do an explicit example regarding the polynomial parametrization of $\mathbb{A}_\gamma$ as well. The reason why we do not provide a printed expression for such a parametrization is well-known: "*a derivation of the uncertain polynomial governing stability involves so much algebra that only a masochist would attempt it by hand.*" [Bar94, p.29].

In the next step, we tackle the worst-case LQ problem over $\mathcal{A}_\gamma$, as the inner maximization of the RLQR problem (3-1.2). This problem is defined as

$$\sup_{A_{c\ell}\in\mathcal{A}_\gamma(\widehat{A})} \mathcal{J}(A_{c\ell}, Q_{c\ell}), \tag{3-2.2}$$

[5]This choice of (θ_1, θ_2) over $(\Delta_{A13}, \Delta_{A31})$ is purely driven by visualization purposes.

for some given $Q_{c\ell} \succeq 0$ $(Q + K^\top RK)$ and $\sqrt{\alpha}$-stable $\widehat{A}_{c\ell}$ $(\widehat{A} + BK)$. Denote the solution to (3-2.2) by $A^\star_{c\ell}(\gamma) := \widehat{A}_{c\ell} + D\Delta^\star_A(\gamma)$.

Proposition 3-2.3 (Worst-case LQ cost). *Consider the problem (3-2.2) with the $\sqrt{\alpha}$-stable nominal closed-loop system matrix $\widehat{A}_{c\ell}$, the structural information matrix D, some $\alpha \in (0,1)$, the initial data $\Sigma_0, \Sigma_v \in \mathcal{S}^n_{++}$, and the closed-loop cost matrix $Q_{c\ell} \in \mathcal{S}^n_+$. Given some $\delta \in \mathbb{R}_{\geq 0}$, let us assume that $(\delta^{-1} I_d - \alpha D^\top SD) \succ 0$ is satisfied for the symmetric (minimal) positive semi-definite solution S to the algebraic equation*

$$S = Q_{c\ell} + \alpha \widehat{A}^\top_{c\ell} S \widehat{A}_{c\ell} + \alpha^2 \widehat{A}^\top_{c\ell} SD(\delta^{-1} I_d - \alpha D^\top SD)^{-1} D^\top S \widehat{A}_{c\ell}.$$

Then define

$$\Delta^\star_A(\delta) = \alpha(\delta^{-1} I_d - \alpha D^\top SD)^{-1} D^\top S \widehat{A}_{c\ell}. \tag{3-2.3}$$

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

$$\widetilde{\Sigma}_x = \alpha\big(\widehat{A}_{c\ell} + D\Delta^\star_A(\delta)\big)\widetilde{\Sigma}_x\big(\widehat{A}_{c\ell} + D\Delta^\star_A(\delta)\big)^\top + W \tag{3-2.4}$$

which in its turn defines the function

$$\widetilde{h}(\delta) = \left\langle \big(\Delta^\star_A(\delta)\big)^\top \Delta^\star_A(\delta), \widetilde{\Sigma}_x \right\rangle. \tag{3-2.5}$$

Then, $\Delta^\star_A(\gamma) = \Delta^\star_A(\delta)$ and $\mathcal{J}^\star = \langle \widetilde{\Sigma}_x, Q_{c\ell} \rangle$ are the optimizer (worst-case uncertainty) and the optimal value of the problem (3-2.2) with the parameter $\gamma = \widetilde{h}(\delta)$.

Proof of Lemma 3-2.3. Consider the problem

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

If γ satisfies $h(\delta) = \gamma$ then from the Lemma A-0.1 the solution to $\mathcal{P}_a(\gamma)$ can be directly retrieved from the (negated) problem

$$\mathcal{P}_b(\delta) : \begin{cases} \underset{\Delta_{A_{c\ell}} \in \mathbb{R}^{d\times n}}{\operatorname{argmin}} & \underset{x_0, v}{\mathbb{E}} \left[\sum_{k=0}^{\infty} \alpha^k \left(\delta^{-1} w_k^\top w_k - x_k^\top Q_{c\ell} x_k \right) \right] \\ \text{subject to} & x_{k+1} = \widehat{A}_{c\ell} x_k + D w_k + v_k, \\ & v_k \overset{i.i.d.}{\sim} \mathcal{P}(0, \Sigma_v),\ x_0 \sim \mathcal{P}(0, \Sigma_0), \\ & w_k = \Delta_{A_{c\ell}} x_k. \end{cases} \tag{3-2.6}$$

Under the conditions from Proposition 3-2.3 the program $\mathcal{P}_b(\delta)$ can be solved using Dynamic Programming, *e.g.* see chapter 3 from [Ber07], regarding feasibility one can always select $w_k = 0\ \forall k$, moreover $(\delta^{-1} I_d - \alpha D^\top SD) \succ 0$ 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 (3-2.6), under a policy $\nu := \{w_0, w_1, \ldots\}$ be parameterized by $V^\nu(x) = -x^\top Sx + q$, $S \in \mathcal{S}^n_+$, $q \in \mathbb{R}$. 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 (3-2.6)

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

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

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

where

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

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

Now we are at the stage to present one of the main results.

Theorem 3-2.4 (RLQR solution under $\mathcal{A}_\gamma$ (see Definition 3-1.1)). *Consider the RLQR problem (3-1.2) with the nominal $\sqrt{\alpha}$-stabilizable model $(\widehat{A}, B)$, the structural information matrix D, $\alpha \in (0,1)$, the cost matrices $Q \succeq 0, R \succ 0$ and the covariance matrices $\Sigma_v, \Sigma_0 \in \mathcal{S}^n_{++}$. Given the parameter $\delta \in \mathbb{R}_{\geq 0}$, assume that the algebraic equation*

$$P = Q + \alpha\widehat{A}^\top P\Lambda^{-1}\widehat{A}, \quad \Lambda := I_n + \alpha(BR^{-1}B^\top - \delta DD^\top)P.$$

in P admits a symmetric minimal[6] *positive semi-definite solution denoted $P(\delta)$ and define $\Lambda(\delta)$ correspondingly. Furthermore, define*

$$\Delta^\star_A(\delta) = \alpha\delta D^\top P(\delta)(\Lambda(\delta))^{-1}\widehat{A}. \tag{3-2.7}$$

Consider the expressions for $\widetilde{\Sigma}_x$ and $\widetilde{h}(\delta)$ as in (3-2.4) and (3-2.5) respectively, which are now functions of K as well, to emphasize the difference, the tildes are dropped, i.e., *define:*

$$\Sigma_x = \alpha(\widehat{A} + D\Delta^\star_A(\delta) + BK^\star(\gamma))\Sigma_x(\widehat{A} + D\Delta^\star_A(\delta) + BK^\star(\gamma))^\top + W \tag{3-2.8}$$

$$h(\delta) = \left\langle (\Delta^\star_A(\delta))^\top\Delta^\star_A(\delta), \Sigma_x \right\rangle. \tag{3-2.9}$$

Then, for $\gamma = h(\delta)$,

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

$$K^\star(\gamma) = -\alpha R^{-1}B^\top P(\delta)(\Lambda(\delta))^{-1}\widehat{A} \tag{3-2.10}$$

is (the minimizing part of) the solution to the RLQR problem.

[6] See Lemma 3-2.9 for the definition and more information.

(ii) Furthermore, the matrix $\Delta^\star_A(\delta)$ in (3-2.7) is the worst-case[7] *model uncertainty, i.e., the maximizing solution is $A^\star_{c\ell}(\gamma) = \widehat{A} + BK^\star(\gamma) + D\Delta^\star_A(\delta)$, in other words, the worst-case system matrix is given by $A^\star(\gamma) = \widehat{A} + D\Delta^\star_A(\delta)$.*

(iii) Moreover,

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

(b) the map $h(\delta)$ is (real) analytic on the interval $[0, \overline{\delta})$.

See section 3-2-2 for a game theoretic interpretation of this "*breakdown point*" $\overline{\delta}$, which contains an intuitive and formal proof of the theorem as well. Note that we have chosen to interpret (3-2.7) as an ***additive*** uncertainty, but by construction, we could have interpreted the adversarial disturbance as an ***multiplicative*** uncertainty as well, *e.g.*, $A^\star(\gamma) = [I_n + \alpha\delta D^\top P(\delta)(\Lambda(\delta))^{-1}]\widehat{A} = \Delta \cdot \widehat{A}$. The implications of this observation are discussed in section 3-3.

In principle, one could use Theorem 3-2.4 to solve certain[8] Robust LQ regulator problems. However, in the spirit of [HS07] and references therein, like the famous *Lucas Critique* [Luc76], one could use the proposed game theoretic formulation also for model analysis. If society acts adversarially, how does it react to a policy change?

It is important to remark that although problem (3-1.2) is well-defined for all $\gamma \in \mathbb{R}_{\geq 0}$, Theorem 3-2.4 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 $h(\overline{\delta}) = \overline{\gamma}$. We do not necessarily have $\lim_{\delta \uparrow \overline{\delta}} h(\delta) = \infty$. This also explains the implicit formulation of the Theorem. We elaborate on this in section 3-4-2-2.

We make another remark.

Remark 3-2.5 (Nested sets). *It can be observed from definition 3-1.1 that for a fixed $\widehat{A}$ and $\gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_n$ we have $\mathcal{A}_{\gamma_1}(\widehat{A}) \subseteq \mathcal{A}_{\gamma_2}(\widehat{A}) \subseteq \ldots \mathcal{A}_{\gamma_n}(\widehat{A})$.*

The (local) continuity of the map h allows for showing that the nestedness is actually invariant under feedback.

Lemma 3-2.6 (Controlled uncertainty sets are nested). *Let γ_1 correspond be feasible in the sense of Theorem 3-2.4. Then there exists a $\gamma_2 \geq \gamma_1$ such that $\mathbb{A}_{\gamma_1}(\widehat{A} + \widehat{B}K^\star(\gamma_1)) \subseteq \mathbb{A}_{\gamma_2}(\widehat{A} + \widehat{B}K^\star(\gamma_2))$.*

So if one can solve Theorem 3-2.4 for some γ_1 and γ_2 satisfying $\gamma_1 \leq \gamma_2$, then the controller under γ_2 works ($\sqrt{\alpha}$-stabilizes) for both uncertainty sets, see for example Figure 3-4a. This also means that given some uncertainty set Δ, if it can be shown that $\Delta \subseteq \mathbb{A}_\gamma$, it is known there exists a single K which can $\sqrt{\alpha}$-stabilize the entire set Δ. This should be contrasted with standard norm-balls on (Δ_A, Δ_B) for which it is not immediately clear if there is a K which can handle the full set.

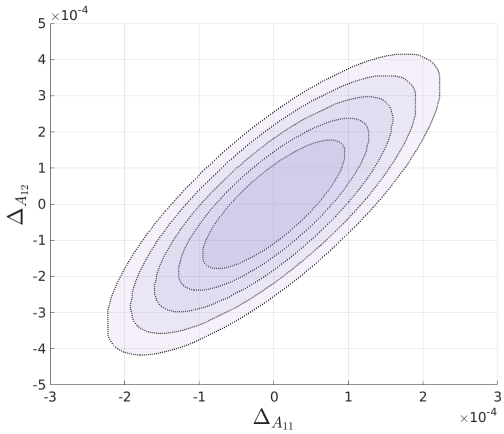
[7] In remark 3-2.10 we affirmatively answer the question if this worst-case model is actually a least-favourable model.

[8] Of course, we still have to see if this set is actually structurally appealing.

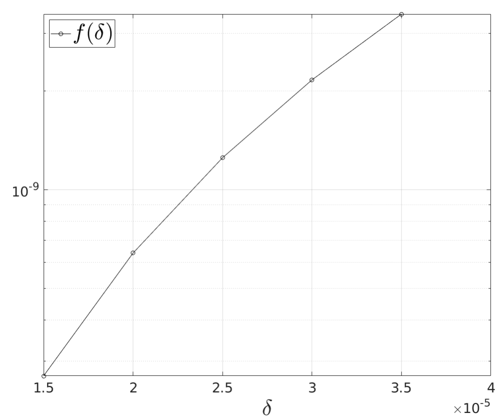
Proof of Lemma 3-2.6. Let $f(\Delta_{A_{c\ell}},\delta)$ be a real-valued function over a subset of $\mathbb{R}^{d\times n}\times\mathbb{R}_{\geq 0}$ defined by

$$f(\Delta_{A_{c\ell}},\delta) := \begin{cases} & \langle \Delta_{A_{c\ell}}^\top \Delta_{A_{c\ell}}, \Sigma_x\rangle, \\ \text{subject to} & \Sigma_x = A_{c\ell}\Sigma_x A_{c\ell}^\top + W, \quad \Sigma_x \succ 0, \\ & A_{c\ell} = \widehat{A} + \widehat{B}K^\star(\delta) + D\Delta_{A_{c\ell}}. \end{cases}$$

By the continuity result from Theorem 3-2.4.(iiib) the levelsets of f cannot intersect. Then map $\delta \mapsto h(\delta) =: \gamma$ and project that coordinate on the space of $\Delta_{A_{c\ell}}$, which by monotonicity of h preserves the lack of intersecting sets. Then due to the game theoretic interpretation we have $\Delta^\star_{A_{c\ell}}(\gamma_1) \in \partial \mathbb{A}_{\gamma_1}$ and for $\gamma_2 \geq \gamma_1$, $\Delta^\star_{A_{c\ell}}(\gamma_2) \in \partial\mathbb{A}_{\gamma_2}$ but $\Delta^\star_{A_{c\ell}}(\gamma_2) \notin \mathrm{Int}(\mathbb{A}_{\gamma_1})$ such that by continuity of $\Delta^\star_A$ in γ the sets must be nested. □



(a) Sets of $\mathbb{A}_\gamma$.



(b) Function $f(\delta)$ (see below).

Figure 3-4: Given the parameters from Section 3-4-3, we show in Figure 3-4a that for a small range of γ's (in fact $\delta \in [1.5, 3.5]\cdot 10^{-5}$), the sets $\mathbb{A}_\gamma(\widehat{A} + BK^\star(\gamma))$ are indeed nested. Although these sets appear to be ellipsoidal, in Figure 3-4b we show, using game-theoretic notation, $f(\delta) := \|L^\star(\delta)\|^2_{F,\Sigma_x(L^\star(\delta))} - \|L^\star(\delta)\|^2_{F,\Sigma_x(-L^\star(\delta))}$. Since $f(\delta)$ is strictly positive, the level-sets cannot contain $L^\star$ and $-L^\star$, and thus are not ellipsoidal. More on this in Section 3-3-1.

.

3-2-1 Uncertainty in the Pair (A, B)

Proposition 3-2.3 and Theorem 3-2.4 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 [JSM19] left off.

Let us be given some controller K $\sqrt{\alpha}$-stabilizing $\widehat{A}_{c\ell} := \widehat{A} + \widehat{B}K$ and $Q_{c\ell} := Q + K^\top RK$, being the closed-loop cost matrix. Then consider the problem

$$\sup_{A_{c\ell}\in\mathcal{A}_\gamma(\widehat{A}_{c\ell})} \mathcal{J}(A_{c\ell}, Q_{c\ell}), \tag{3-2.11}$$

Denote the solution to (3-2.2) by $A^\star_{c\ell}(\gamma) := \widehat{A}_{c\ell} + D\Delta^\star_{A_{c\ell}}(\gamma)$. Now we can directly apply Proposition 3-2.3 and obtain the next Corollary to it.

Corollary 3-2.7 (Decomposition of $\Delta^\star_{A_{c\ell}}(\gamma)$)**.** *If (3-2.11) is feasible and $\widetilde{h}(\delta) = \gamma$, then some 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 (3-2.3) 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 3-2.7 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 [YUP02], although for the infinite-horizon case. The same idea holds for Theorem 3-2.4, consider the problem

$$\inf_{K\in\mathbb{R}^{n\times m}} \sup_{A_{c\ell}\in\mathcal{A}_\gamma(\widehat{A}+\widehat{B}K)} \mathcal{J}(A_{c\ell}, Q + K^\top RK). \tag{3-2.12}$$

Assume that (3-2.12) is feasible in the sense of Theorem 3-2.4. 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 3-2.7 carries through:

Lemma 3-2.8 (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*

$$\begin{aligned}\Delta^\star_A(\gamma) &= \alpha(\delta^{-1}I_d - \alpha D^\top P(\delta)D)^{-1}D^\top P(\delta)\widehat{A},\\ \Delta^\star_B(\gamma) &= \alpha(\delta^{-1}I_d - \alpha D^\top P(\delta)D)^{-1}D^\top P(\delta)\widehat{B}.\end{aligned} \tag{3-2.13}$$

Proof of Lemma 3-2.8. This follows directly from Theorem 3-2.4 whereas the decomposition follows from any standard proof of Lemma 3-2.9, *e.g.,* solving the first step in the corresponding Bellman-Isaacs equation (*cf.* [BB95]). □

Using Lemma 3-3.3.(iv) and 3-4.4 from below, we clearly see the "*growing*" effect of δ, and by monotonicity in h, of γ. Indeed, γ functions as a radius.

3-2-1-1 The Uncertainty Set for (A, B)

Corollary 3-2.7 and Lemma 3-2.8 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 3-2.2 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 (3-1.1) 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 $K^\star, L^\star$ that $\Delta^\star_B = -\delta R$. Note however that this construction is usually not possible since commonly $m < n$, while $D = I_n$ and L is not rank-deficient. Nevertheless, in section 4-3-3 we have a similar discussion on the simple relation between $K^\star$ and $L^\star$ in the context of policy gradient.

Using this decomposition idea, we observe something extra, which we briefly highlight following [Pol87]. Consider for $Q \succeq 0$, $R \succ 0$ the usual discounted LQ problem and let without loss of generality the optimal control gain be $K(A,B)$. Here (A,B) is the real, yet unknown system, which we approximate with $(\widehat{A},\widehat{B})$. Let C be defined by $C^\top C = Q$ and define the sets

$$\begin{aligned} E_C &= \{(A,B) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \;:\; (\sqrt{\alpha}A, B, C) \text{ minimal realization}\},\\ G &= \{(\widehat{A},\widehat{B}) \in E_C \;:\; \widehat{A} + \widehat{B}K(\widehat{A},\widehat{B}) = A + BK(\widehat{A},\widehat{B})\}. \end{aligned}$$

So, G is the set of system matrices such that the estimated closed-loop system matrix equals the real closed-loop system matrix. Then the decomposition for $\Delta_{A_{c\ell}}$ implies that on G we have $\Delta_A + \Delta_B K(\widehat{A},\widehat{B}) = 0$. This lower dimensional manifold (think of a hyperplane) is of special interest in the study of self-tuning regulators, but for us it is simply the set of uncertainties we can deal with for free. A final remark on this decomposition is that since we parametrize a dn-dimensional object with $d(n+m)$ parameters, we lose compactness (recall the discussion corresponding to for example the set in Figure 3-2 and see Figure 3-14b for an example set of (a,b).).

Moreover, instead of decomposing the solution of Theorem 3-2.4, we could also introduce an uncertainty set for the pair (A,B) directly. Let this set be defined as

$$\mathcal{U}_\gamma((\widehat{A},\widehat{B});K) = \left\{(A,B) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \;:\; A + BK \in \mathcal{A}_\gamma(\widehat{A} + \widehat{B}K)\right\}. \tag{3-2.14}$$

Then a solution to

$$\inf_{K\in\mathbb{R}^{m\times n}} \sup_{(A,B)\in\mathcal{U}_\gamma\left((\widehat{A},\widehat{B});K\right)} \mathcal{J}(A+BK, Q + K^\top RK) \tag{3-2.15}$$

is given by (3-2.10) and (3-2.13). Of course, this description is rather implicit, but it generalizes all the (arbitrary) decompositions from above.

3-2-1-2 Other Methods to Incorporate B

At last, we highlight another approach to include uncertainties in B without arbitrary decompositions. This approach, as taken in [GBA94], hinges on extending the state space as proposed in [Bar83]. 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 3-2.4 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 $u_k^e := K^e x_k^e$. Back to the original problem, let $u_k := Kx_k$ for some K. Then from $u_{k+1} = K_x^e x_k + K_u^e u_k$ 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.

3-2-2 Game Theoretic Interpretation

In this section the main results are briefly explained via a introduction to dynamic game theory. It should be highlighted that the link between dynamic game theory and robust control is well studied, see [BB95, HS07] for an accessible and illuminating introduction.

We first introduce the concept of a dynamic game (*cf.* [BB95, BO99]). Consider the stochastic (discounted) two-player zero-sum ***dynamic game***

$$\begin{aligned} \inf_{\{\mu_k\}_{k\in\mathbb{N}}} \sup_{\{\nu_k\}_{k\in\mathbb{N}}} \quad & \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^{\infty}\alpha^k\left(x_k^\top Q x_k + u_k^\top R u_k - \delta^{-1} w_k^\top w_k\right)\right], \\ \text{s.t.} \quad & x_{k+1} = Ax_k + Bu_k + Dw_k + v_k, \\ & v_k \overset{i.i.d.}{\sim} \mathcal{P}(0,\Sigma_v), \quad x_0 \sim \mathcal{P}(0,\Sigma_0), \\ & u_k = \mu_k(x_k), \quad w_k = \nu_k(x_k). \end{aligned} \tag{3-2.16}$$

The μ-player selects a policy μ_k for the input u_k and the ν-player selects a policy ν_k for the input w_k. Here the parameter $\delta \in \mathbb{R}_{\geq 0}$ penalizes the input of the ν-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*"[9] in the sense that there are no cross-terms in the cost, thus the program largely relies on the single parameter δ. This parameter is constrained to live in the interval $[0,\overline{\delta})$, where $\overline{\delta}$ is referred to as the ***breakdown point***, *beyond* this value, the ν-player has to pay so little that the cost becomes unbounded[10].

To see a relationship between dynamic game theory and parametric uncertainty sets, suppose (3-2.16) has a solution and consider the following. The ν-policy aims to maximize the cost. But since the optimal μ-policy can handle this *worst-case* policy, it must also be able to handle policies of a *less* powerful adversary. This effectively gives rise to a whole *family* of state feedback policies the μ-player can handle. It turns out that in (3-1.2) we maximize over precisely this family.

To be a bit more formal we need one key Lemma:

Lemma 3-2.9 (*cf.* chapter 3 from [BB95] for the non-discounted deterministic case)**.** *Given a game (3-2.16) 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 $(\delta^{-1}I_d - \alpha D^\top PD) \succ 0$[11], where P is the sym-*

[9] 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 $S \succeq 0$.

[10] See ch.8 [HS07] for more on the relation between this breakdown point and $\mathcal{H}_\infty$ control. Also, see section 3-4-2-2 for explicit examples of what can happen *at* $\overline{\delta}$.

[11] An equivalent condition as promoted by [HS07] is to check $\operatorname{logdet}(\delta^{-1}I_d - \alpha D^\top PD) > -\infty$

metric minimal[12] *positive semi-definite solution to the Generalized Algebraic Riccati Equation (GARE):*

$$P = Q + \alpha A^\top P \Lambda^{-1} A, \quad \Lambda = \left(I_n + \alpha \left(BR^{-1}B^\top - \delta DD^\top\right) P\right), \tag{3-2.17}$$

then the optimal[13] *strategies are time-invariant, linear in x_k for $K^\star(\delta) \in \mathbb{R}^{m\times n}$ and $L^\star(\delta) \in \mathbb{R}^{d\times n}$ 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 $\mathcal{J}^\star = \langle P, \Sigma_0\rangle + \alpha(1-\alpha)^{-1}\langle P, \Sigma_v\rangle$.

To indicate why this might be of interest, see that we can interpret the policy $\nu_k^\star(x_k)$ as a perturbation $\Delta = DL^\star(\delta) = D\Delta_A$ to the system matrix A. Thus the controller $K^\star(\delta)$ can accommodate some uncertainty in A, namely this $D\Delta_A$. Now, the idea is the following, under the results from Lemma 3-2.9 we can rewrite the game into:

$$\begin{aligned}\inf_{K\in\mathbb{R}^{m\times n}} \sup_{\Delta_A\in\mathbb{R}^{n\times d}} \quad & \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^{\infty}\alpha^k x_k^\top\left(Q + K^\top RK - \delta^{-1}\Delta_A^\top\Delta_A\right)x_k\right],\\ \text{s.t.} \quad & x_{k+1} = (A + BK + D\Delta_A)x_k + v_k,\\ & v_k \overset{i.i.d.}{\sim} \mathcal{P}(0,\Sigma_v), \quad x_0 \sim \mathcal{P}(0,\Sigma_0).\end{aligned}$$

Then assume we know there is some $\gamma \in \mathbb{R}_{\geq 0}$ such that we can take the adversarial part out of the cost and put it into the constraints:

$$\begin{aligned}\inf_{K\in\mathbb{R}^{m\times n}} \sup_{\Delta_A\in\mathbb{R}^{n\times d}} \quad & \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^{\infty}\alpha^k x_k^\top\left(Q + K^\top RK\right)x_k\right],\\ \text{s.t.} \quad & x_{k+1} = (A + BK + D\Delta_A)x_k + v_k,\\ & v_k \overset{i.i.d.}{\sim} \mathcal{P}(0,\Sigma_v), \quad x_0 \sim \mathcal{P}(0,\Sigma_0),\\ & \langle \Delta_A^\top\Delta_A, \Sigma_x\rangle \leq \gamma, \quad \Sigma_x = \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^{\infty}\alpha^k x_k x_k^\top\right].\end{aligned}$$

This is a robust LQR problem under the set introduced in Definition 3-1.1. The question remains, can we find a link between γ and δ, *e.g.,* of the form $h(\delta) = \gamma$, such that we can perhaps relate their solutions as well? This question is affirmatively answered in Theorem 3-2.4.

With the intuition from this section in mind we present the proof of Theorem 3-2.4.

[12] In the terminology of p.81 ch.3 [BB95], 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.

[13] Not a general saddle-point (see *e.g.* [Mag76])

Proof of Theorem 3-2.4 Now this apparent link between the solution to a robust LQR problem and a dynamic game as set forth in section 3-2-2 is formalized, which constitutes the main result, Theorem 3-2.4. This is not new, see for example [DB92, HS07], where in the latter[14], the pair (γ,δ) is interpreted via multiplier theory (*cf.* [Lue69, Ber99]) 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 $(\{u_k\}_k,\{w_k\}_k)$ which eventually allows for numerically finding a solution depending on δ, given γ (see Lemma 3-4.1).

Recall Definition 3-1.1 and the RLQR problem (3-1.2). Let a solution to (3-1.2) be denoted by the pair $(K^\star(\gamma),\Delta_A^\star(\gamma))$ whereas a solution to (3-2.16), if it exists, is $(K^\star(\delta),L^\star(\delta))$. Then the next proof allows us to link the solution from the dynamic game (3-2.16) to the solution of the robust LQ regulator (3-1.2). This proof of Theorem 3-2.4 is split up into a few parts.

Proof of Theorem 3-2.4 part (i),(ii) and (iiia). Regarding (iiia), first consider the game (3-2.16). By Lemma 3-2.9 the cost can be equivalently written as $f(K,L)-\delta^{-1}g(K,L)$ for $u_k = Kx_k$, $w_k = Lx_k$, $x_{k+1} = Ax_k + Bu_k + Dw_k + v_k$ and the pair $f(K,L)$, $g(K,L)$ being defined by

$$f(K,L) = \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^{\infty}\alpha^k x_k^\top\left(Q+K^\top RK\right)x_k\right], \tag{3-2.18}$$

$$g(K,L) = \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^{\infty}\alpha^k w_k^\top w_k\right] = \left\langle L^\top L,\Sigma_x\right\rangle, \tag{3-2.19}$$

with $\Sigma_x = \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^{\infty}\alpha^k x_k x_k^\top\right]$[15]. Then $\sup_L\{f(K',L)-\delta^{-1}g(K',L)\}$ corresponds to program $\mathcal{P}_2$ from Lemma A-0.1 with the map h from (3-2.5) and an additional (fixed) parameter K'. The map $h(\delta)$ is non-decreasing on some interval $[0,\bar\delta)\subset\mathbb{R}_{\geq 0}$, $\bar\delta<\infty$. 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 [Whi90, HS07] the parameter $\bar\delta$ resembles their "*breakdown*" point $\underline{\theta}$.

Regarding (i)-(ii), by construction of the result for (iiia), the programs (3-1.2) and (3-2.16) are of the form

$$\widetilde{\mathcal{P}}_1(\gamma):\begin{cases}\inf\limits_{K\in\mathbb{R}^{m\times n}}\ \sup\limits_{L\in\mathbb{R}^{d\times n}}\ f(K,L)\\ \qquad\qquad\quad \text{s.t.}\quad g(K,L)\leq\gamma,\end{cases}\qquad \widetilde{\mathcal{P}}_2(\delta):\inf_{K\in\mathbb{R}^{m\times n}}\ \sup_{L\in\mathbb{R}^{d\times n}}\ f(K,L)-\delta^{-1}g(K,L),$$

respectively, for $f(K,L)$ and $g(K,L)$ defined by (3-2.18) and (3-2.19).

These programs $(\widetilde{\mathcal{P}}_1(\gamma),\widetilde{\mathcal{P}}_2(\delta))$ correspond to $\mathcal{P}_1(\gamma)$ and $\mathcal{P}_2(\delta)$ from Lemma A-0.1 but with an outer minimization step over K. Let the corresponding solutions to the inner maximization problems be denoted by $L_1^\star(\gamma,K)$ and $L_2^\star(\delta,K)$. Then by Lemma A-0.1 we have $L_1^\star(\gamma,K)=L_2^\star(h^{-1}(\gamma),K)$. Moreover, when $h(\delta)=\gamma$ then $L_1^\star(\gamma,K)=L_2^\star(\delta,K)$ and thereby $g(K,L_1^\star(\gamma,K))=g(K,L_2^\star(\delta,K))$.

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

[15]This step relies on the Bounded Convergence Theorem (*cf.* p.57 [AL06]) 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.*, $\lim_{n\to\infty}\int_{\mathcal{X}} f_n+g_n\mathrm{d}\mu=\int_{\mathcal{X}}\lim_{n\to\infty}f_n\mathrm{d}\mu+\int_{\mathcal{X}}\lim_{n\to\infty}g_n\mathrm{d}\mu$.

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 A-0.1 for the sake of contradiction it is not. For $\widetilde{\mathcal{P}}_1$ we effectively consider $\inf_K\{f(K, L_1^\star(\gamma, K))\}$ where it is known that $g(K, L_1^\star(\gamma, K)) \leq \gamma$ holds. However, since $h(\delta)=\gamma$ we can equivalently consider $\inf_K\{f(K, L_2^\star(\delta, K))\}$. Then to continue the contradictive argument assume there is some $\widetilde{K}$ such that

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

By construction we have $h(\delta)=\gamma$, and thus $g(\widetilde{K}, L_2^\star(\delta,\widetilde{K})) = \gamma = g\Big(K^\star(\delta), L_2^\star(\delta, K^\star(\delta))\Big)$ 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 $(K^\star(\delta), L^\star(\delta))$ exists, it is an optimal solution to both (3-2.16) and (3-1.2).

Thus, when there is a $\delta \geq 0 : h(\delta)=\gamma$, which we have by construction of the Theorem, then the solution to (2-1.2) is given by the pair $(K^\star(\delta), L^\star(\delta))$, for which the expressions are given by Lemma 3-2.9. 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 3-2.9 (feasibility of (3-2.16), *e.g.,* (A, B, C) being a minimal realization). □

Remark 3-2.10 (Least-favourable model). *It does not change any of the results, but we can in fact establish that our worst-case model is a least-favourable model. The proof is analogous to the one above.*

At last we characterize the regularity of the map h in the context of Theorem 3-2.4, which is again very useful with numerical algorithms in mind. This is done in the spirit of the work by Polderman [Pol86a, Pol87].

Proof of Theorem 3-2.4 (iiib). We will first show that $\overline{P}^+(\delta)$[16] 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

$$\ell(\delta, P) = P - Q - \alpha A^\top P\left(I_n + \alpha\left(BR^{-1}B^\top - \delta DD^\top\right)P\right)^{-1}A. \tag{3-2.20}$$

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 $\stackrel{L}{=}$, in Δ_P:

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

These steps hinge on geometric series for matrices, and a few linear algebraic identities[17]. Now

[16] See Lemma 3-2.9 for more on this notation.
[17] Most notably: $P(1+QP)^{-1} = (1+PQ)^{-1}P$ and $(I+P)^{-1} = I - (I+P)^{-1}P$.

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 [Pol86a]) for such a point $(\widetilde{\delta}, \overline{P}^{+}(\widetilde{\delta}))$. Therefore, we can apply the Implicit Function Theorem (*cf.* [KP03]), 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 $\overline{P}^{+}(\delta) \in C^{\omega}((0, \overline{\delta}))$ since $P(\delta)|_{\delta \in U_{\widetilde{\delta}}}$ are unique (see [SW94]) and stabilizing by continuity. This implies that $L^{\star}(\delta)$ is C^{ω} in δ and by Theorem E.1.4[18]. from [vS18], 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. □

See section 3-4-2-2 for some graphs of the map h.

3-2-3 A Relation to Contemporary Results, Inner and Outer Approximations

In Theorem 3-2.4 one solves a robust optimal control problem over $\mathcal{A}_{\gamma}$. To immediately link this to contemporary end-to-end frameworks we can find inscribed balls of Δ_A and Δ_B.

Since the algebraic Riccati equation (corresponding to discrete-time LQR) is analytic in (A, B, C) on the space of minimimal realizations [Del84], which is itself open in $\mathbb{R}^{n^2+mn+np}$, arbitrary small perturbations $(\Delta_{A_{\varepsilon}} = \varepsilon_A)$ will not blow up the cost. Therefore, we can always find a ball around 0. However, here we try to quantify this ball.

Lemma 3-2.11 (An Inscribed Ball of $\Delta_{A_{c\ell}}$)**.** *Consider* (3-1.2) *and assume* $Q + (K(\gamma)^{\star})^{\top} RK(\gamma)^{\star} \succ 0$ *then* $\mathcal{E}^{-} \subseteq \mathbb{A}_{\gamma}(\widehat{A} + \widehat{B}K^{\star}(\gamma))$ *for:*

$$\mathcal{E}^{-} := \left\{ \Delta_{A_{c\ell}} \in \mathbb{R}^{d \times n} \; : \; \|\Delta_{A_{c\ell}}\|_F^2 \leq \gamma \left[\kappa(Q + (K(\gamma)^{\star})^{\top} RK(\gamma)^{\star}) \mathrm{Tr}(\Sigma_x^{\star}(\gamma)) \right]^{-1} \right\}.$$

As previously remarked, these balls are by no means natural in terms of control theoretic uncertainty sets, but from a statistical point of view they are relevant (*cf.* [AYPS11, SMT+18]). Regarding an outscribed ball $\mathcal{E}^{+}$, one can observe that if we write the discrete Lyapunov equation compactly as $\Sigma_x(\Delta_A) = \alpha A_{c\ell} \Sigma_x(\Delta_A) A_{c\ell}^{\top} + W$ and assume $W \succ 0$ then

$$\mathbb{A}_{\gamma}(\widehat{A}_{c\ell}) \subseteq \left\{ \Delta_{A_{c\ell}} \in \mathbb{R}^{d \times n} \; : \; \|\Delta_{A_{c\ell}}^{\top}\|_{F,W}^2 \leq \gamma \right\} =: \mathcal{E}^{+}.$$

Lemma 3-2.11 applies to some uncertainty in $A_{c\ell}$, or if you like, in A. This can be extended to a rectangular set of uncertainties in A and B.

Lemma 3-2.12 (An Inscribed rectangular set of (Δ_A, Δ_B))**.** *Assume* $\{\Delta_{A_{c\ell}} \; : \; \|\Delta_{A_{c\ell}}\|_2^2 \leq r\} \subset \mathbb{A}_{\gamma}(\widehat{A} + \widehat{B}K^{\star}(\gamma))$ *and define* $\widetilde{r} := r_{\Delta_A} + r_{\Delta_B} r_K + 2\sqrt{r_{\Delta_A}}\sqrt{r_{\Delta_B}}\sqrt{r_K}$ *for some positive scalars* $r_{\Delta_A}, r_{\Delta_B}, r_K$*. Then, for all* $(\Delta_A, \Delta_B, K^{\star}(\gamma))$ *such that* $\|\Delta_A\|_F^2 \leq r_{\Delta_A}$, $\|\Delta_B\|_F^2 \leq r_{\Delta_B}$, $\|K^{\star}(\gamma)\|_F^2 \leq r_K$ *we have* $\Delta_A + \Delta_B K^{\star}(\gamma) \in \mathbb{A}_{\gamma}(\widehat{A} + \widehat{B}K^{\star}(\gamma))$ *when* $r \geq \widetilde{r}$.

Given these results, we can in principle find *some* finite-sample guarantees for our framework, which are by no means sharp or even computationally attractive. The point is that under mild conditions we have a non-empty interior.

[18] Effectively, by the results from Polderman [Pol87]

Proof of Lemma 3-2.11. Using Lemma A-0.2 we can bound $\sup_{\Delta_A \in \partial \mathbb{A}_\gamma} \|\Sigma_x(\Delta_A)\|_2$. First, observe that by construction we have

$$\left\langle Q + (K^\star(\delta))^\top RK^\star(\delta), \Sigma_x^\star \right\rangle \geq \left\langle Q + (K^\star(\delta))^\top RK^\star(\delta), \Sigma_x(\Delta_A) \right\rangle \forall \Delta_A \in \partial \mathbb{A}_\gamma.$$

If $Q + (K^\star(\delta))^\top RK^\star(\delta) \succ 0$ then we can directly apply Lemma A-0.2 and find

$$\sup_{\Delta_A \in \partial \mathbb{A}_\gamma} \|\Sigma_x(\Delta_A)\|_2 \leq \kappa \left(Q + (K^\star(\delta))^\top RK^\star(\delta) \right) \mathrm{Tr}(\Sigma_x^\star).$$

Now we find a bound via

$$\begin{aligned} \gamma =& \langle \Delta_A^\top \Delta_A, \Sigma_x(\Delta_A) \rangle \\ \leq& \|\Delta_A\|_F^2 \sup_{\Delta_A \in \partial \mathbb{A}_\gamma} \|\Sigma_x(\Delta_A)\|_2 \\ \leq& \|\Delta_A\|_F^2 \kappa (Q + (K(\delta)^\star)^\top RK(\delta)^\star) \mathrm{Tr}(\Sigma_x^\star). \end{aligned}$$

These inequalities imply that for all $\Delta_A \in \partial \mathbb{A}_\gamma$ we have

$$\left(\frac{\gamma}{\kappa (Q + (K(\delta)^\star)^\top RK(\delta)^\star) \mathrm{Tr}(\Sigma_x^\star)} \right)^{1/2} \leq \|\Delta_A\|_F$$

which defines an inscribed ball:

$$\mathcal{E}^- = \left\{ \Delta_A \in \mathbb{R}^{d \times n} \ : \ \|\Delta_A\|_F \leq \left(\gamma \left[\kappa (Q + (K(\delta)^\star)^\top RK(\delta)^\star) \mathrm{Tr}(\Sigma_x^\star) \right]^{-1} \right)^{1/2} \right\}.$$

□

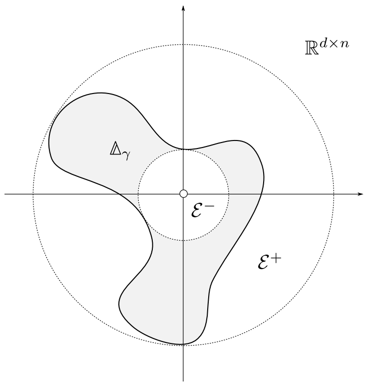


Figure 3-5: Schematic representation of $\mathcal{E}^-$ and $\mathcal{E}^+$ with respect to $\mathbb{A}_\gamma$.

Proof of Lemma 3-2.12. As before, let $\Delta_{A_{c\ell}} := \Delta_A + \Delta_B K^\star(\gamma)$. Now if we have an uncertainty set defined by all (Δ_A, Δ_B, K) such that $\|\Delta_A\|_F \leq \sqrt{r_{\Delta_A}}$, $\|\Delta_B\|_F \leq \sqrt{r_{\Delta_B}}$ and $\|K\|_F \leq \sqrt{r_K}$, then a sufficient condition to solve over this uncertainty set is that $\|\Delta_{A_{c\ell}}\|_2 \leq$

$\sqrt{r}$ is a valid inscribed ball for $r \geq r_{\Delta_A} + r_{\Delta_B} r_K + 2\sqrt{r_{\Delta_A}}\sqrt{r_{\Delta_B}}\sqrt{r_K}$. This simply follows from

$$\begin{aligned} r \geq & \|\Delta_A\|_F^2 + \|\Delta_B\|_F^2 \|K^\star(\gamma)\|_F^2 + 2\|\Delta_A\|_F \|\Delta_B\|_F \|K^\star(\gamma)\|_F \\ \geq & \|\Delta_A^\top \Delta_A\|_F^2 + \|(K^\star(\gamma))^\top \Delta_B^\top \Delta_B K\|_F^2 + 2\|\Delta_A^\top \Delta_B K\|_F \\ \geq & \|(\Delta_A + \Delta_B K^\star(\gamma))^\top (\Delta_A + \Delta_B K^\star(\gamma))\|_2 \\ = & \|\Delta_{A_{c\ell}}\|_2^2. \end{aligned}$$

□

Obtaining sharper bounds would be interesting, since then the ratio $\mathrm{Vol}(\mathcal{E}^+)/\mathrm{Vol}(\mathcal{E}^-)$, possibly as a function of γ, would provide further information about the symmetry of our set.

3-3 More on the Qualitative Properties of Game Theoretic Robust LQ Regulators

In this section we highlight several structural properties of our uncertainty set and worst-case model which bring about new insights in potential applications, or the lack thereof.

Linear Dynamical systems are unfortunately often compared based on how close they are in a certain, *non-system-theoretic*, norm. For example, when looking at the system matrix, a stable- and unstable system can be arbitrary close in a any induced p-norm, say $\|\widehat{A} - A\|_2$, while *qualitatively* they are obviously different.

Moreover, from a robust control point of view, the choice of distance metric greatly influences the shape of the confidence set around some $(\widehat{A}, \widehat{B})$ in $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n}$. As indicated via Example 2-2.1 and 2-2.3, these sets are for example by no means convex.

There are of course many methods to measure the distance between systems *e.g.*, stability radii and frequency domain norms, but a particular natural approach is to take closed-loop stability- and thereby design specifications, directly into account. To some readers this approach might be reminiscent of the *Gap*- and *Graph* topologies (*e.g.*, see [Zhu89]). In short, given a nominal transfer function $H(P_0, C_0)$ for P_0 the plant and C_0 the controller, they effectively study neighbourhoods of P_0

$$U(P_0, \epsilon) = \{P \, : \, C_0 \text{ stabilizes } P,\ \|H(P_0, C_0) - H(P, C_0)\| < \epsilon\}$$

inducing a topology T. Then if the systems P and P_0 are ϵ-close in T, they are necessarily stabilizable with the same controller.

Instead of measuring distances between general systems, we can first try to classify them. Mechanically speaking, think of separating *spring*- and *damper*-like scalar systems. Therefore, we investigate to what extend our framework preserves topological properties, as introduced in section 2-2-4. Here we will follow the exposition by Kuiper and Robbin in [Rob72, KR73], focusing on ***linear endomorphisms***, which are simply linear maps $f : V \to V$ for some vector space V. We speak of ***automorphisms*** when the map f is also invertible.

We use the following definition[19]

[19] See appendix B-2 for an explanation of how this definition can be interpreted.

Definition 3-3.1 (Topological Equivalence)**.** *Two endomorpshims $f : V \to V$ and $g : W \to W$ are* topologically equivalent (conjugate) *if and only if there exists a homeomorphism $\varphi : V \to W$ such $g = \varphi \circ f \circ \varphi^{-1}$,* i.e. *the diagram*

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\varphi"'] & V \arrow[d, "\varphi"] \\ W \arrow[r, "g"] & W \end{tikzcd}$$

commutes. This relation will be denoted by $f \overset{t}{\sim} g$.

Instead of Definition 3-3.1 one encounters *linear equivalence* more often in System & Control Theory, *e.g.*, $\widetilde{A} = TAT^{-1}$ for some $T \in \mathsf{GL}(n, \mathbb{R})$. A simple, yet illuminating example from [KR73] is given by $f(x) = 2x$ and $g(x) = 8x$. When thought of as dynamical systems, we cannot speak of linear equivalence since the eigenvalues are clearly different, but qualitatively they are the same. Thus indeed, we observe that $f \overset{t}{\sim} g$ since $\varphi(x) = x^3$ is the corresponding homeomorphism[20]. In fact, these scalar linear maps have just 7 equivalence classes, greatly simplifying their study.

Consider a linear dynamical system as an endomorphism on some finite n-dimensional vector space V, say $\mathbb{R}^n$. Now decompose $\mathbb{R}^n$ as $\mathbb{R}^n = W^\infty(f) \oplus W^+(f) \oplus W^-(f) \oplus W^0(f)$, defined by

$$W^\infty(f) = \bigoplus_{\lambda=0} E_\lambda(f), \quad W^+(f) = \bigoplus_{0<|\lambda|<1} E_\lambda(f), \quad W^-(f) = \bigoplus_{|\lambda|>1} E_\lambda(f), \quad W^0(f) = \bigoplus_{|\lambda|=1} E_\lambda(f), \tag{3-3.1}$$

for $E_\lambda(f)$ the usual eigenspace of the pair (f, λ). Moreover, let $f|_a = f|_{W^a(f)}$ be the automorphic part, *i.e.* $W^a = W^+ \oplus W^- \oplus W^0$. In Conjecture A of [KR73] several conditions are proposed to assert topological equivalence. We focus on one of them: *orientation*. See [AMR88, ch.6] for a formal discussion on orientation. We call a linear automorphism 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 $\mathrm{Or}(f) = -1$.

Lemma 3-3.2 (Orientation, a necessary condition)**.** *Let $f : V \to V$ and $g : W \to W$ be two linear automorphims. Then $g = \varphi \circ f \circ \varphi^{-1}$ for some homeomorphism $\varphi : V \to W$, only if $\mathrm{Or}(f) = \mathrm{Or}(g)$.*

(informal) Proof of Lemma 3-3.2 . If we could pick φ to be linear, the result is obvious since we merely have a similarity transformation which only holds when $\det(f) = \det(g)$. The general result follows from the fact that φ is a homeomorphism, thereby either φ and φ^{-1} preserve orientation, or both reverse it. Formally put, as is widely known, orientation is a *topological invariant.* □

For example, take $f(x) = 0.5x$ and $g(x) = -0.5x$, then since the orientations are different, they are not topologically equivalent. When looked upon as a dynamical system, *e.g.*, $x_{k+1} = f(x_k)$, then it is clear that for non-zero initial conditions there does not exist a

[20] Simply check that $8x \circ x^3 = x^3 \circ 2x = 8x^3$.

homeomorphism mapping $\mathrm{Im}(f)$ to $\mathrm{Im}(g)$, since such a homeomorphism must be necessarily monotone. A prototypical example of a n-dimensional orientation preserving dynamical system is a mechanical system moving through space. There is no way that such a rigid object can turn through itself under its own dynamics, *e.g.,* see Figure 3-6. More examples can be put together in for example the area of graph-theoretic models due to permutation matrices having determinant ± 1. To continue, given some endomorphism f, we can without

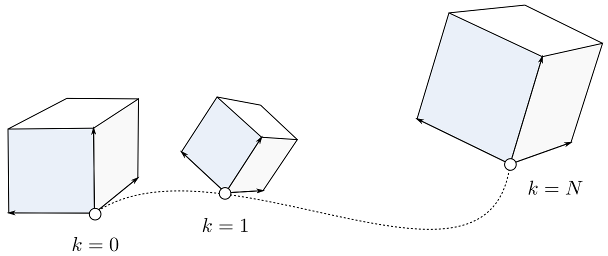


Figure 3-6: The unit-cube remains positively oriented under a orientation preserving automorphism.

loss of generality appeal to Lemma 3-3.2, by using $f|_a$ instead[21]. Of course, the g under consideration must satisfy $\dim(W^a(f)) = \dim(W^a(g))$.

As indicated before, the type of uncertainty set we consider is difficult to *quantify* due to the dependence on $K^\star(\gamma)$[22]. We can however observe several *qualitative* features.

Lemma 3-3.3 (Qualitative features of extremizers in (3-1.2), implications of Theorem 3-2.4)**.** *For simplicity assume $D = I_n$, then*

(i) *The worst-case closed-loop uncertainty resides on the boundary of the uncertainty set,* i.e., *when $h(\delta) = \gamma$ we know that $\Delta^\star_{A_{c\ell}}(\gamma) \in \partial \mathbb{\Delta}_\gamma(\widehat{A} + \widehat{B}K^\star(\gamma))$.*

(ii) *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 $\mathbb{\Delta}_\gamma(\widehat{A} + BK^\star(\gamma)) \subseteq \{\Delta_A \in \mathbb{R}^{n\times n} : \mathrm{Ker}(\widehat{A}) \subseteq W^+(\sqrt{\alpha}(\widehat{A} + \Delta_A))\}$ (see Example 3-3.5 below).*

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

(iv) *For $A^\star(\gamma) = \widehat{A} + \Delta^\star_A(\gamma)$, we have $\|A^\star(\gamma)\|_F > \|\widehat{A}\|_F$ almost surely. Moreover, using decomposition (3-2.13) we additionally have,* a.s., *$\|B^\star(\gamma)\|_F > \|\widehat{B}\|_F$.*

Orientation is just one part of topological equivalence, but to put it in simple words, item (iii) tells us that an adversarial player does not reveal itself that easily. Moreover, it means that

[21] For example, $f(x) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x$, $g(x) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} x$ and $h(x) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} x$ have the same orientation. See that $f(x)$ and $h(x)$ are even linearly equivalent.

[22] See for example (3-1.2), we maximize over $\mathcal{A}_\gamma(\widehat{A} + BK)$, the uncertainty depends on the solution $K^\star(\gamma)$.

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 under $\widehat{A} \mapsto \widehat{A} + \Delta_A^\star(\gamma) =: A^\star(\gamma)$[23]. In addition, item (ii) and (iii) imply that the $W^\infty(\widehat{A}x)$ is invariant under the worst-case perturbation, with our without feedback. These observations, together with item (iv), introduce new challenges for *unbiased* identification algorithms, one of them is explained in Figure 3-7[24], whereas section 3-4-4 highlights benefits in the context of *biased* identification. See appendix B-5 for a brief introduction of why Least-Squares identification can lead (under for example ergodic or episodic assumptions) to ellipsoidal sets of estimates.

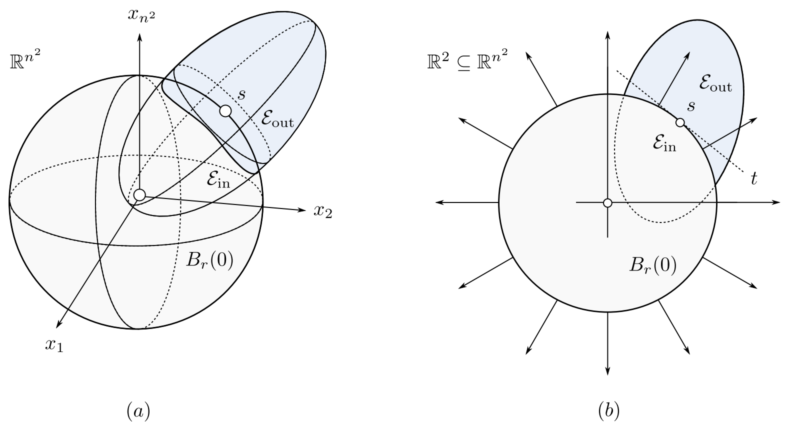


Figure 3-7: Let s represent the real system matrix $(\mathrm{vec}(A))$. Using unbiased least-squares we can form an ellipsoid around s given by $\mathcal{E} := \mathcal{E}_{\text{in}} \cup \mathcal{E}_{\text{out}}$, containing estimates of s, denoted $\widehat{s}$. Then since $\|A^\star\|_F > \|\widehat{A}\|_F$ *a.s.*, an estimate $\widehat{s} \in \mathcal{E}_{\text{min}}$ might lead to a worst-case model close to s, while the worst-case model related to a $\widehat{s} \in \mathcal{E}_{\text{max}}$ is even further away from s then the initial $\widehat{s}$. Think of the vectors in (b). The critical observation is however that $\mathrm{Vol}(\mathcal{E}_{\text{out}}) > \mathrm{Vol}(\mathcal{E}_{\text{in}})$ such that the push in the wrong direction is likely to dominate, hence, leading to bad performance. This observation will be further highlighted in sections 3-4-3-1,3-4-4.

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, see for example (3-4.8) from section 3-4-2-1 or think of T as the "*radius*" in matrix polar decomposition. Moreover, it be observed that the spectrum of T does not intersect the open unit-disk, which is exemplified in Figure 3-16 below. Another interpretation is the most clearly illustrated when considering a standard LQ regulator problem, say for $B = R = I_n$ and $\alpha = 1$ such that $K^\star(0) = -(I_n + P)^{-1}PA$. Since $P \in \mathcal{S}^n_+$ we know that $(I_n + P)^{-1}P$ is symmetric. Going back to our LQ problem, $R \succ 0$ prevents us from simply selecting $K = -A$, instead we end up with closed-loop dynamics of the form $x(k+1) = (I_n - (R+P)^{-1}P)Ax(k)$. So indeed, the cost determines *how* $x(k) \to 0$ for $k \to \infty$, but we also see this symmetric scaling factor in front of A since for the cost the influence of state $x_i(k)$ on $x_j(k+1)$ or the influence of state $x_j(k)$ on $x_i(k+1)$ are equally important.

To prove Lemma 3-3.3 we need one useful property of $\Lambda(\delta)$:

[23] 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.

[24] To see how $\mathrm{Vol}(\mathcal{E}_{\text{out}})/\mathrm{Vol}(\mathcal{E}_{\text{in}})$ 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 $\mathrm{Vol}(\mathcal{E}_{\text{out}})/\mathrm{Vol}(\mathcal{E}_{\text{in}}) \gtrsim C^n$, $C \in (1,2]$, for an even idealized scenario.

Lemma 3-3.4 ($\Lambda(\delta)$ is an orientation preserving map)**.** *The matrix* $\Lambda(\delta)$ *has positive eigenvalues and thus* $\det(\Lambda^{-1}(\delta)) > 0$.

Proof. The map $\Lambda(\delta) = (I_n + \alpha(BR^{-1}B^\top - \delta DD^\top)P(\delta))$ 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 $\delta \in [0,\bar{\delta})$. ☐

Proof of Lemma 3-3.3. We do the proof per item:

(i) This follows directly from Lemma 3-2.6 and the proof of Theorem 3-2.4 since the worst-case uncertainty $\Delta_A^\star(\delta)$ from Theorem 3-2.4 satisfies $h(\delta) = \langle(\Delta_A^\star(\delta))^\top\Delta_A^\star(\delta)\Sigma^\star(\delta)\rangle = \gamma$.

(ii) The fact that the worst-case closed-loop system can be written as $(\Lambda^\star(\delta))^{-1}\widehat{A}$ follows from Lemma 3-2.9 and Lemma 3-3.4 or just by direct computation. This also holds for $\gamma \to 0$ since it also holds for the standard LQR closed-loop system[Pol86b]. The last part follows from (3-2.10), $K^\star(\gamma)$ 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 $\mathrm{Ker}(\widehat{A}) \subseteq W^+(\sqrt{\alpha}(\widehat{A}+\Delta_A))$ when $\Sigma_0 \succ 0$.

(iii) Lemma 3-3.4 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 [BB95] 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.

(iv) 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 $\mathrm{vec}(A^\star(\gamma)) = (I_n \otimes T)\mathrm{vec}(\widehat{A})$. 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 (3-2.13), using the identity $(I+(I-P)^{-1}P) = (I-P)^{-1}$ we can write $B^\star(\gamma)$ as $(I_n - \alpha\delta P(\delta))^{-1}\widehat{B}$. Then the result follows from $(\delta^{-1}I_n - \alpha P(\delta)) \succ 0$, symmetry of P and a similar line of arguments as above.

☐

Example 3-3.5 (Kernel of LQ regulators). *Consider the matrices*

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

Then $c(1\ \ 1)^\top \in \mathrm{Ker}(\widehat{A})\ \forall\ c \in \mathbb{R}$. *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* $c \in \mathbb{R}$*:*

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

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* $E_{\lambda=1.2}(A^\star) = \mathrm{Ker}(\widehat{A})$, *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* $\widehat{A}$.

Lemma 3-3.4 has another implication:

Corollary 3-3.6 (Closed-loop Topological Equivalence). *Let* $D = I_n$, $\alpha = 1$ *and* $\Sigma_v = 0$, i.e., *make* (2-1.1), *up to* x_0, *deterministic. Moreover, interpret* $\Delta^\star(\gamma)$ *as either* $\Delta^\star_A(\gamma)$ *or* $\Delta^\star_{A_{c\ell}}(\gamma)$ *for some feasible formulation of* (3-1.2) *or* (3-2.15). *Then, the nominal-, robust- and worst-case robust closed-loop systems given by* $f(x) := (\widehat{A} + \widehat{B}K^\star(\gamma))x|_{\gamma=0}$, $g(x) := (\widehat{A} + \widehat{B}K^\star(\gamma))x|_{\gamma\in(0,\bar{\gamma})}$, $h(x) := (\widehat{A} + \widehat{B}K^\star(\gamma) + \Delta^\star(\gamma))x|_{\gamma\in(0,\bar{\gamma})}$, *respectively, are topologically equivalent. In other words,* $\forall \delta \in [0, \bar{\delta})$ *there exist a homeomorphisms* $\varphi, \psi : \mathbb{R}^n \to \mathbb{R}^n$ *such that* $\varphi \circ f \circ \varphi^{-1} = g = \psi \circ h \circ \psi^{-1}$.

Proof. Recall that $f(x) = \Lambda^{-1}(\delta)\widehat{A}x|_{\delta=0}$, $g(x) = (I_n - \delta P(\delta))\Lambda^{-1}(\delta)\widehat{A}x|_{\delta\in(0,\bar{\delta})}$ and $h(x) = \Lambda^{-1}(\delta)\widehat{A}x|_{\delta\in(0,\bar{\delta})}$. Since $(\delta^{-1}I_n - P) \succ 0$ must hold, the kernel and orientation carry over. Next, decompose $\mathbb{R}^n$ as above: $\mathbb{R}^n = W^\infty(\widehat{A}x) \oplus W^+$. Then the result follows from Proposition 5.2 in [KR73] and Lemma 3-3.4. □

Corollary 3-3.6 indicates that the uncertainties we consider are somewhat natural[25], at least for a deterministic system. Think of an identified model $(\widehat{A}, \widehat{B})$ used for simulation and controller design. Then, for our uncertainty set (3-1.1) there is a K, *i.e.,* $K^\star(\gamma)$ such that the simulated behaviour extends *structurally* to the real system, with or without worst-case uncertainty. Think of Figure 3-8 (a), if c_1 is a simulated trajectory for some scalar system, then the robust and worst-case trajectories are of the form c_2, c_3, but never like curve w. This is the beauty. However, in contrast to uncertainty sets arising from statistics, these kind of structure preserving uncertainties are not symmetric; for example, like region (3) around scalar system $\widehat{a}$ in Figure 3-8 (b), which is clearly not symmetric.

[25] Previously we used mechanical intuition to introduce topological equivalence, but one needs to be careful. For example, $\mathrm{diag}(e,e)x \overset{t}{\sim} \mathrm{diag}(-e,-e)x$, which hinges on orientation.

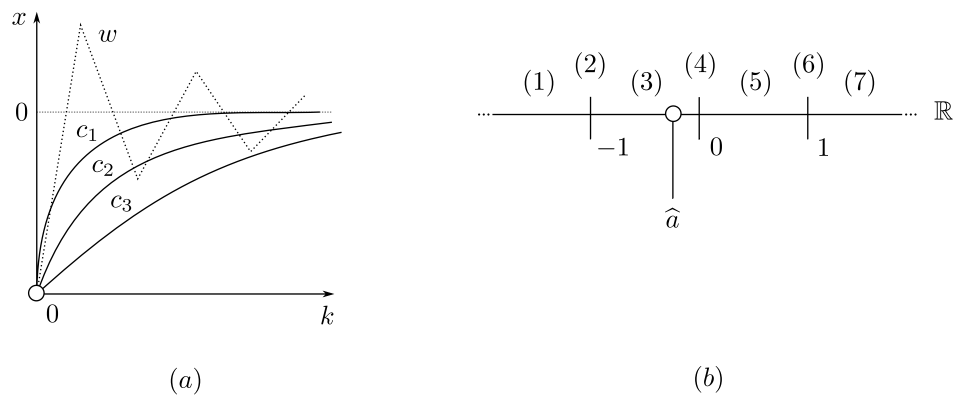


Figure 3-8: (a) The systems from Corollary 3-3.6 share qualitative properties like $\{c_1, c_2, c_3\}$. (b) $\mathbb{R}/\overset{t}{\sim}$ consists of 7 classes, which are by no means symmetric around a model within.

3-3-1 Almost Surely Conservatism

Item (iii) from Lemma 3-3.3 has a different implication. Before stating this result we introduce some terminology, but only the bare minimum. Classic references on this material are Halmos [Hal70] and Folland [Fol99]. However, instead of general topological spaces, we will be concerned with $\mathbb{R}^n$. A set $A \subset X$ is ***dense*** in X when $\mathrm{cl}(A) = X$. The local complement of dense sets are especially important for us. A set $X \subset \mathbb{R}^n$ is ***nowhere dense*** in $\mathbb{R}^n$ when there does not exist any Euclidean ball $B_r(x)$ around some $x \in X$, which is fully contained in X. Furthermore, we denote by μ the n-dimensional Lebesgue measure, which coincides with the standard volume, so for any (measurable) set $S \subset \mathbb{R}^n$ we have $\mu(S) = \mathrm{Vol}(S)$. Indeed, any nowhere dense set $X \subset \mathbb{R}^n$ is of measure zero, *i.e.,* $\mu(X) = 0$. Now following [Fol99], let $(X, \mathcal{F}, \mu)$ be some measure space, where a statement T is true for all $x \in X$, except for some set $Y \subset X$, with $\mu(Y) = 0$. Then T is said to be true for almost every $x \in X$, or differently put, T is true ***almost everywhere***. When the underlying measure space is actually a probability space, *i.e.,* $\mu(X) = 1$, we say that T is true ***almost surely*** (*a.s.*). At last we give two important examples of sets which are open and dense in their ambient spaces under the standard topology. First, the general linear group $\mathsf{GL}(n, \mathbb{R})$ is open and dense in $\mathbb{R}^{n\times n}$ (*cf.* [DK99]). Secondly, controllable pairs (A, B) are open and dense in $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$, following essentially the same argument as for $\mathsf{GL}(n, \mathbb{R})$. Thus, any triple (A, B, C) sampled from $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n}$ is almost surely a minimal realization.

Now we can continue the discussion on uncertainty sets. Classically, uncertainty sets like proposed in [GBA94] were of the form

$$\left\{(\Delta_A, \Delta_B) \;:\; \begin{pmatrix}\Delta_A & \Delta_B\end{pmatrix} = DF\begin{pmatrix}E_1 & E_2\end{pmatrix},\; F^\top F \preceq I\right\} \tag{3-3.2}$$

for some fixed D, E_1, E_2 and free variable F, all of appropriate dimension. At that point these sets were mostly of academic interest. However, recently, the purpose of solving robust LQ regulators comes from the desire to establish end-to-end performance bounds, *e.g., regret*- and *sample complexity* bounds. Therefore, the modern description of the uncertainty is fully driven by identification algorithms and measure concentration inequalities. Besides the operator-norm [SMT+18, DMM+18], *e.g.,* $\|\Delta_A\|_2 \leq \varepsilon_A$, $\|\Delta_B\|_2 \leq \varepsilon_B$, a different set, based on the Frobenius-norm, is often used and closely resembles our description. This latter group, *e.g.,* [AYS11, AL18, CKM19], builds on [AYPS11] (and references therein), which gives a concentration inequality for ℓ_2-regularized least-squares. In that work, the authors use the *Optimism in the Face of Uncertainty* (OFU) principle. Here, a confidence set around a nominal model is created, whereafter the control law is designed for the most *optimistic* model in the set. This should be contrasted with our approach so far, which could have been called *Pessimism in the Face of Uncertainty.* To be brief but concrete, we informally state part of Theorem 1 from [AYS11]:

Theorem 3-3.7 (Part of Theorem 1 [AYS11])**.** *Let $\Theta^\top := \begin{pmatrix} A & B \end{pmatrix}$, $z_k^\top := \begin{pmatrix} x_k^\top & u_k^\top \end{pmatrix}$ such that $x_{k+1} = \Theta^\top z_k + v_k$ and assume that the noise is sub-Gaussian. Now let $\widehat{\Theta}_k$ be the usual ℓ_2-regularized least-squares estimator of Θ. Then with probability at least $1-\delta$ we have*

$$\mathrm{Tr}\left((\Theta - \widehat{\Theta}_k)^\top V_k (\Theta - \widehat{\Theta}_k)\right) \leq \beta_k(\delta)$$

for $V_k \in \mathcal{S}^{n+m}_{++}$ defined by $V_k := \lambda I + \sum_{i=0}^{k-1} z_k z_k^\top$ and some function β_k.

Therefore, when we define $\Delta_{\Theta_k} := \Theta - \widehat{\Theta}_k$, we have

$$\mathbb{P}\left\{\Delta_{\Theta_k} \in \left\{\Delta_\Theta \in \mathbb{R}^{n\times(n+m)} \ : \ \|\Delta_\Theta\|^2_{F,V_k} \leq \beta_k(\delta)\right\} \geq 1-\delta\right\}, \tag{3-3.3}$$

which is precisely of the form as the uncertainty set $\mathbb{\Delta}_\gamma$ (see Definition 3-1.1), *e.g.,* after an embedding in $\mathbb{R}^{n^2+nm}$ a standard ellipsoid[26]. See that these formulations are remarkably similar if we assume $u_k = Kx_k$. Specifically, let $\lambda \to 0$, $\alpha \to 1$ and truncate $\Sigma_x = \mathop{\mathbb{E}}_{x_0,v}\left[\sum_{k=0}^\infty \alpha^k x_k x_k^\top\right]$:

$$\begin{aligned}\mathrm{Tr}\left((\Theta - \widehat{\Theta}_k)^\top V_k (\Theta - \widehat{\Theta}_k)\right) &= \mathrm{Tr}\left(\begin{pmatrix}\Delta_A & \Delta_B\end{pmatrix}\begin{pmatrix}I_n \\ K\end{pmatrix}\sum_{i=0}^{k-1} x_i x_i^\top \begin{pmatrix}I_n \\ K\end{pmatrix}^\top \begin{pmatrix}\Delta_A & \Delta_B\end{pmatrix}^\top\right) \\ &= \mathrm{Tr}\left(\Delta_{A_{c\ell}}^\top \sum_{i=0}^{k-1} x_i x_i^\top \Delta_{A_{c\ell}}\right) = \left\langle \Delta_{A_{c\ell}}^\top \Delta_{A_{c\ell}}, \sum_{i=1}^{k-1} x_i x_i^\top \right\rangle.\end{aligned} \tag{3-3.4}$$

Where the last step follows from the invariance of $\mathrm{Tr}(\cdot)$ under cyclic permutation and the symmetric term in the middle. Thus, as seen from (3-3.4), the set in (3-3.3) and Definition 3-1.1 are closely related. For example, when one can find a feasible γ such that $V_k \succeq \sup_{\Delta_A \in \mathbb{\Delta}_\gamma} \Sigma_x(\Delta_A)$ and $\beta_k(\delta) \leq \gamma$, then a robust controller, with high probability stability guarantees, can be synthesized. This is of course merely a reformulation of the approximations in section 3-2-3.

Now, Proposition 3-2.1 showed that our set can be non-convex such that one might immediately conclude that when we solve the robust LQR problem over one of these inscribed ellipsoids we just discussed, the control law is necessarily conservative. However, one might argue as well that since our set is connected (ch.4) with an *a.s.* smooth boundary[27] it becomes ellipsoidal if we just take γ to be sufficiently small? This would be great, since then we might be able to efficiently solve the robust LQ regulator problem corresponding to (3-3.3). It turns out that this is not the case, as was already hinted at in Figure 3-4.

[26] Recall, an ellipsoid is usually defined as the solid $\mathcal{E} := \{x \in \mathbb{R}^n \ : \ x^\top Q x \leq 1\}$ for some $Q \succ 0$, see *e.g.,* ch. V [Bar02].

[27] This can be shown using the tools from [Pol86a].

Lemma 3-3.8 (Most adversaries do not live on ellipsoids). *Let $D = I_n$ and for simplicity of notation, consider only some uncertainty in A. Then we can make two comments regarding the relation between (3-3.3) and the uncertainty sets as seen in dynamic games,* i.e., *(3-3.5) below.*

(i) Consider uncertainty sets of the form (3-3.3) with only some uncertainty in A such that $V_k \in \mathcal{S}^n_{++}$. Then, there is a.s. *no $\gamma > 0$ such that $\{\Delta_A : \|\Delta_A\|^2_{F,V_k} \leq \gamma\} = \{\Delta_A : \|\Delta_A\|^2_{F,\Sigma_x} \leq \gamma\}$ with $V_k = \Sigma_x$,* i.e., *we should not expect to simply directly link (3-3.3) via (3-3.4) to a dynamic game.*

(ii) Moreover, the set

$$\mathbb{A}_\gamma(\widehat{A} + BK^\star(\gamma)) = \left\{\Delta_A \in \mathbb{R}^{n\times n} \ : \ \|\Delta_A^\top\|^2_{F,\Sigma_x} \leq \gamma\right\} \tag{3-3.5}$$

or better yet, its canonical embedding, via $\mathrm{vec} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n^2}$, *in* $\mathbb{R}^{n^2}$

$$\left\{\mathrm{vec}(\Delta_A^\top) \in \mathbb{R}^{n^2} \ : \ \|\mathrm{vec}(\Delta_A^\top)\|^2_{2,(I_n\otimes\Sigma_x)} \leq \gamma\right\}$$

is almost surely no ellipsoid.

Proof of Lemma 3-3.8. Regarding (i), for the set to be ellipsoidal like (3-3.3) we need Σ_x to be constant, at least locally. This Discrete Lyapunov solution in Definition 3-1.1 is unique, depending on $A_{c\ell} := \widehat{A} + \Delta_A + BK^\star(\gamma)$ in a Kronecker product fashion, *i.e.*,

$$\mathrm{vec}(\Sigma_x) = (I_{n^2} - \alpha A_{c\ell} \otimes A_{c\ell})^{-1}\mathrm{vec}(W),$$

where $W \succ 0$ contains the covariance matrices. Therefore, some different closed-loop systems, $A_{c\ell}$ and $A'_{c\ell}$, can only give rise to the same Σ_x when they are both feasible and

$$A_{c\ell} \otimes A_{c\ell} = A'_{c\ell} \otimes A'_{c\ell}. \tag{3-3.6}$$

However, we know from Lemma 3-3.9 (see below) that for some fixed $A_{c\ell}$, all $A'_{c\ell}$ which satisfy (3-3.6) are nowhere dense in the ambient space $\mathbb{R}^{n\times n}$. Moreover, $A'_{c\ell}$ uniquely defines Δ'_A by translation. Hence, the set of Δ'_A satisfying (3-3.6) is nowhere dense in $\mathbb{R}^{n\times n}$ as well. This means that $\Sigma_x(\Delta_A)$ is almost surely not constant, not even locally, which means that we can never recreate a set similar to (3-3.3).

This does however not immediately imply that our game theoretic uncertainty set is most likely no ellipsoid. For example, consider the two sets

$$\mathcal{E}_1 = \left\{x \in \mathbb{R}^n \ : \ x^\top F(x) x \leq \gamma\right\}, \quad \mathcal{E}_2 = \left\{x \in \mathbb{R}^n \ : \ x^\top Q x \leq \gamma\right\}.$$

Then $\mathcal{E}_1 = \mathcal{E}_2$ for $n = 1$, $f(x) = x^2$, $Q = 1$ and $\gamma = 1$, showing that the lack of a constant weighting matrix does not rule out being ellipsoidal.

Regarding the second item (ii). If we demand that $\partial\mathbb{A}_\gamma$ is ellipsoidal[28], than it must contain the worst-case uncertainty $L^\star$, but also $-L^\star$ by symmetry. Since they are on the boundary

[28] With some abuse of notation, we mean, of the form $\{x \ : \ x^\top Q x = 1\}$ for some $Q \succ 0$.

we know that, for the scalar case[29], (3-3.6) must hold. That means we consider for $\widehat{A}_{c\ell} := \widehat{A} + \widehat{B}K^{\star}(\gamma)$ (or for a known B) the set

$$\mathcal{L}(\widehat{A}_{c\ell}) := \left\{ L \,:\, (\widehat{A}_{c\ell} - L) \otimes (\widehat{A}_{c\ell} - L) = (\widehat{A}_{c\ell} + L) \otimes (\widehat{A}_{c\ell} + L) \right\}. \tag{3-3.7}$$

If $\widehat{A}_{c\ell} \neq 0$ then $\mathcal{L}(\widehat{A}_{c\ell}) = 0$, otherwise any appropriately sized L is a member of that set. Thus, a necessary condition for $\mathbb{A}_{\gamma}$ to be an ellipsoid is that $\widehat{A}_{c\ell} = 0$, which almost surely never holds since the required $(\widehat{A}, \widehat{B})$ live on a lower dimensional hyperplane. Although this is unlikely to happen, we see exactly this behaviour in Section 3-4-2-1, where $\widehat{a}$ and $k(\delta)$ cancel each other out for $\delta \to 1$.

Now, for the general case ($n \geq 2$), we compare $\Sigma_x(L^{\star})$ and $\Sigma_x(-L^{\star})$ and claim that $\Sigma(L^{\star}) \succ \Sigma(-L^{\star})$ almost surely. Of course, when $\Sigma_x(-L^{\star})$ is not defined then the set is not ellipsoidal anyway, thus assume the contrary. Let $C^{\top}C := Q$, then assume for now that (A, C) is observable such that we can appeal to Lemma A-0.3 since $Q + \left(K^{\star}(\delta)\right)^{\top} RK^{\star}(\delta) - \delta^{-1}\left(L^{\star}(\delta)\right)^{\top} L^{\star}(\delta) \succ 0$ by Lemma 3.5 from [BB95], which shows that $\Sigma_x(-L^{\star}) \succ \Sigma_x(L^{\star})$ would be in conflict with optimality of $L^{\star}$. Then, $\Sigma_x(L^{\star}) = \Sigma_x(-L^{\star})$ fails to hold almost surely, as discussed before via (3-3.6) and (3-3.7). This leads to our claim. Now, when $(\widehat{A}, \widehat{B})$ is controllable (Q relates to (A, C) being observable already), we know that $P \succ 0$ such that from (3-2.7) we know that $L^{\star} \in \mathsf{GL}(n, \mathbb{R})$ when $\widehat{A} \in \mathsf{GL}(n, \mathbb{R})$. Under these constraints $\langle (L^{\star})^{\top} L^{\star}, \Sigma_x(L^{\star}) - \Sigma_x(-L^{\star}) \rangle$ is almost surely unequal to 0, which concludes the argument since these controllable/observable pairs and the general linear group are dense in their ambient spaces. $\square$

Note that we only used symmetry, such that we can use the same line of thought and show that our set is unlikely to be for example a $\|\cdot\|_2$-ball as well.

Lemma 3-3.9. *Given some $Y \in \mathbb{R}^{n \times n}$ and consider the standard $\|\cdot\|_F$-norm topology on $\mathbb{R}^{n \times n}$, then the set*

$$\mathcal{X} := \left\{ X \in \mathbb{R}^{n \times n} \,:\, X \otimes X = Y \otimes Y \right\} \tag{3-3.8}$$

is nowhere dense in $\mathbb{R}^{n \times n}$.

Proof. For $n = 1$ we simply get $\mathcal{X} = \pm\sqrt{y}$, which is indeed nowhere dense in $\mathbb{R}$. This can be extended to higher dimensions. Let $X' \in \mathcal{X}$, then we show that there does not exist an arbitrary small non-zero perturbation matrix E, such that $X + E \in \mathcal{X}$. Consider the element $(X' \otimes X')_{11} =: x_{11}$, then we must have $x_{11}^2 = y_{11}^2$. Now, any component-wise perturbation e_{11} must satisfy $(x_{11} + e_{11})^2 = y_{11}^2$ such that $e_{11} = 0$ or $e_{11} = -2x_{11}$. Therefore, there does not exist an arbitrary small $\|\cdot\|_F$-norm ball around X', completely living in $\mathcal{X}$. $\square$

So, any feasible ellipsoid of Δ_A, inscribed in $\mathbb{A}_{\gamma}$, is inherently conservative in the sense that the corresponding controller is almost surely synthesized to hedge against a larger set. The crux is that Σ_x is a function of Δ_A and $\Sigma_x(\Delta_A)$ is *a.s.* not constant on any neighbourhood of some Δ_A. Thus, as elegant as the relation between simple dynamic games and (3-3.3) may seem, Lemma 3-3.8 tells us that its practical relevance is questionable. Of course, since the

[29] We cannot immediately generalize this to the general case since $\langle L^{\top}L, \Sigma_x \rangle = \gamma = \langle L^{\top}L, \Sigma_2 \rangle$ does not imply $\Sigma_1 = \Sigma_2$ when $n \geq 2$.

result is *almost surely* true, there are exceptions, and one exception is given in the scalar example of section 3-4-2-3, where the set becomes symmetric for $\delta \to 1$. This should be contrasted with Figure 3-4a, which *appears* to be ellipsoidal, but is not.

Extending this to some uncertainty in the pair (A, B) via the canonical decomposition from section 3-2-1 does not change the conclusion as already indicated via (3-3.4), without loss of generality we can take $\Delta_{A_{c\ell}}$ instead of Δ_A in Lemma 3-3.8. Then, under the assumption that $u_k = Kx_k$ generated V_k, item (i) changes to the *a.s.* non-existence of a $\gamma > 0$ such that for $\Delta_\Theta = \begin{pmatrix}\Delta_A & \Delta_B\end{pmatrix}$, $\{\Delta_\Theta \ : \ \|\Delta_\Theta\|^2_{F,V_k} \leq \gamma\} = \{\Delta_\Theta \ : \ \|\Delta_\Theta\|^2_{F,\Sigma'_x} \leq \gamma\}$ with $V_k = \Sigma'_x$ for

$$\Sigma'_x = \begin{pmatrix} I_n \\ K \end{pmatrix} \Sigma_x \begin{pmatrix} I_n \\ K \end{pmatrix}^\top . \tag{3-3.9}$$

Item (ii) immediately extends by taking $\Delta_A \triangleq \Delta_{A_{c\ell}}$ and writing

$$\Delta_{A_{c\ell}} = \Delta_A + \Delta_B K^\star(\gamma) = \begin{pmatrix}\Delta_A & \Delta_B\end{pmatrix} \begin{pmatrix} I_n \\ K^\star(\gamma) \end{pmatrix}, \tag{3-3.10}$$

from where it can be observed that if the set (3-3.5) is not ellipsoidal in Δ_A it is obviously not ellipsoidal in $\Delta_{A_{c\ell}}$. Moreover, it can be observed that the decomposition of $\Delta_{A_{c\ell}}$ results in a set which is never compact[30] and of infinite volume in the product space for $\Delta_A \times \Delta_B$. So any compact rectangular set for the pair (Δ_A, Δ_B) can never be of the same volume. Hence, inscribed balls remain inherently conservative, even if B enters the picture independently of A.

After the proof of Lemma 3-3.8 a comment is made that our set is almost surely no $\|\cdot\|_2$-ball as well, we do however not elaborate on this.

Now one might say that we should not be concerned with volumes of these sets but with the corresponding costs. Really, we are only conservative when the *actual* uncertainty induces (*a.s.*) a lower cost than the *worst-case* uncertainty. Assume that B is known and our A matrix is given by $\widehat{A} + \Delta_A$ for some unknown Δ_A. How likely is it that this additive uncertainty can be represented as $\widehat{A} + \Delta_A = T\widehat{A}$ for $T \in \mathcal{S}^n_{++}$? This is interesting, since by Lemma 3-3.3 (iii) such a representation is a necessary condition in a our framework.

Lemma 3-3.10. *Consider the standard $\|\cdot\|_F$-norm topology on $\mathbb{R}^{n\times n}$. Then, for $n \geq 2$ the set*

$$\left\{\Delta_A \in \mathbb{R}^{n\times n} \ : \ \exists T \in \mathcal{S}^n_{++} : T\widehat{A} = \widehat{A} + \Delta_A\right\} \tag{3-3.11}$$

is nowhere dense in $\mathbb{R}^{n\times n}$.

Proof of Lemma 3-3.10. This follows immediately from $\mathsf{Sym}(n, \mathbb{R})$ being nowhere dense in $\mathbb{R}^{n\times n}$, plus the fact that an action of any $\widehat{A} \in \mathbb{R}^{n\times n}$, from the right, cannot change this[31]. We will however give a more explicit argument below. First, consider $\det(\widehat{A}) = 0$, then from $(T - I_n)\widehat{A} = \Delta_A$ it is found that the set of all $\Delta_A \in \mathbb{R}^{n\times n}$ satisfying this equation cannot be locally dense since $\det(\Delta_A) = 0$ while $\mathsf{GL}(n, \mathbb{R})$ is open and dense in $\mathbb{R}^{n\times n}$. Now, say

[30] There are several interpretations here, one can resort to the discussion from Section 3-2-1-1 or observe from (3-3.10) that after decomposition, the weighting matrix $\Sigma'_x \in \mathcal{S}^{n+m}_+$ (3-3.9) is inherently rank-deficient.
[31] In fact, recall that $\mathsf{Sym}(n, \mathbb{R}) \simeq \mathbb{R}^{n(n+1)/2}$.

$\det(\widehat{A}) \neq 0$, then $\Delta_A \widehat{A}^{-1}$ must be at least symmetric. Assume that symmetry holds for the pair $(\Delta'_A, \widehat{A}')$. We will show that an arbitrary small perturbation in Δ'_A can destroy the symmetry. Assume that the ij^{th} element of $\widehat{A}'^{-1}$ is nonzero and consider an arbitrarily small matrix-perturbation $\Delta_A + E$ with $E = \varepsilon e_i e_j^\top$, $i \neq j$, for some arbitrarily small $\varepsilon > 0$. Under such a perturbation $(\Delta'_A + E)\widehat{A}'^{-1}$ cannot be symmetric. Hence, even in this case the existence of any interior around Δ_A is excluded since a valid (set of) T's does not exist. Note that this argument cannot be made when $n = 1$ (in that case, feasible intervals do exist). $\square$

Lemma 3-3.10 implies that it is very unlikely that some additive perturbation coincides with a worst-case uncertainty as derived from a dynamic game. Differently put, given the observed ***behaviour***[32] of some partially unknown dynamical system (2-1.1), then the mean state process can almost surely not be explained from a game theoretic point of view. Therefore, if one solves a feasible robust LQR problem under this framework, then the solution is *almost surely* conservative[33]. Then, again, extending the result to include B does not change the conclusion since $\Delta_{A_{c\ell}} = \Delta_A + \Delta_B K$ constrains the pair (Δ_A, Δ_B) to some lower dimensional set. However, $\partial \mathbb{A}_\gamma$ is also a lower dimensional set, which by Lemma 3-3.3 contains all worst-case models, so the fact that the set is simply lower-dimensional is not immediately negative. Nevertheless, looking at the corresponding argument we see that this idealized scenario can only occur for $n = 2$, since then $n^2 - 1 = n(n+1)/2$. Hence, especially for $n > 2$, our worst-case models live on a geometrically negligible set.

Remark 3-3.11 (More on symmetry). *The reader might think that the previous result holds in part due to Q and R being symmetric. What if they are merely positive (semi)-definite and not symmetric? Indeed, then P is not necessarily symmetric. However, any of these non-symmetric cost matrices lead to a quadratic form which can always be parametrized by a symmetric matrix since $x^\top Q x = \frac{1}{2} x^\top (Q + Q^\top) x$. Thus, no new pairs of $(K^\star, L^\star)$, in the sense of Lemma 3-2.9, are introduced. Finding this symmetric representation is however not even needed, one can simply alter the Riccati equation, which becomes a bit more involved. The point is, we can assume without loss of generality that Q and R are symmetric.*

Someone might say, given a $\widehat{A}$, we just saw that we cannot easily create any desirable A, but what if we have some clever "worst-case" identification algorithm which selects $\widehat{A}$ such that the worst-case model is A? Well, due to the relation between $\widehat{A}$ and $A^\star(\gamma)$, as just examined, it follows that the set of $\widehat{A}$ obeying this requirement is of measure zero, putting rather impossible requirements on the identification algorithm.

This section assumed that $D = I_n$, which is the most natural setting when one has no further information about the structure of the uncertainty. Including D in the *system-theoretic* framework as an optimization variable is an interesting future problem.

Item (iii) from Lemma 3-3.3 has yet another different implication. Say, someone wonders if the unknown perturbation Δ_A in $x_{k+1} = (\widehat{A} + \Delta_A + BK)x_k$, $x_0 \sim \mathcal{P}(0, \Sigma_0)$ could have been generated by a game theoretic adversary? Gather the data of E episodes, individually

[32] Recall (*cf.* [PW98]), we can define a dynamical system by the tuple $\Sigma := (\mathbb{T}, \mathbb{W}, \mathfrak{B})$, where the behaviour $\mathfrak{B} \subset \mathbb{W}^{\mathbb{T}} := \{f : f : \mathbb{T} \to \mathbb{W}\}$, *e.g.* in the case of discrete LTI systems the behaviour is defined by $\mathbb{T} := \mathbb{Z}_{\geq 0}$ and some transition matrix $A_{c\ell}$ giving rise to sequences (the behaviour) the system can attain.

[33] On a more intuitive level in the context of GANs [GPM+14], where game theory is heavily used, this would imply that if one wants to generate pictures of animals, the algorithm only gives you pictures of sausage dogs.

running for N timesteps, *i.e.*, $\{x_k^{(e)}\}_{k=0}^N$, $e = 1, \dots, E$. Then a simple necessary condition for this data to be generated by a dynamic game, or worst-case uncertainty in A, is that the convex[34] program

$$\inf_{T \in \mathcal{S}_{++}^n} \quad \frac{1}{EN} \sum_{e=1}^{E} \sum_{k=0}^{N-1} \left\| x_{k+1}^{(e)} - T\widehat{A}x_k^{(e)} - BKx_k^{(e)} \right\|_2^2$$

attains a cost arbitrary[35] close to 0. Note, this condition does not depend on γ, making it very easy to check in practice.

3-4 Analytical and Numerical Experiments

In this section we provide ample analytic and numerical examples, showing, for the better or worse, the rich structure this basic game theoretic formulation already displays.

3-4-1 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 section 3-4-2-2 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

$$\begin{aligned} \sup_{\delta \in \mathbb{R}_{\geq 0}} \quad & \delta \\ \text{subject to} \quad & \delta^{-1} I_d - \alpha D^\top P(0) D \succeq 0 \end{aligned} \tag{3-4.1}$$

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 3-2.9), $P(0)$ exists, such that the solution to (3-4.1) is given by $\delta^\star = \|\alpha D^\top P(0) D\|_2^{-1}$. Of course, for a meaningful bound we must assume that $DP(0) \neq 0$. Also, $\delta^\star$ is not necessarily feasible (see Figure 3-9 (1)). 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 3-9.

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.

[34]Recall that we can formulate $\inf_{A \in \mathbb{R}^{n\times n}} \|A - A'\|_F^2$ as

$$\begin{aligned} \inf_{A \in \mathbb{R}^{n\times n}, B \in \mathcal{S}_+^n} \quad & \mathrm{Tr}(B) \\ \text{s.t.} \quad & \begin{pmatrix} I & (A - A') \\ (A - A')^\top & B \end{pmatrix} \succeq 0 \end{aligned}$$

. So indeed, we can easily impose the constraint $T \succ 0$.

[35]Demanding the cost to be identically 0 can of course be numerically challenging in practice, *i.e.*, needing perfect state measurements.

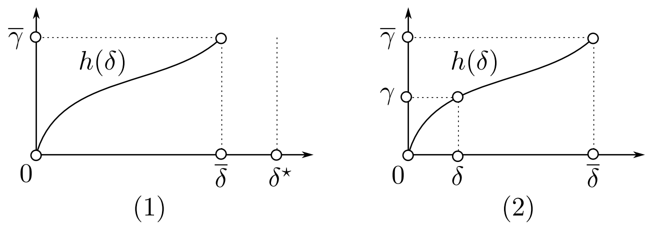


Figure 3-9: Let us be given a Robust LQR problem for some γ. First (1), compute $\delta^\star$ from (3-4.1) and find $\overline{\delta}$ using any algorithm similar to bisection. As a byproduct, $\overline{\gamma}$ is given such that feasibility of γ can be readily checked. Note that in practice one 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 $\gamma\leq\overline{\gamma}_\varepsilon$.

Lemma 3-4.1 (Finding δ)**.** *Given a desired γ and assume it is feasible in the sense of Theorem 3-2.4. Let the (local) Lipschitz constant of the map h be $L>0$ on $[0,\delta)$ and select $\beta\leq L^{-1}$. Then, the algorithm*

$$\delta_{k+1}=\delta_k+\beta(\gamma-h(\delta_k)),\quad \delta_0=0, \tag{3-4.2}$$

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

Proof of Lemma 3-4.1. Consider the algorithm

$$\delta_{k+1}=\delta_k+\beta_k(\gamma-h(\delta_k)),\quad \delta_0=0. \tag{3-4.3}$$

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 $L>0$

$$|h(\delta_2)-h(\delta_1)|\leq L|\delta_2-\delta_1|,\quad \delta_1,\delta_2\in[0,\overline{\delta}). \tag{3-4.4}$$

Therefore, by (3-4.4) 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 $e_{k+1}=e_k-\widehat{L}^{-1}(\gamma-h(\delta_k))=(1-\varepsilon)e_k$, for some $\varepsilon\in(0,1]$ such that, the cruder $\widehat{L}$ is, the smaller ε and thus the slower $e_{k+1}\to 0$. $\square$

In the light of figures 3-10a and 3-10b we do emphasize that estimation of the Lipschitz constant is critical the closer $\delta : h(\delta)=\gamma$ is to $\overline{\delta}$.

At last, we make a brief remark on how the Generalized Algebraic Riccati Equation can be solved.

3-4-1-1 Solving the Generalized Algebraic Riccati Equation (GARE)

There are several methods to obtain the minimal solution to the GARE (3-2.17), but a particularly simple method is to iteratively compute

$$P_{k+1}(\delta) = Q + \alpha \widehat{A}^{\top} P_k(\delta)\left(\Lambda(\delta)\right)^{-1}\widehat{A}, \quad P_0 = Q. \tag{3-4.5}$$

until $\|P_{k+1}(\delta) - P_k(\delta)\|_p \leq \varepsilon_P$[36] for some $\varepsilon_P > 0$. Of course, (3-4.5) only converges under certain conditions, as expressed in Lemma 3-2.9. Moreover, although (3-4.5) is simple, it is not straight forward to find a ε_P which implies that the corresponding P_{k+1} will lead to a $\sqrt{\alpha}$-stable controller, since theoretically one needs $P_\infty = \lim_{k\to\infty} P_{k+1}$ instead of $P_{M<\infty}$. Different methods to solve the GARE exist, but it is not clear to the authors which method is superior, possibly as function of the problem size. For more information, see [SW94, BB95].

3-4-2 Scalar Examples

Here we study a few scalar examples to provide more insights in the structure of the control law, limits of the map h and a small hint at the existence of concentration inequalities.

3-4-2-1 Scalar Control Gains

In the first part we compare the *full* admissible intervals of uncertainties in the system matrix A, for LQR and RLQR scalar control gains. Although these scalar systems have closed-form solutions, it will be shown they are by no means simple and elegant. Consider the (nominal) optimal control problem:

$$\inf_{k\in\mathbb{R}} \mathcal{J}(1+k, 1+k^2)$$

for $\alpha = \sigma_0 = \sigma_v = 0.5$[37]. For this simple problem the positive solution to the corresponding Riccati equation is $p = \sqrt{2}$ and the optimal control gain is given by $k^\star(0) = 1 - \sqrt{2}$. Thereby, the closed-loop mean of x_{k+1} is given by $(2-\sqrt{2})x_k$, which is $\sqrt{\alpha}$-stable. Moreover, when considering only the system matrix, this controller can stabilize all the uncertainties $\Delta_a \in (-2, -2+2\sqrt{2})$. At last we compute the LQ cost as a function of the control gain and the nominal system matrix $(k, \widehat{a})$, which is given by

$$\mathcal{J}(\widehat{a}+k, 1+k^2) = \frac{1+k^2}{1-\frac{1}{2}(\widehat{a}+k)^2} \implies \mathcal{J}(\widehat{a}+k^\star(0), 1+\left(k^\star(0)\right)^2) = \frac{2(2-\sqrt{2})}{1-\frac{1}{2}(\widehat{a}+1-\sqrt{2})^2}.$$

Thus, positive perturbations to $\widehat{a} = 1$, increase $\mathcal{J}$ exponentially fast.

The next part shows that the robust control framework anticipates on this last observation. Consider the dynamic game

$$\inf_{k\in\mathbb{R}} \sup_{\ell\in\mathbb{R}} \mathcal{J}(1+k+\ell, 1+k^2-\delta^{-1}\ell^2) \tag{3-4.6}$$

[36] Or until some relative error is small.

[37] To keep the notation simple and in line with the previous sections we will refer to σ_v as the variance of v instead of the more common σ_v^2 notation.

for again $\alpha = \sigma_0 = \sigma_v = 0.5$. A series of algebraic manipulations reveal that the minimal solution to the GARE is given by the root:

$$p(\delta) = \frac{-\frac{1}{2}\delta + \sqrt{2 - 2\delta + \frac{1}{4}\delta^2}}{1-\delta}. \tag{3-4.7}$$

See that $\lim_{\delta\downarrow 0} p(\delta) = \sqrt{2}$ indeed and that $\overline{\delta} = 1$ since for $\delta \to 1$ we have $\delta^{-1} - \alpha p(\delta) \to 0$. Also when we consider the worst-case closed-loop matrix $a_{c\ell}(\delta) = \lambda^{-1}(\delta)\widehat{a} = (1 - 1/4\delta + 1/2\sqrt{2 - 2\delta + 1/4\delta^2})^{-1}$, it can again be observed that $\lim_{\delta\downarrow 0} a_{c\ell}(\delta) = (1 + 1/2\sqrt{2}) = 2 - \sqrt{2}$ indeed. Then at last, the controller as parametrized by δ is given by

$$k(\delta) = -\frac{-\frac{1}{2}\delta + \sqrt{2 - 2\delta + \frac{1}{4}\delta^2}}{(1-\delta)\left(2 - \frac{1}{2}\delta + \sqrt{2 - 2\delta + \frac{1}{4}\delta^2}\right)}$$

such that again $\lim_{\delta\downarrow 0} k(\delta) = -\sqrt{2}(2+\sqrt{2})^{-1} = 1 - \sqrt{2}$. Obviously, this controller can stabilize the additive uncertainties $\Delta_a \in (-1 - k(\delta) - \sqrt{2}, -1 - k(\delta) + \sqrt{2})$. To see how this set behaves under $\delta \in [0, \overline{\delta})$ we compute $\lim_{\delta\uparrow\overline{\delta}} k(\delta) = -1$. This limiting controller can be interpreted as the most pessimistic, assuming big trouble, the most safe location is to render the nominal closed-loop system matrix 0.

Then similar to the controller, we can find a closed-form expression for the worst-case disturbance

$$\ell(\delta) = \frac{-\frac{1}{2}\delta^2 + \delta\sqrt{2 - 2\delta + \frac{1}{4}\delta^2}}{(1-\delta)\left(2 - \frac{1}{2}\delta + \sqrt{2 - 2\delta + \frac{1}{4}\delta^2}\right)}. \tag{3-4.8}$$

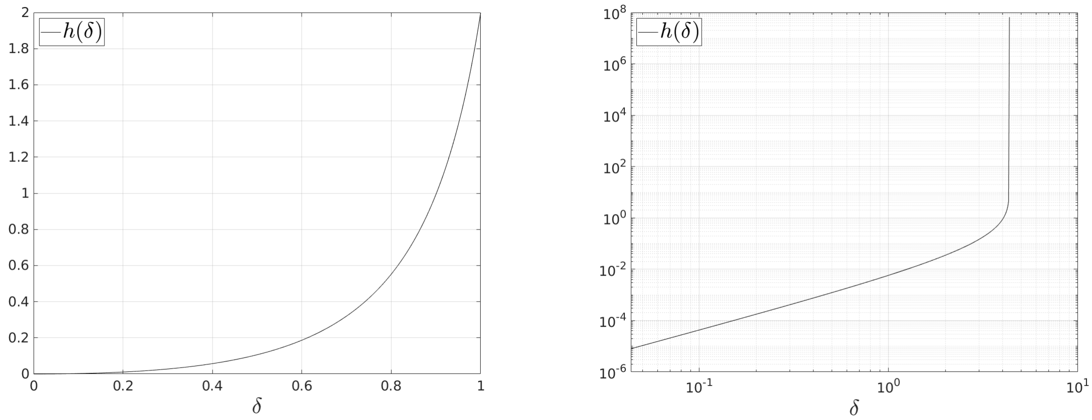
Again, the limiting cases are interesting, by expansion we find that $\lim_{\delta\downarrow 0} \ell(\delta) = 0$ and $\lim_{\delta\uparrow\overline{\delta}} \ell(\delta) = 1$. Of course, since $\widehat{a} = 1$, positive perturbations are the cheapest method to destabilize the system, you pay for the perturbation.

Finally, given all these expressions, we can find the closed-form solution to the discrete-time Lyapunov equation, which is simply $\sigma_x(\delta) = \left(1 - \alpha a_{\mathrm{cl}}^2(\delta)\right)^{-1}$. This allows for plotting the map $h(\delta) = \ell^2(\delta)\sigma_x(\delta)$. The graph of $(\delta \in (0,1), h(\delta))$ is shown in Figure 3-10a. We observe strictly monotonic and smooth behaviour, but also a clear limit: $\lim_{\delta\uparrow 1} h(\delta) = 2$, which is hardly a surprise when looking at (3-4.7). To further study the map $h(\delta)$ we can look at its closed-form expression:

$$h(\delta) = \frac{\delta^3}{(\delta-1)^2(\delta-4)} + \frac{4\delta^2(4-3\delta)}{(\delta-1)^2(\delta-4)\left((\delta-4)\sqrt{\delta^2 - 8\delta + 8} - \delta^2 + 8\delta - 8\right)}. \tag{3-4.9}$$

We observe from (3-4.9) that, as seen before, $h(1)$ is not well defined, plus for $\delta > 4 - \frac{1}{2}\sqrt{32} > 1$ the expression becomes complex-valued. So, even for such a simple dynamical system the expression for $h(\delta)$ is very involved and the authors do not know of a local inverse. Therefore, algorithms like presented in Lemma 3-4.1 are rather a necessity.

Now it might seem odd that $\lim_{\delta\uparrow\overline{\delta}} h(\delta)$ can have a finite limit since $\widetilde{h}(\delta)$ from (3-2.5) cannot. This observation can however be justified.



(a) Given the parameters from section 3-4-2-1 we show the map $h(\delta)$ over $(0, 1)$, which is indeed smooth, monotone and apparently bounded.

(b) Adjusting the parameters from section 3-4-2-1 to $\widehat{a} = 0.5$ and $q = 0.25$ we observe different limiting behaviour (unbounded).

Figure 3-10: The map h from Theorem 3-2.4 can have finite or infinite limiting values when approaching the breakdown point $\bar{\delta}$.

3-4-2-2 Limits of the Map h

In the previous section it was observed that $\lim_{\delta\uparrow\bar{\delta}} h(\delta) = 2$. More formally, instead of looking at (3-4.9), which is quite involved, we look at the definition of $h(\delta)$. Then, $\lim_{\delta\uparrow\bar{\delta}} h(\delta) = \ell^2(\delta)\left(1-\alpha a_{\mathrm{cl}}^2(\delta)\right)^{-1} = 2$ since $\alpha = 0.5$, $\lim_{\delta\uparrow\bar{\delta}} \ell(\delta) = 1$ and $\lim_{\delta\uparrow\bar{\delta}} a_{\mathrm{c}\ell}(\delta) = 1$. Thus, we see that although problem (2-1.2) is well-defined for all $\gamma \in \mathbb{R}_{\geq 0}$, once $\gamma > 2$ then the corresponding RLQR problem to the game (3-4.6) is not well-defined in *our solution framework*.

What is the interpretation of this limiting value, and does it always exist? One might expect that once you approach $\bar{\delta}$ the closed-loop spectrum approaches the boundary of $\mathbb{D}_{\alpha^{-1/2}}$ from inside. As formally proven in section 3.8 of [BB95] for a closely related $\mathcal{H}_\infty$ problem, if the limiting controller $\lim_{\delta\uparrow\bar{\delta}} K^\star(\delta)$ exists, then the nominal closed-loop system is bounded-input bounded-state (BIBS) stable. However, we cannot conclude anything regarding the worst-case closed-loop system, which is precisely the term we use in the definition of $h(\delta)$. Therefore it seems hard to say anything about $\lim_{\delta\uparrow\bar{\delta}} h(\delta)$.

To emphasize this, we update example system (3-4.6) to an example like proposed on page 93 of [BB95], and simply change $\widehat{a}$ to 0.5 and q to 0.25. Now we can again plot $h(\delta)$ (see Figure 3-10b). Indeed, this time $\lim_{\delta\uparrow\bar{\delta}} h(\delta) = \infty$, which can be explained from the fact that the worst-case closed-loop system converges to $\alpha^{-1/2}$, thus becoming unstable in the $\sqrt{\alpha}$-sense. This should be contrasted with 3-4-2-1, where the limiting worst-case closed-loop system was strictly $\sqrt{\alpha}$-stable.

So, to understand the limit, we could just investigate the limiting spectral radius of the worst-case closed-loop system: $\lim_{\delta\uparrow\bar{\delta}} \rho(\Lambda^{-1}(\delta)A)$. We will implicitly do this by looking (semi-formally) at the boundedness of the worst-case cost.

First, it is easy to see that $\delta^{-1}I_d - \alpha D^\top PD \succeq 0$ is a necessary condition for a dynamic game to be well-defined (bounded cost). For example, consider solving the robust Bellman (or Isaacs) equation

$$V(x) = \inf_u \sup_w \left(x^\top Q x + u^\top R u - \delta^{-1} w^\top w + \alpha \mathop{\mathbb{E}}_{x_0,v} \left[V(x') | x \right] \right).$$

To solve this equation, assume that $V(x) = x^\top P x + q$, then we find

$$x^\top P x + s = \inf_u \sup_w \begin{pmatrix} x \\ w \\ u \end{pmatrix}^\top \begin{pmatrix} Q + \alpha A^\top P A & \alpha A^\top P D & \alpha A^\top P B \\ \alpha D^\top P A & -\delta^{-1} I_d + \alpha D^\top P D & \alpha D^\top P B \\ \alpha B^\top P A & \alpha B^\top P D & R + \alpha B^\top P B \end{pmatrix} \begin{pmatrix} x \\ w \\ u \end{pmatrix}, \tag{3-4.10}$$

for s the part containing q. A necessary condition to keep the right-hand side of (3-4.10) bounded under extremization over w is that $\delta^{-1} I_d - \alpha D^\top P D \succeq 0$ holds. This conditions gives rise to an interval of valid δ, specifically, $\delta \in [0, \bar{\delta}]$, for $\bar{\delta}$ the "*breakdown point*" (see section 3-2-2). A sufficient condition would follow from making the inequality strict, *i.e.* $\delta \in [0, \bar{\delta})$. Thus, the most interesting value for δ is this breakdown point $\bar{\delta}$, what happens over there? To answer this question, interpret the inner maximization of (3-4.10) as a QP for some $J \succeq 0$ and $H = H^\top$:

$$\operatorname*{argmax}_w \begin{pmatrix} z \\ w \end{pmatrix}^\top \begin{pmatrix} G & H \\ H^\top & J \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} = \operatorname*{argmax}_w w^\top J w + 2 z^\top H w. \tag{3-4.11}$$

Then another condition (in combination with $J \succeq 0$ necessary and sufficient) for (3-4.11) to have a bounded cost is that $J w^\star = -H^\top z$[38]. When written in the notation of (3-4.10), with additionally the dependence of P on δ, this becomes:

$$\big(-\delta^{-1} I_d + \alpha D^\top P(\delta) D \big) w_k^\star = -\alpha D^\top P(\delta)(A x_k + B u_k).$$

First of all, spot the link with (3-2.13) from Theorem 3-2.4. We know that the optimizing input will be $u_k^\star = K^\star(\delta) x_k$ and $w_k^\star = L^\star(\delta) x_k$ such that the question becomes: "Does

$$\lim_{\delta \uparrow \bar{\delta}} \big(-\delta^{-1} I_d + \alpha D^\top P(\delta) D \big) L^\star(\delta) x_k = \lim_{\delta \uparrow \bar{\delta}} -\alpha D^\top P(\delta) (A + B K^\star(\delta)) x_k \tag{3-4.12}$$

hold?" The crux is that the left-hand side might lower its rank at $\bar{\delta}$. Now consider again the example from 3-4-2-1, indeed, there we had $\lim_{\delta \uparrow \bar{\delta}} (a + b k^\star(\delta)) = 0$ and in fact both sides of (3-4.12) converge to 0 such that even *at* the breakdown point the cost is bounded. However, for the other example, with $a = 0.5$ and $q = 0.25$ this conditions fails and the cost becomes unbounded. So if (3-4.12) is satisfied, player 1 cannot be fooled. In other words, to get $\lim_{\delta \uparrow \bar{\delta}} h(\delta) = \infty$ the cost must approach ∞ since $h(\delta)$ is a function of $\Sigma_x := \mathop{\mathbb{E}}_{x_0,v} \sum_{k=0}^{\infty} x_k x_k^\top$. However, we see that we can relate the question of unboundedness to a *degenerate* QP being a function of δ, the cost matrices (Q, R) but also of the dynamical system Σ (2-1.1). It turns out that some problem parameters maintain a well-defined QP, even for $\delta \to \bar{\delta}$ such that the cost becomes unbounded, not *at* $\bar{\delta}$, but *beyond* $\bar{\delta}$, making $\lim_{\delta \uparrow \bar{\delta}} h(\delta) < \infty$. This has an important implication, namely that although $K^\star(\gamma)$ might exist for all $\gamma \in \mathbb{R}_{\geq 0}$ we cannot

[38] See appendix B-1, or consider the QP: $\inf_x \ \frac{1}{2} x^\top Q x + r^\top x$ for $Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \succeq 0$ and two options of r, $r_1^\top = \begin{pmatrix} 1 & 0 \end{pmatrix}$, $r_2 = \begin{pmatrix} 0 & 1 \end{pmatrix}$. Under r_2 the cost is unbounded, while under r_1 we have $x^\star = -r_1$ and indeed $Q x^\star = -r_1$. In our context both Q and r are parametrized by δ, the cost matrices and the dynamics.

always find it via our game theoretic framework. We are restricted to $\gamma \in [0, \overline{\gamma})$, $\overline{\gamma} := h(\overline{\delta})$ which *might* coincide with $\mathbb{R}_{\geq 0}$, but not for each problem. Schematically, one can think of Figure 3-11, if $\overline{\gamma} < \infty$ then in the current framework we cannot utilize the full potential of linear control, we never hedge against the full set some K can $\sqrt{\alpha}$-stabilize. Moreover, this implies that if $\lim_{\delta \uparrow \overline{\delta}} K^\star(\delta)$ stabilizes (A, B) then we cannot be *sure* that this pair is part of our uncertainty set.

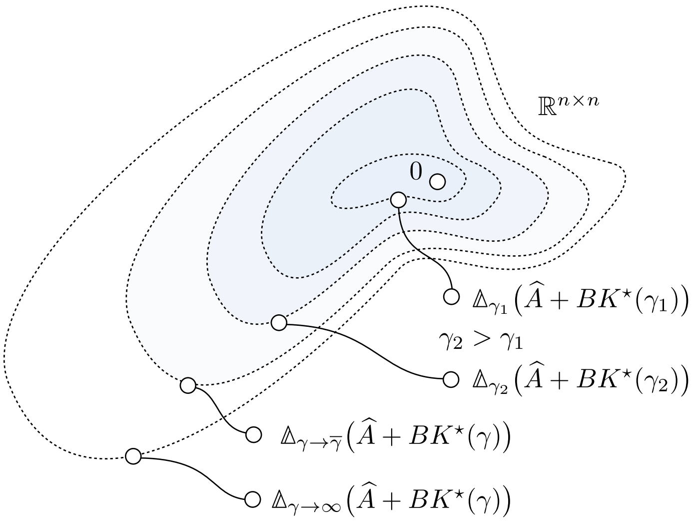


Figure 3-11: Although $\mathbb{A}_{\gamma \to \infty}(\widehat{A} + BK^\star(\gamma))$ is well-defined, Theorem 3-2.4 only gives us a controller which can provably $\sqrt{\alpha}$-stabilize the set $\mathbb{A}_{\gamma \to \overline{\gamma}}(\widehat{A} + BK^\star(\gamma))$, which is equal or *smaller*.

Now, one can derive all sorts of conditions from (3-4.12), we do not embark on this and refer the reader to a somewhat similar discussion, but then for the entropy interpretation, see [HS07, ch.8].

3-4-2-3 Uncertainty Sets

To efficiently start the discussion scalar uncertainty sets, we first consider a generic example.

Example 3-4.2 (1D $\mathcal{A}_\gamma$). *Consider a 1D version of $\mathcal{A}_\gamma$ where $d = 1$, $\alpha \in (0, 1)$ and $\widehat{a}$ is $\sqrt{\alpha}$-stable, plus, without loss of generality we simplify the Lyapunov equation to $\sigma_x = \alpha a_{c\ell}^2 \sigma_x + w$, $w > 0$, $a_{c\ell} = \widehat{a} + \Delta_a$. Then it follows that the feasible uncertainties Δ_a are parametrized by the map $f_1 : \mathbb{R} \to \mathbb{R}$:*

$$\mathbb{A}_\gamma(\widehat{a}) = \left\{ \Delta_a \in \mathbb{R} : -(1 + \alpha\gamma/w)\Delta_a^2 - 2\alpha(\gamma/w)\widehat{a}\Delta_a + (\gamma/w)(1 - \alpha\widehat{a}^2) = f_1(\Delta_a) \geq 0. \right\}.$$

Here, the constraint $1 - \alpha(\widehat{a}^2 + 2\widehat{a}\Delta_a + \Delta_a^2) = f_2(\Delta_a) > 0$ is implicit. Since both f_1 and f_2 have $\partial^2 f_i < 0$ we can purely focus on their respective roots since they define the endpoints of the feasible connected interval $\mathbb{A}_\gamma(\widehat{a})$. The function f_2 relates to $\sqrt{\alpha}$-stability such that its feasible interval is given by $(-\alpha^{-1/2} - \widehat{a}, \alpha^{-1/2} - \widehat{a}) \supseteq \mathbb{A}_\gamma(\widehat{a})$. The roots of f_1 are given by

$$\Delta_a^{(1,2)}(\gamma, w, \widehat{a}, \alpha) = \frac{(\gamma/w)\alpha\widehat{a} \pm \sqrt{(\gamma/w)(1 + \alpha\gamma/w - \alpha\widehat{a}^2)}}{-(1 + \alpha\gamma/w)}. \tag{3-4.13}$$

Here we impose the order $\Delta_a^{(1)} \leq \Delta_a^{(2)}$. *It can be observed that the ratio* γ/w *determines if* $\Delta_a^{(1,2)}$ *is real- or complex-valued, which happens when* $\alpha\widehat{a}^2 > 1+\alpha\gamma/w$. *Now by construction*[39] *of* (f_1, f_2) *we have* $[\Delta_a^{(1)}, \Delta_a^{(2)}] \in (-\alpha^{-1/2} - \widehat{a}, \alpha^{-1/2} - \widehat{a})$ *such that* $\Delta_a^{(1,2)}$ *completely defines* $\mathcal{A}_\gamma(\widehat{a})$. *Furthermore, as expected* $\lim_{\gamma\to\infty} \Delta_a^{(1,2)} = -\widehat{a} \pm \alpha^{-1/2}$. *Moreover,* $\mathrm{Vol}(\mathcal{A}_\gamma(\widehat{a})) = 2\sqrt{(\gamma/w)(1+\alpha\gamma/w - \alpha\widehat{a}^2)}(1+\alpha\gamma/w)^{-1}$ *such that* $\lim_{\gamma\to\infty} \mathrm{Vol}(\mathcal{A}_\gamma(\widehat{a})) = 2\alpha^{-1/2}$.

Now instead of the entire set of admissible uncertainties for some controller $k(\delta)$, we will compute $\mathbb{A}_\gamma(\widehat{a}+bk^\star(\delta))$, which is the uncertainty we can provably hedge against via Theorem 3-2.4. To that end, consider the scalar robust LQR problem:

$$\inf_{k\in\mathbb{R}} \sup_{a+k\in\mathcal{A}_\gamma(1+k)} \mathcal{J}(a+k, 1+k^2)\,. \tag{3-4.14}$$

Let $(\alpha, \sigma_0, \sigma_v)$ be such that this optimization problem directly relates to the game (3-4.6), therefore $\gamma \in [0,2)$. Section 3-4-2-1 provides us with expressions for $(p(\delta), k(\delta), \ell(\delta), h(\delta))$. Then to construct the uncertainty set $\mathbb{A}_\gamma(1+k(\gamma))$ we use (3-4.13) and obtain the boundary points:

$$\Delta_a^{(1,2)}(\gamma) = \frac{\frac{1}{2}\gamma(1+k(\gamma)) \pm \sqrt{\gamma[1+\frac{1}{2}\gamma - \frac{1}{2}(1+k(\gamma))^2]}}{-(1+\frac{1}{2}\gamma)}. \tag{3-4.15}$$

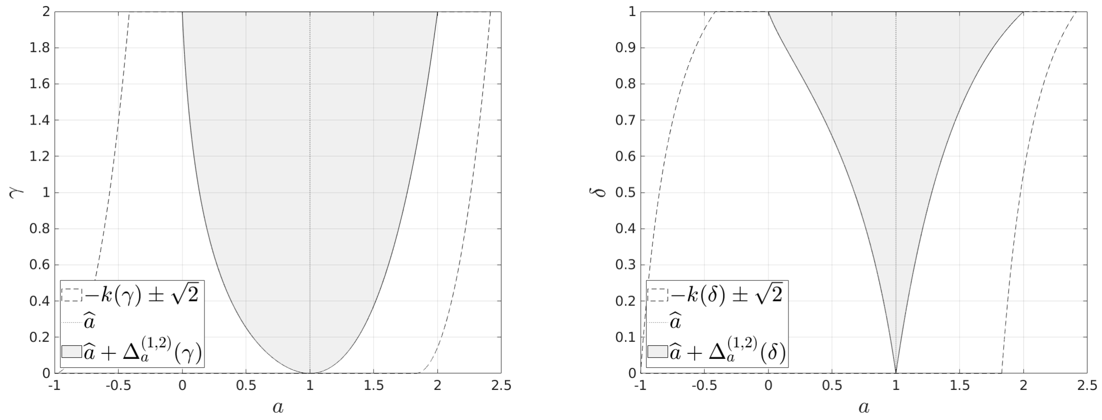
Of course, instead of working with γ, we can work with δ directly via $\gamma = h(\delta)$ and (3-4.9). Then in Figure 3-12a we plot the interval $(\widehat{a} + \Delta_a^{(1)}(\gamma), \widehat{a} + \Delta_a^{(2)}(\gamma))$ of admissible systems for $\gamma \in [0,2)$ and $\widehat{a} = 1$. In Figure 3-12b the set is parametrized directly by δ. Indeed, we observe the limiting case $\lim_{\delta\uparrow\bar{\delta}} \Delta_a^{(1,2)}(\gamma) = \pm 1$. It is interesting to spot the difference between the two parametrizations, this can be explained from the exponential behaviour seen in Figure 3-10a. The figure indicates that exponentially fast updating schemes for γ result in approximately linear growth of the uncertainty set. Also, as predicted, it turns out that right-most line $(\widehat{a} + \Delta_a^{(2)}(\gamma))$ corresponds to the worst-case uncertainties (increasing a is the cheapest). Moreover, to emphasize how our set grows over γ, we show the full set of possible models a which can be stabilized with $k(\gamma)$ (see $k(\gamma) \pm \sqrt{2}$). See that for this particular example we do *not* converge to this boundary, which relates to our prior discussion on limits (see section 3-4-2-2) and indeed is an example of Figure 3-11.

3-4-2-4 Remarks on Model Concentration Bounds

Consider a scalar linear dynamical system and assume that someone is given $b = 1$, but is unsure about a. It is only known that (a,b) can be stabilized by $k^{(i)}(\overline{\gamma}) := \lim_{\delta\uparrow\bar{\delta}} k^{(i)}(\delta)$. Define the set $\mathcal{K}_\alpha^{(i)} := \{a \in \mathbb{R} \;:\; |a + k^{(i)}(\overline{\gamma})| < \alpha^{-1/2}\}$ such that $a \in \mathcal{K}_\alpha^{(i)}$. Then recall the our controlled uncertainty sets are nested (see Lemma 3-2.6), such that the set $\mathbb{A}_\gamma$ under $k^{(i)}(\overline{\gamma})$ is the largest set we can provably hedge against. Now, given some nominal $\widehat{a}$, we can compute the probability, as a function of γ, that our system with unknown a can be provably stabilized via Theorem 3-2.4

$$\mathbb{P}\left\{a+k(\gamma) \in \mathcal{A}_\gamma(\widehat{a}+k(\gamma))\right\}(\gamma) := \frac{\mathrm{Vol}\left(\mathcal{A}_\gamma(\widehat{a}+k(\gamma)) \cap \mathcal{K}_\alpha\right)}{\mathrm{Vol}(\mathcal{K}_\alpha)}(\gamma) = \frac{\sqrt{\alpha}}{2}\mathrm{Vol}\left(\mathcal{A}_\gamma(\widehat{a}+k(\gamma))\right)(\gamma). \tag{3-4.16}$$

[39] f_1 follows from $\Delta_a^2 \leq \gamma/w(1-a^2-2a\Delta_a - \Delta_a^2)$ subject to $f_2(\Delta_a) > 0$.



(a) The set $\widehat{a} + \mathbb{\Delta}_\gamma$ over $\gamma \in [0, 2)$.

(b) The set $\widehat{a} + \mathbb{\Delta}_\gamma$ parameterized by δ

Figure 3-12: Given the parameters from section 3-4-2-3, we show the admissible system matrices a for program (3-4.14), parametrized by either γ or δ.

.

Example 3-4.2 give us the expression to compute (3-4.16). We do this for the case where $(\widehat{a}, q) = (1, 1)$, defining controller $k^{(1)}(\gamma)$, and for the case where $(\widehat{a}, q) = (0.5, 0.25)$, defining controller $k^{(2)}(\gamma)$. Computing the limits provide us with fundamental probabilistic bounds. First, for $k^{(1)}(\gamma)$ we know that $\overline{\gamma} = 2$ such that:

$$\lim_{\gamma \uparrow \overline{\gamma}} \frac{\sqrt{\alpha}}{2} \mathrm{Vol}\left(\mathcal{A}_\gamma(1 + k^{(1)}(\gamma))\right)(\gamma) = \sqrt{\alpha} < 1.$$

Thus, when it is only known that $a \in \mathcal{K}_\alpha^{(1)}$, our framework can never provide higher probabilistic bounds than $\sqrt{\alpha}$, $\alpha \in (0, 1)$. For the second controller this is different since $\overline{\gamma} = \infty$ and

$$\lim_{\gamma \to \infty} \frac{\sqrt{\alpha}}{2} \mathrm{Vol}\left(\mathcal{A}_\gamma(0.5 + k^{(2)}(\gamma))\right)(\gamma) = 1.$$

Of course, these limits are implications of our discussion in section 3-4-2-2. For both controlled systems we can plot (3-4.16), which is done in Figure 3-13. Here we observe the theoretical limit under $k^{(1)}$, but overall *log-normal* behaviour.

Note that this logarithmic scale was already presented subconsciously in for example Figure 3-2a. Moreover, this behaviour can be expected from the close relation to the following problem. Given some $\sigma \in \mathbb{R}_{>0}$ we can consider a LQ problem with exponential cost (3-4.17), or exponential utility function if you like. It turns out that this problem is closely related to a special[40] LQR problem:

[40] The idea is that *Certainty Equivalence* control laws (sometimes rightfully called *Naive Feedback Controllers* (sec. 5.4 [Ber76])), like standard LQR, neglect the noise intensity, whereas the LEQR formulation (3-4.17) tries to take this covariance directly into account.

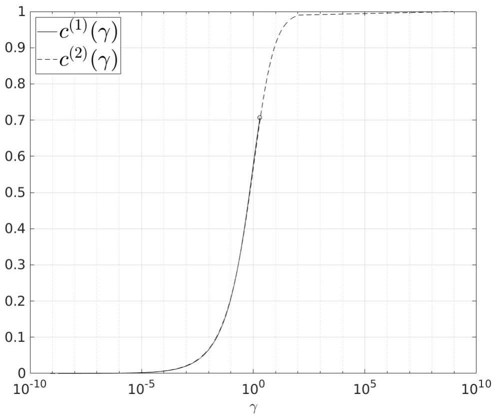


Figure 3-13: The function (3-4.16) of γ for the examples in section 3-4-2-4. Here, $c^{(1)}(\gamma) := \mathbb{P}\left\{a+k^{(1)}(\gamma)\in\mathcal{A}_{\gamma}(1+k^{(1)}(\gamma))\right\}(\gamma)$ and $c^{(2)}(\gamma) := \mathbb{P}\left\{a+k^{(2)}(\gamma)\in\mathcal{A}_{\gamma}(0.5+k^{(2)}(\gamma))\right\}(\gamma)$. Note that formally speaking these are not per se CDFs since $\int_{\mathbb{R}_{\geq 0}} c^{(1)}(\gamma)d\gamma < 1$.

$$\begin{aligned} \inf_{\{u_k\}_{k\in\mathbb{N}}} \quad & \frac{2}{\sigma}\log\mathbb{E}\left[\exp\left(\frac{1}{2}\sigma\sum_{k=0}^{\infty}\alpha^k\left(x_k^{\top}Qx_k+u_k^{\top}Ru_k\right)\right)\right] \\ \text{s.t.} \quad & x_{k+1}=Ax_k+Bu_k+D\xi_k, \quad \xi_k \overset{i.i.d.}{\sim} \mathcal{N}(0,\Sigma_{\xi}^{-1}). \end{aligned} \tag{3-4.17}$$

Problem (3-4.17) is the discounted infinite horizon version of the problem pioneered by Jacobson [Jac73]. He was the first to observe that when $\Sigma_{\xi}^{-1}=I_d$ and $\sigma=\delta$ then the minimizing input in (3-4.17) is precisely the feedback in a dynamic game, and thus it equals (3-2.10). With this link in mind it is clear that increasing δ at a *linear* rate, entails increasing the uncertainty set *exponentially* fast, like is seen in Figure 3-12b. In its turn, this explains the *exponential* relation as seen in Figure 3-10, thereby the behaviour as seen in Figure 3-12a. This more *logarithmic* growth is then further observed in Figure 3-13 and 3-14a indeed.

3-4-2-5 Towards Topologically Equivalent Drift Terms

It can be argued that an ideal identification routine establishes in which topological class the unknown system lives, whereafter a robust controller hedges against a certain subset of those models. Remember, we should not fit a damper to a spring. We saw in Lemma 3-3.3 that at least orientation is preserved. Can we do more?

Consider the program (3-4.14) and let the uncertainty fully act in a, but then for $\widehat{a}=0.5$ and $q=0.25$, *i.e.*,

$$\inf_{k\in\mathbb{R}} \sup_{a+k\in\mathcal{A}_{\gamma}(0.5+k)} \mathcal{J}(a+k,0.25+k^2) \tag{3-4.18}$$

with maximizing solution $a^{\star}_{c\ell}(\gamma) := a^{\star}(\gamma)+k^{\star}(\gamma)$ for the worst-case drift: $a^{\star}(\gamma)=\widehat{a}+\Delta_a^{\star}(\gamma)$. Since $\widehat{a}x$ is part of the *structurally stable*[41] set $\{f(x) : f(x)=sx,\ s\in(0,1)\}$ it is

[41] Let f be a linear endomorphism. Then f is ***structurally stable*** if and only if there is a neighbourhood U of f such that $g\overset{t}{\sim}f\ \forall g\in U$. So, a marginally stable system is not structurally stable.

hypothesized that there is a $\gamma > 0$ such that the solutions of (3-4.18)

$$\begin{aligned} \inf_{k\in\mathbb{R}} \sup_{a+k\in\mathcal{A}_\gamma(0.5+k)} \quad & \mathcal{J}(a+k, 0.25+k^2) \\ \text{s.t.} \quad & \widehat{a}x \overset{t}{\sim} ax \end{aligned} \tag{3-4.19}$$

and (3-4.19) are equivalent.

In Figure 3-14a we show that for $\gamma < \sqrt{2}-1$ our hypothesis is true. Note that this is different from simply using the definition of structural stability, we quantify the neighbourhood and optimize over it.

In principle, for these kind of scalar systems we can omit graphical methods and instead derive a sufficient condition on γ to assert topological equivalence between the nominal- and worst-case drift. Let us consider the system corresponding to (3-4.18) and let $p(\delta)$ be the stabilizing solution to the GARE. Then first we want a bound on δ such that $\widehat{a}+\alpha\delta d p(\delta)\big(\lambda(\delta)\big)^{-1}\widehat{a} < 1$. However, since we know that $\big(\lambda(\delta)\big)^{-1}\widehat{a} < 1/\sqrt{\alpha}\ \forall\ \delta \in [0,\bar{\delta}) \subset \mathbb{R}_{\geq 0}$, we get, after plugging in the problem parameters, the following bound on δ:

$$\delta < \frac{1-\widehat{a}}{\sqrt{\alpha}dp(\delta)} \iff \delta < \frac{-\left(\sqrt{2}\left(19-8\sqrt{2}\right)^{1/2}+3\sqrt{2}-2\right)}{(\sqrt{2}-4)} \approx 2.38. \tag{3-4.20}$$

Since the map h is monotone, we can easily translate (3-4.20) to a bound on γ.

Indeed, in the context of periodic orbits or volume-preserving maps ($\widehat{A} \in \mathsf{SL}(n,\mathbb{R})$) we recover the standard lack of structural stability. From the analysis in section 3-3 we see that if $\widehat{A} \in \mathsf{SL}(n,\mathbb{R})$ then $A^\star(\gamma) \in \mathsf{SL}(n,\mathbb{R})$ only for the pathological case of $P(\delta) = 0_{n\times n}$ due to the structure of $T \in \mathcal{S}^n_{++}$, even when $\widehat{A}$ is hyperbolic.

Extending the ideas of this section to a higher dimensional setting is an interesting open problem.

3-4-2-6 Uncertainty in a and b

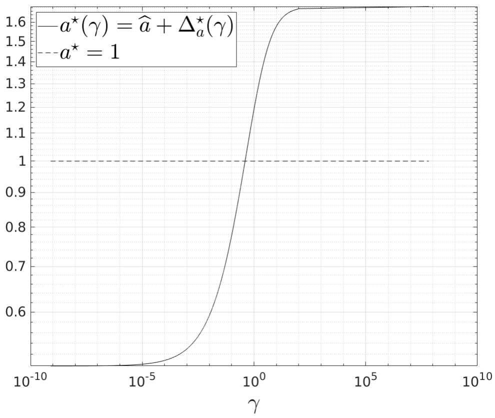
Using the ideas set forth in section 3-2-1 we reconsider program (3-4.14) and construct an uncertainty set for both a and b. To that end we use (3-2.14) with $\widehat{a} = \widehat{b} = 1$ and consider the program

$$\inf_{k\in\mathbb{R}} \sup_{(a,b)\in\mathcal{U}_\gamma((1,1);k)} \mathcal{J}(a+bk, 1+k^2).$$

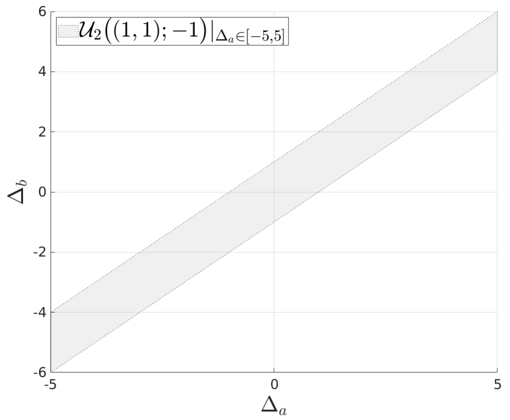
To illustrate how $\mathcal{U}_\gamma$ looks like, pick the upper-bounding radius $\gamma = 2$. Section 3-4-2-3 told us that $\gamma = 2$ corresponds to $\Delta_{a_{c\ell}} \in [-1,1]$ and section 3-4-2-1 showed that $k^\star(\gamma=2) = -1$. Therefore, we have

$$\mathcal{U}_{\gamma=2}((1,1),-1) = \left\{(\Delta_a,\Delta_b) \in \mathbb{R}^2\ :\ \Delta_b = \Delta_a - z,\ z \in [-1,1]\right\}. \tag{3-4.21}$$

A compact subset of the unbounded set (3-4.21) is shown in Figure 3-14b.



(a) The maximizing solution of (3-4.18) as a function of γ, with the same behaviour as Figure 3-13.



(b) A compact subset of (3-4.21), which can thought of as a fat hyperplane.

3-4-3 Worst-Case models

Although section 3-3-1 showed that controllers resulting from (3-1.2) are generally conservative, the framework has another application regardless. Namely, in this section we show what one can do with Proposition 3-2.3. For visualization purposes, we start with a 2-dimensional state-space example with an uncertainty only in the system matrix A.

Consider the controllable pair $(\widehat{A}, B)$ and the structural matrix D defined as

$$\widehat{A} = \begin{pmatrix} 1.2 & 0.5 \\ 0 & 1.2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{3-4.22}$$

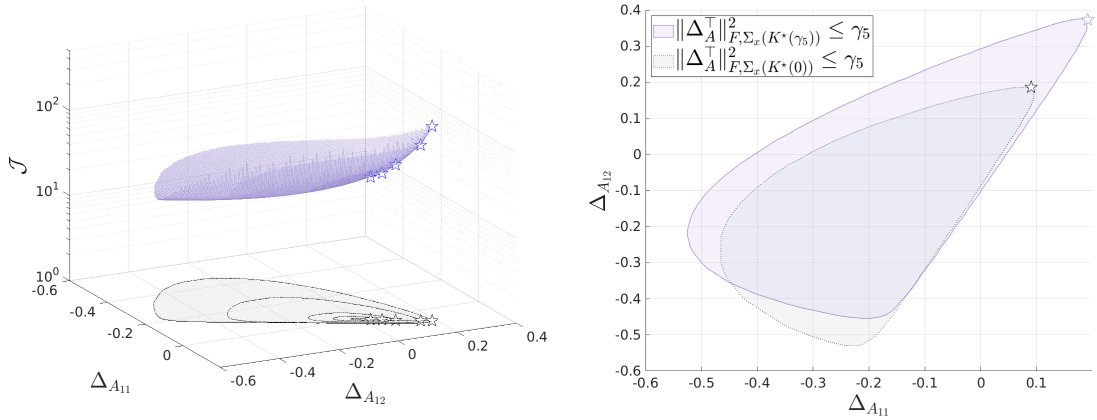
Also define the covariance matrices $\Sigma_v = 0.1I_2$, $\Sigma_0 = I_2$, the cost matrices $Q = 0.1I_2$, $R = 10$, and the discount factor $\alpha = 0.95$.

Then, set K to the nominal discounted LQ regulator[42], *i.e.,* $K = K^\star(0)$. Now, Figure 3-15a depicts the level sets of $\mathbb{A}_\gamma(\widehat{A} + BK^\star(0))$ as defined by Definition 3-1.1 for different levels $\gamma \in \Gamma := \{0.005, 0.03, 0.09, 0.4, 1\}$[43]. We further solve the worst-case model uncertainty problem (3-2.2). The solutions to the worst-case model uncertainty for different γ, denoted by $\Delta_A^\star(\delta)$, are proposed by Proposition 3-2.3. Let us recall that the mapping $\widetilde{h}$ defined in (3-2.5) provides the relation $\gamma = \widetilde{h}(\delta)$ between the different values of γ. In this example, the corresponding δ are $10^{-3} \cdot \{2, 3.9, 5.5, 7.3, 7.7\}$. The locations of these worst-case models are marked by a star symbol in Figure 3-15a.

Looking at the cost in Figure 3-15a anyone would have guessed where these worst-case uncertainties might reside. However, now imagine having a high-dimensional problem (see the next section) for which the critical part of a dynamical system is not that easy to observe, then Proposition 3-2.3 might help.

[42] $K = -\alpha(R+\alpha B^\top PB)^{-1}B^\top P\widehat{A}$ for $P = Q + \alpha\widehat{A}^\top P\widehat{A} - \alpha^2\widehat{A}^\top PB(R+\alpha B^\top PB)^{-1}B^\top P\widehat{A}$ or equivalently $P = Q_{c\ell} + \alpha A_{c\ell}^\top PA_{c\ell}$.

[43] $\gamma = 0$ would yield 0 since the covariance matrices are full-rank, implying $\Sigma_x \succ 0$.



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

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

Figure 3-15: Given the parameters from section 3-4-3 we show the worst-case uncertainties via Proposition 3-2.3 plus how our robust controller anticipates on where the cost increases the sharpest.

At last, it is interesting to see what a robust controller $K^\star(\gamma)$ would do, for say, $\gamma_5 = 1$. See Figure 3-15b for the corresponding sets $\mathbb{A}_{\gamma=1}$ under both types of controller. When compared with Figure 3-15a we clearly see that the robust controller anticipates on where the troubles might occur, *i.e.,* the set is extended in the direction of the worst-case path.

3-4-3-1 Vector-Field Interpretation

In the previous part we fixed $\widehat{A}$, while in practice this matrix might vary based on incoming data. The aim of this section is to 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

$$\widehat{A}(x, y) = \begin{pmatrix} x & y \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{3-4.23}$$

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 solve $\Delta_A^\star(\delta = 10^{-3})$ 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 3-16, 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 3-3.3 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 3-16b we show 1000 Least-Squares estimates[44] for $z^\star := (x, y) = (1.5, 0.5)$, 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 3-7). In Figure 3-19a we show that there are indeed a few trajectories where the robust controllers improves performance compared to $K^\star(0)$, imagine being in the left halfspace of Figure 3-16b, moving towards $(1.5, 0.5)$. Nevertheless, on average the performance deteriorates. As will be shown in section 3-4-4, there are systems for which the Least-Squares estimates *are*, on average, adequate.

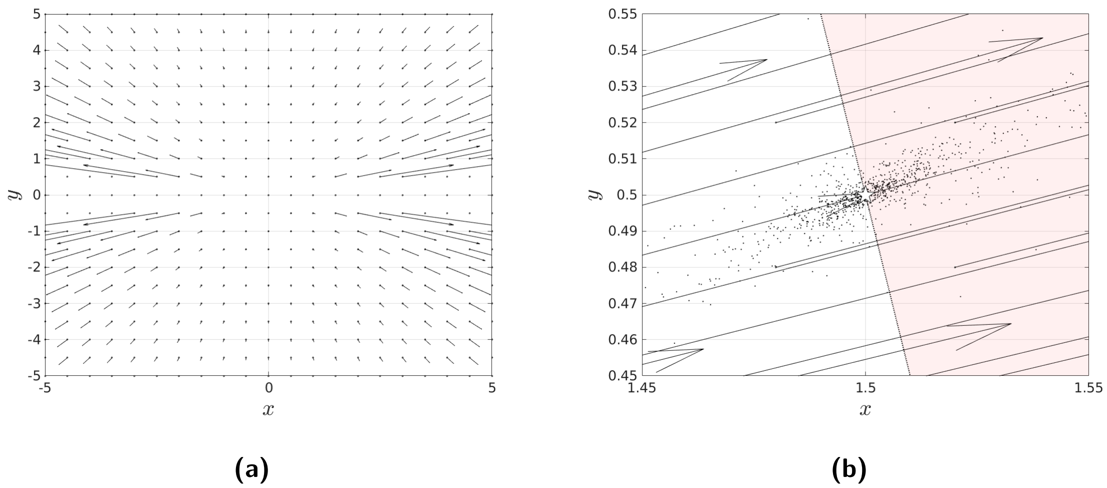


Figure 3-16: (A) Vector-field corresponding to section 3-4-3-1. (B) Zoomed-in version of Figure 3-16a, together with 1000 Least-Squares estimates of $\widehat{A}(1.5, 0.5)$.

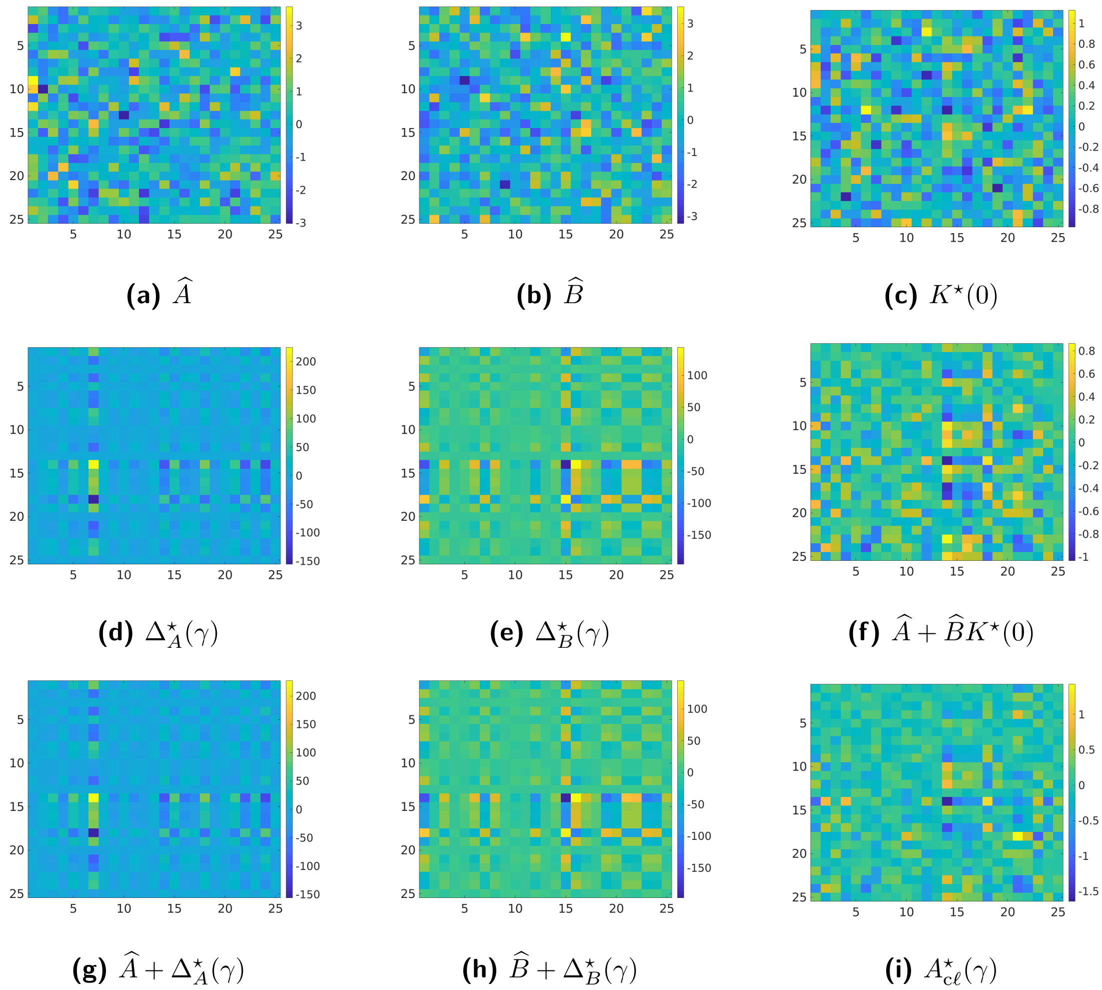
So far we have only considered low-dimensional examples, the reason being analytical insights and visualization. Computationally speaking nothing prohibits us from doing high dimensional examples (on the order 1000-dimensional systems). To still visualize the matrix structure, we continue with an example for $n = 25$.

3-4-3-2 Higher Dimensional Models

For illustration purposes, consider the discrete-time system $x_{k+1} = \widehat{A}x_k + \widehat{B}u_k + v_k$, $x_0 \sim \mathcal{N}(0, I_n)$, $v_k \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_v^2 I_n)$ defined by $n = 25$, $\sigma_v = 0.1$ and the stabilizable pair $(\widehat{A}, \widehat{B})$ randomly drawn from an appropriately sized Gaussian distribution (see Figure 3-17a and 3-17b). Let $Q = R = I_n$, $\alpha = 0.95$ and design a nominal LQ regulator $K^\star(0)$ (see Figure 3-17c). Then for $D = I_n$ we want to investigate $\lim_{\gamma\to\infty} \Delta^\star_{A_{c\ell}}(\gamma)$, and its decomposition, from Corollary 3-2.7. This will tell us where $\widehat{A}$ and $\widehat{B}$ are sensitive with respect to the LQ cost. In Figure 3-17e and 3-17f we show the worst-case uncertainties for $\gamma = 10^4$. Both figures display a clear structure, *e.g.,* the effect of the 7^{th} state on $\widehat{A}$, which would have been difficult to

[44] We used the same procedure as in section 3-4-4, but with $\lambda = 0$, $N = 10$ and B being known.

observe directly from Figures 3-17a and 3-17b. This should be contrasted with the previous example, where we could observe the most sensitive part directly from (3-4.22). Moreover, it is found that indeed a pair of complex closed-loop eigenvalues of approaches the boundary of $\mathbb{D}_{\alpha^{-1/2}}$ (compare Figure 3-17f with Figure 3-17i and see that the worst-case closed-loop matrix clearly increases in norm.).



(a) $\widehat{A}$ **(b)** $\widehat{B}$ **(c)** $K^{\star}(0)$

(d) $\Delta_A^{\star}(\gamma)$ **(e)** $\Delta_B^{\star}(\gamma)$ **(f)** $\widehat{A}+\widehat{B}K^{\star}(0)$

(g) $\widehat{A}+\Delta_A^{\star}(\gamma)$ **(h)** $\widehat{B}+\Delta_B^{\star}(\gamma)$ **(i)** $A_{c\ell}^{\star}(\gamma)$

Figure 3-17: Matrices corresponding to section 3-4-3-2.

3-4-4 Data-Driven Example

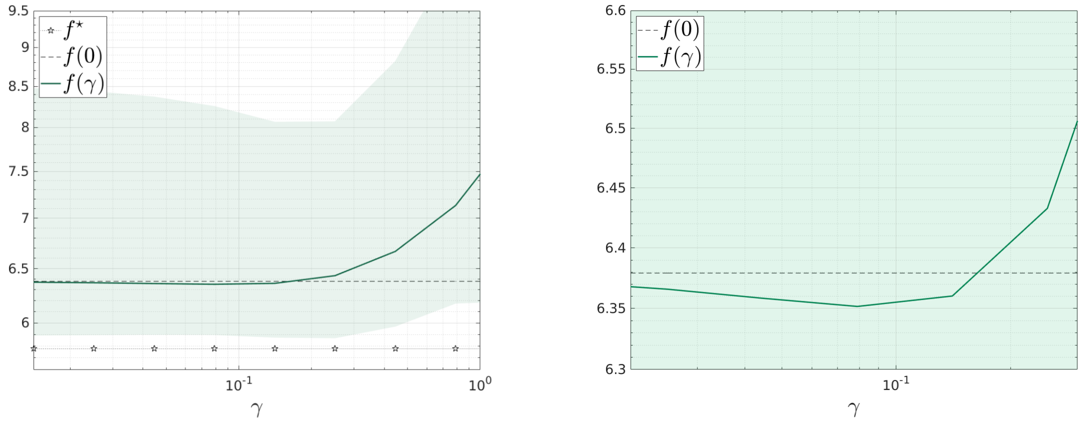
Although section 3-3 indicated that our framework is most likely conservative, intuitively, it is expected that the robustness coming from a game theoretic approach is useful when one is pessimistic about an estimated model. In other words, the real system should be *worse* in some sense, to be precise, 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 (of the form (2-1.1)) as in section 4 of [DMM+18] with $\Sigma_0 = I_3$, $\Sigma_v = 0.1^2 I_3$, $\alpha = 0.95$, $A = \text{tridiag}(0.01, 1.01, 1.01)$, $B = I_3$, $Q = I_3$, $R = I_3$, and give some empirical evidence that our framework can handle these kind of situations. First, we will do $Z = 200$ experiments, for each experiment z, we let the controlled system run for $N = 25$ steps where we have set $u_k^{(z)} \overset{i.i.d.}{\sim} \mathcal{N}(K^\star(0), 0.1^2 I_3)$[45]. The resulting data $\{x_k^{(z)}, u_k^{(z)}\}_{k=0}^N$ is the input to a regularized ($\lambda = 0.001$) Least-Squares problem

$$(\widehat{A}^{(z)}, \widehat{B}^{(z)}) := \underset{A,B}{\text{argmin}} \sum_{k=0}^{N-1} \|x_{k+1}^{(z)} - Ax_k^{(z)} - Bu_k^{(z)}\|_2^2 + \lambda \|A\ B\|_F^2$$

which yields an approximate model of the unknown pair (A, B). Since we have no further structural information, $D = I_3$.

Now 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.*, $(A + BK^{(z)\star}(\widetilde{\gamma})) \in \mathcal{A}_{\widetilde{\gamma}}(\widehat{A}^{(z)} + \widehat{B}^{(z)} K^{(z)\star}(\widetilde{\gamma}))$ and tame the real cost, while surpassing performance of $K^\star(0)$. It is shown in Figure 3-18 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 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 less pronounced.



(a) Induced cost as a function of γ.

(b) Zoomed-in version of Figure 3-18a.

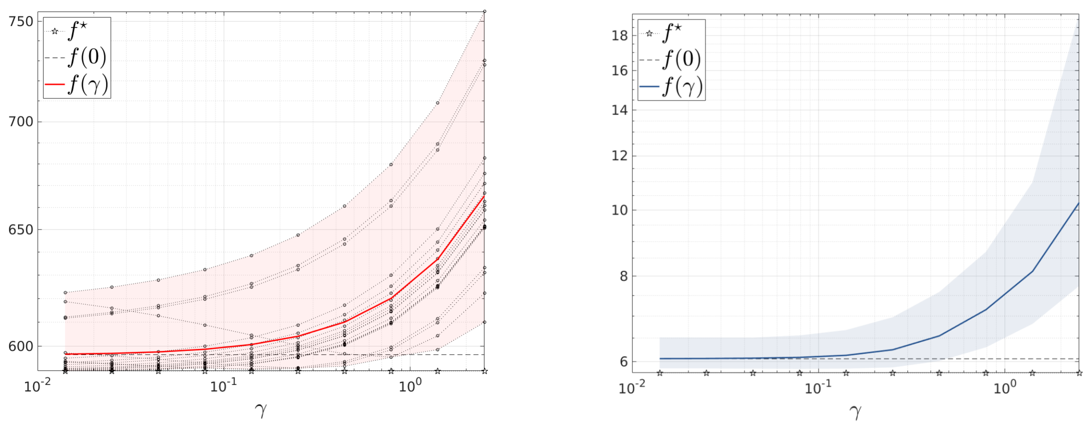
Figure 3-18: For the simple Least-Squares procedure outlined in section 3-4-4, discard the best- and worst 5% of the data. Let $f^\star$ be the best achievable cost, let $f(0)$ be the empirical mean of the induced cost under $K^\star(0)$ (not a function of γ, merely a reference line) and $f(\gamma)$ the empirical mean of the induced cost under $K^\star(\gamma)$. The shaded area is simply the hull of all the remaining 90% of data points.

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,

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

in line with section 3-3. It is an interesting problem to see for which class of systems and identification algorithms our setting can provably outperform a nominal controller. Towards an answer we do a few more simulations and indeed find a well-defined class.

For example, looking back at Figure 3-16b, it is expected that some $\widehat{A}$ on the left side of the hyperplane can give rise to robust controllers doing well on the real system, simply because in that case, the worst-case system is more likely to be close to the real system. In Figure 3-19a we show, for that particular example, 20 cost-trajectories as a function of γ and observe that indeed for a few experiments, this is the case. Even more so, when we reconsider the setup from section 3-4-4, where now B is *known*, then we get a similar result as in Figure 3-19a, the "*sweet-spot*" disappeared (see Figure 3-19b).



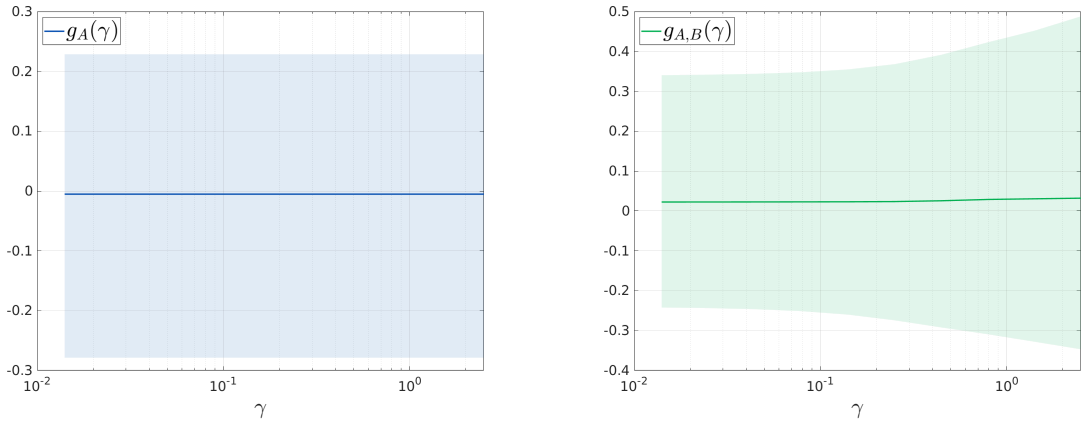
(a) Given the example from section 3-4-3-1, we show 20 cost-trajectories.

(b) The exact same experiment leading up to Figure 3-18, but with only A unknown.

Figure 3-19: When only A is partially unknown, the probability of improving performance under the robust scheme is low. Here, $f^\star$ is the best achievable cost, $f(0)$ is the empirical mean of the induced cost under $K^\star(0)$ (not a function of γ, merely a reference line) and $f(\gamma)$ is the empirical mean of the induced cost under $K^\star(\gamma)$.

Can this behaviour be explained? To that end we recall that in this section we have $D = I_n$ and that Lemma 3-3.3.(iv) 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 3-2-1 we know that we can interpret the worst-case system in many was, *e.g.,* (1) as $A^\star = \widehat{A} + \Delta_A^\star$, but we can also think of $\Delta_A^\star$ as (2) $\Delta_{A_{c\ell}}^\star = \Delta_A^\star + \Delta_B^\star K^\star$. This lead to the comparison made in Figure 3-20, regardless of the interpretation of the uncertainty, $\|A\|_F$ should be bigger than $\|\widehat{A}\|_F$ for it to be in the direction of a worst-case model and potentially improve out-of-sample performance. In line with our hypothesis and Figure 3-18, as Figure 3-20 shows, when B is also known, then the average gap is negative while it becomes positive once B is unknown as well. Note, when B is known, this gap is not a function of γ, but constant; yet, this style is chosen to keep the visualization consistent.

Biased Identification, more Regularization It appears from Figure 3-20 that for our framework to perform well, we need to hope for $\|A\|_F - \|\widehat{A}^{(z)}\|_F \geq 0$. Of course, there is a heuristic to enforce $\|\widehat{A}\|_F \leq \|A\|_F$: sufficiently increasing the ℓ_2-*regularization* parameter $\lambda \in \mathbb{R}_{\geq 0}$ to



(a) The empirical mean is below 0.

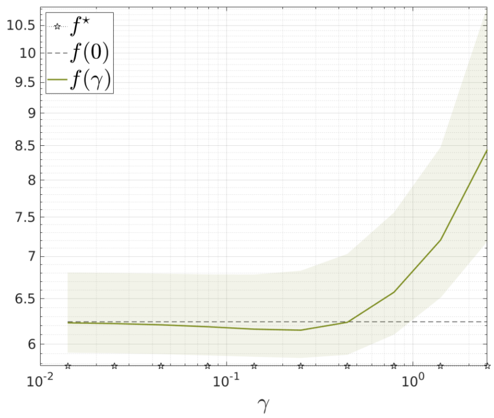
(b) The empirical mean is always above 0.

Figure 3-20: Given the parameters from section 3-4-4, let $g_A(\gamma) := Z^{-1}\sum_{z=1}^{Z}\|A\|_F - \|\widehat{A}^{(z)}\|_F$ and $g_{A,B}(\gamma) := Z^{-1}\sum_{z=1}^{Z}\|A + \left(B - \widehat{B}^{(z)}\right)K^{(z)\star}(\gamma)\|_F - \|\widehat{A}^{(z)}\|_F$.

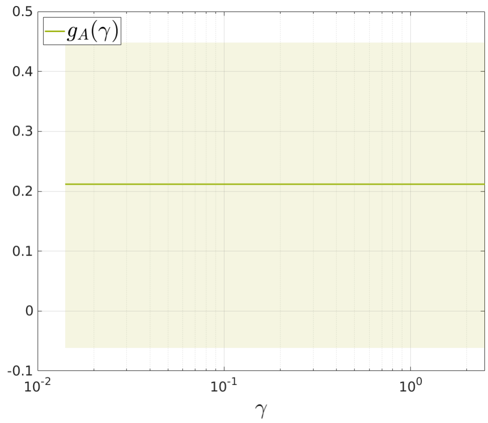
introduce a, for us favourable, bias, *i.e.,* $\|\widehat{A}|_{\lambda>0}\|_F \leq \|\widehat{A}|_{\lambda=0}\|_F$. However how to select λ? Too small is useless and too big is as if we solve a completely different problem.

The result of increasing $\lambda = 10^{-3}$ to $\lambda = 10^{-1}$ is shown in Figure 3-21, 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 to provide some numerical stability such that demanding $\lambda > 0$ is far from unrealistic (see Appendix B-5-1).

Selection of γ via a Holdout Method At last we make a brief digression in a possible heuristic to select γ in practice, using a method referred to as *holdout*, which is the most basic form of cross-validation. To that end, we split, for each z, the data $\{x_k^{(z)}, u_k^{(z)}\}_{k=0}^{N}$ into *training* (N_T) and *validation* (N_V) data via $N = N_T + N_V$, with $N_T = rN$, $N_V = (r-1)N$, $r \in \{0.5, 0.7, 0.9\}$. Just as before, we let B be known, make $\lambda = 0$ again and do 200 experiments for $N \in \{50, 70, 120, 190, 310\}$. Here, we compute $\widehat{A}^{(z)}$ using the N_T dataset and compute a validation system matrix ($\widehat{A}_V^{(z)}$) using the N_V dataset. Next we design a robust controller $K^{(z)\star}(\gamma)$ based on $(\widehat{A}^{(z)}, B)$, for all $\gamma \in 0.5 \cdot \{0, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\} =: \Gamma$ and select the $K^{(z)\star}(\gamma)$ which achieves the lowest cost on the validation system (*i.e.,* under $\widehat{A}_V^{(z)}$), denoted by $K^\star(\widehat{\gamma}^{(z)})$ with $\widehat{\gamma}^{(z)} \in \Gamma$. We call this cost our *certificate* and denote it by $\widehat{\mathcal{J}}^{(z)}$. Moreover, let $\mathcal{J}^{(z)}$ be the cost (on the real system) induced by $K^\star(\widehat{\gamma}^{(z)})$, which we compare with $\mathcal{J}_0^{(z)}$, the cost induced by $K^\star(0)$ (the nominal control law, which also relies on z). In Figure 3-22 we show how well this certificate helps us in selecting γ. Overall we see that, as before, the robust scheme does not help in improving the cost with respect to the nominal scheme, on average (recall Figure 3-7). Also, the more data we have to validate on, the more likely we are to select $\widehat{\gamma}^{(z)} = 0$; for $r = 0.5$ we select $\widehat{\gamma}^{(z)} = 0$ half of the time. If we re-introduce regularization and make $\lambda = 10^{-1}$, then we observe that we can improve upon



(a) For $\lambda = 10^{-1}$ the "*sweet-spot*" reappeared.



(b) The empirical mean is indeed above 0.

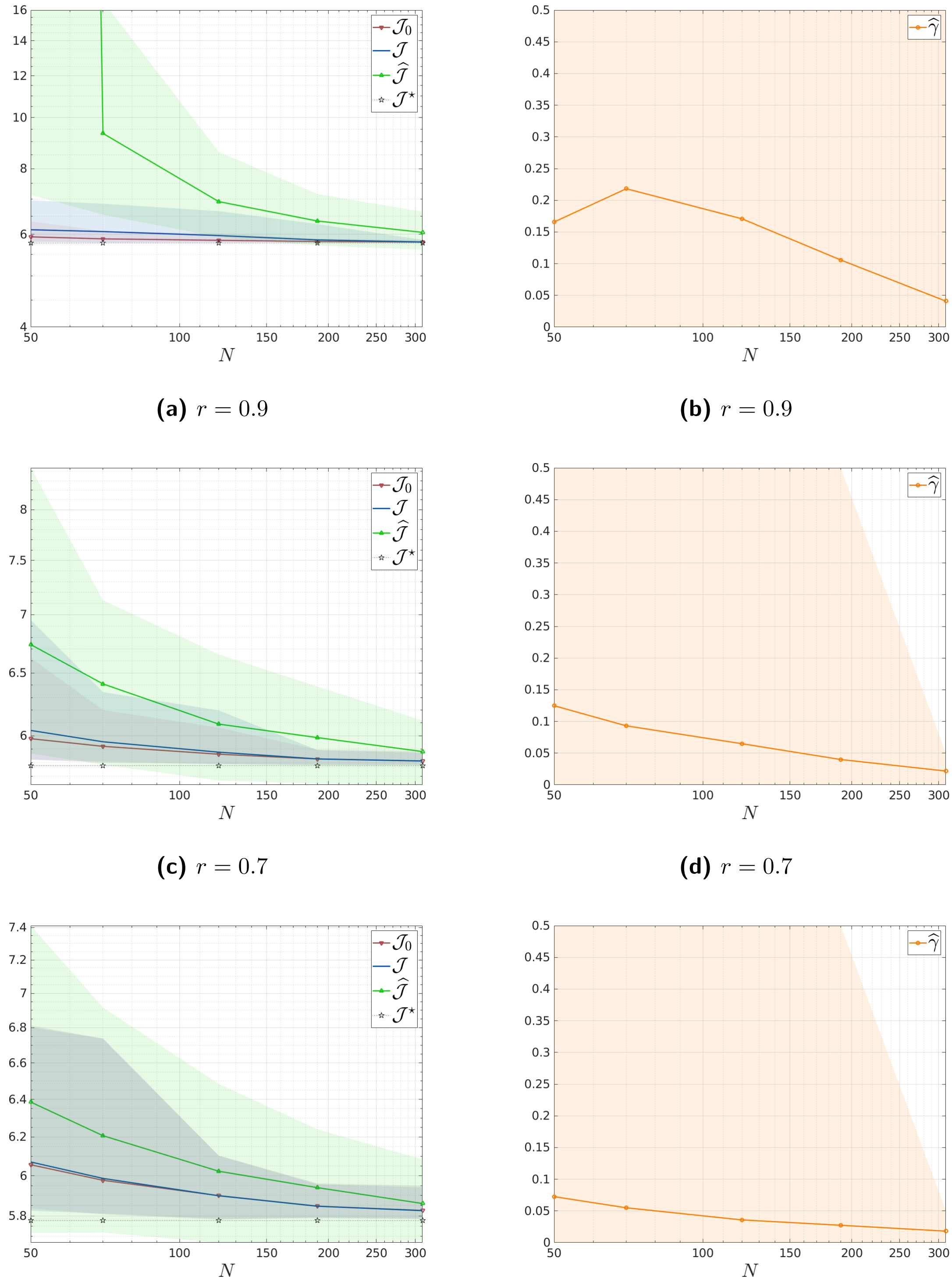
Figure 3-21: Given the parameters from section 3-4-4, let $g_A(\gamma) := Z^{-1}\sum_{z=1}^{Z} \|A\|_F - \|\widehat{A}^{(z)}\|_F$ and perform the exact same experiment leading up to Figure 3-18, but now with only A being unknown *and* $\lambda = 10^{-1}$ (Figure 3-19b used $\lambda = 10^{-3}$). Here, $f^\star$ is the best achievable cost, $f(0)$ is the empirical mean of the induced cost under $K^\star(0)$ (not a function of γ, merely a reference line) and $f(\gamma)$ is the empirical mean of the induced cost under $K^\star(\gamma)$.

the nominal cost by 1.15%; where we however select $\widehat{\gamma}^{(z)} = 0$ for 75% of the time. Moreover, when we set $\alpha = 0.999$, such that $\sqrt{\alpha}A$ is unstable, the results are largely unchanged; the improvement is at most 1.4%.

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

Optimal Selection of γ Finally, to upper-bound possible performance by the previous holdout method, we select γ such that it achieves the smallest cost on the real system, denoted $\gamma^{(z)\star}$, and compare that again to the nominal scenario. Here we will take $N \in \{25, 35, 60, 95, 155\}$ (half the previous set) since that is where we expect potential improvement. The results are shown in Figure 3-23 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 3-7, $\gamma^{(z)\star} = 0$ 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 Figure 3-23c-3-23d. Indeed, here we select $\gamma^{(z)\star} = 0$ less than 7% of the time. Note, we looked at 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 Figure 3-23e-3-23f. 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 also not completely understood why and when an uncertain B seems to help.

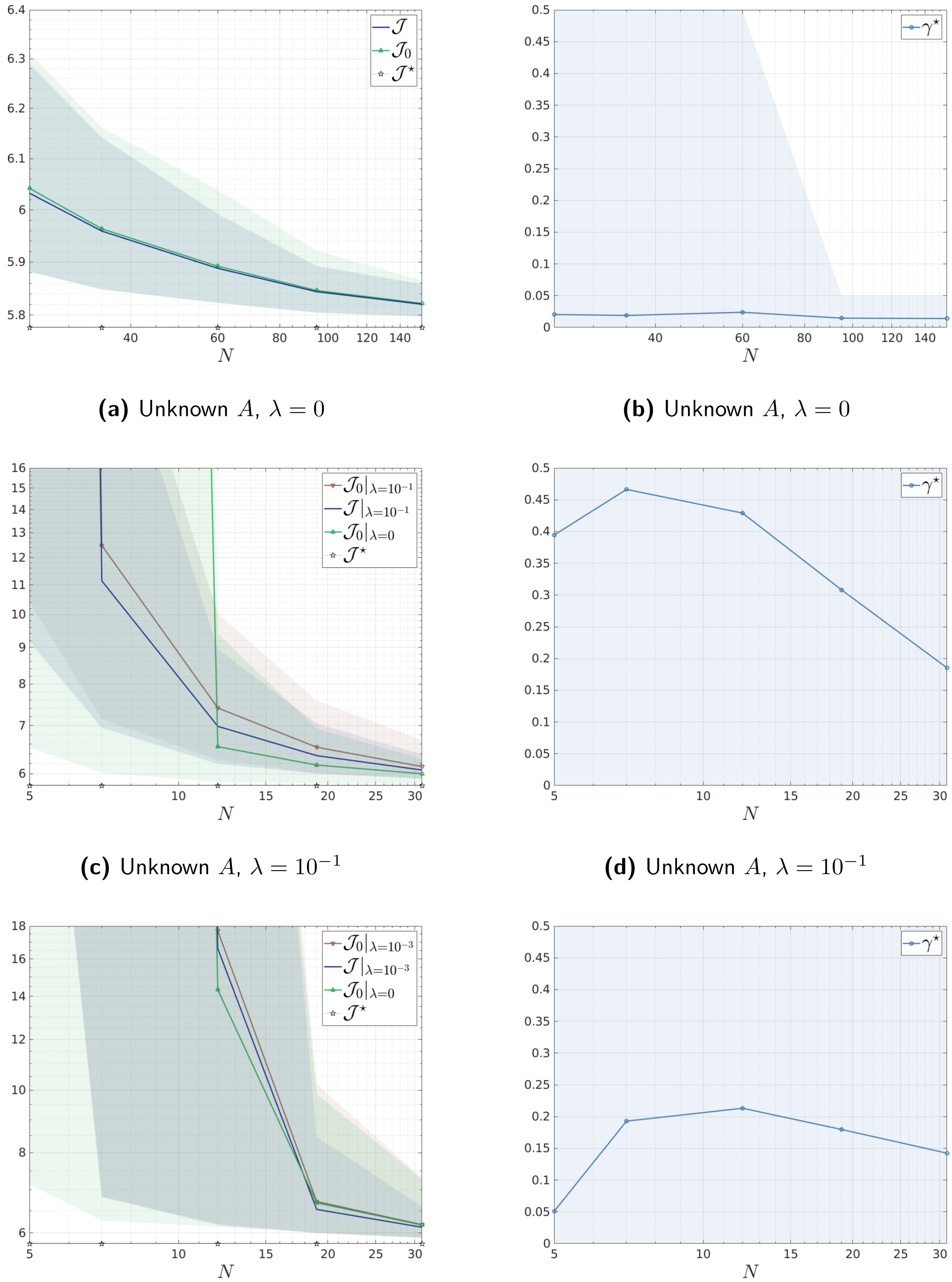


(a) $r = 0.9$ **(b)** $r = 0.9$

(c) $r = 0.7$ **(d)** $r = 0.7$

(e) $r = 0.5$ **(f)** $r = 0.5$

Figure 3-22: Given the holdout method from section 3-4-4, we discard the data corresponding to the top- and bottom 10% of the certificate cost ($\widehat{\mathcal{J}}$) and show the best achievable cost $\mathcal{J}^\star$, the induced cost under $K^\star(0)$: $\mathcal{J}_0$ and the induced cost under $K^\star(\widehat{\gamma})$: $\mathcal{J}$, where $\widehat{\gamma}$ is selected using a holdout method for several r. All thick lines represent the empirical mean over z.



(a) Unknown A, $\lambda = 0$

(b) Unknown A, $\lambda = 0$

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

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

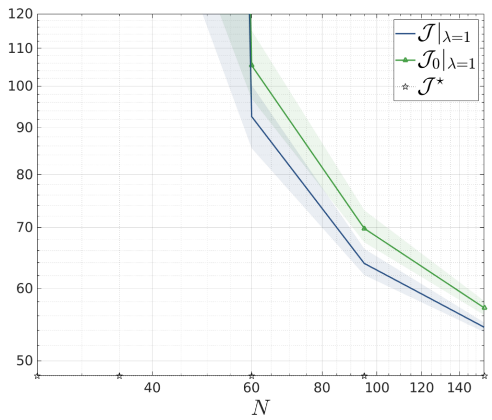
(e) Unknown (A, B), $\lambda = 10^{-3}$

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

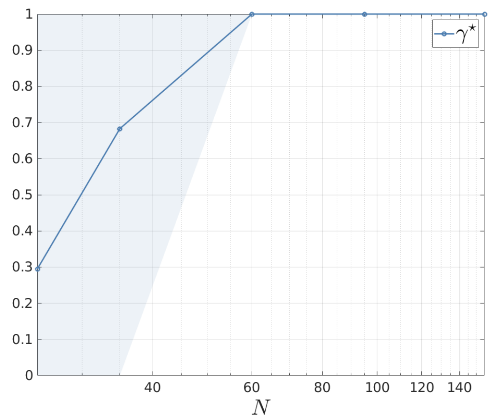
Figure 3-23: Select γ optimally (section 3-4-4), discard the top- and bottom 10% of the cost data for $K^\star(\gamma^\star)$ ($\mathcal{J}$) and show the best achievable cost ($\mathcal{J}^\star$) plus the cost for $K^\star(0)$ ($\mathcal{J}_0$). In 3-23a-3-23b for A unknown, $\lambda = 0$, in 3-23c-3-23d for A unknown, $\lambda = 10^{-1}$; and in 3-23e-3-23f for (A, B) unknown, $\lambda = 10^{-3}$. All thick lines represent the empirical mean over z.

3-4-4-1 High-dimensional Example

If our geometric interpretation is correct, then we should expect an increase in performance for larger n, since for a sufficiently large $\lambda \in \mathbb{R}_{\geq 0}$ the estimates concentrate (at the boundary) of $\mathcal{E}_{\text{in}}$ (see Figure 3-7 or Figure 3-27 below). Here, we generalize the previous example and let $A = (T \otimes I_{n^2}) + (I_{n^2} \otimes T) \in \mathbb{R}^{n^2 \times n^2}$ for $T = \text{tridiag}(\varepsilon, \frac{1}{2}(1+\varepsilon), \varepsilon) \in \mathbb{R}^{n \times n}$, $\varepsilon = 0.01$ and $n = 5$. Hence, instead of 9 unknowns, we now have 625 unknowns. The other parameters are higher dimensional generalizations, *e.g.*, $B = I_{n^2}$. We use the optimal selection scheme from before, this time for a given B, $\lambda = 1$, $N \in \{25, 35, 60, 95, 155\}$, $Z = 200$ and $\gamma \in \Gamma \subseteq [0, 1]$. The results are shown in Figure 3-24. Indeed, the maximal cost improvement has risen to 13.7%, while at the same time maxing out the robustness-radius γ once the cost becomes finite. So indeed, when significant ℓ_2-regularization is used, our framework provides a significantly better performing control law than $K^\star(0)$.



(a) Unknown A, $\lambda = 1$



(b) Unknown A, $\lambda = 1$

Figure 3-24: Given the problem from section 3-4-4-1, select γ optimally (section 3-4-4), discard the top- and bottom 10% of the cost data for $K^\star(\gamma^\star)$ ($\mathcal{J}$) and show 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.

3-4-5 Pendubot, Aiding Mechanical Intuition

Here we will apply some of the robust control techniques to an actual underactuated two-link pendulum (pendubot) built by DCSC, see Figure 3-25. This mechanical system allows for verifying our mechanical intuition as hinted at before, the proposed robust framework improves performance when the estimated model is *optimistic*. Parts of this section are based on a previous report[46] by the author and teammate. First, construct a model using the geometric Euler-Lagrange framework as explained in [BL04]. If we take $x = (\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) \in \mathsf{TQ}$ as our state and τ as input then the model can be written in the convenient nonlinear affine control system form $\dot{x} = f(x) + g(x)\tau$.

[46] For the course SC42035

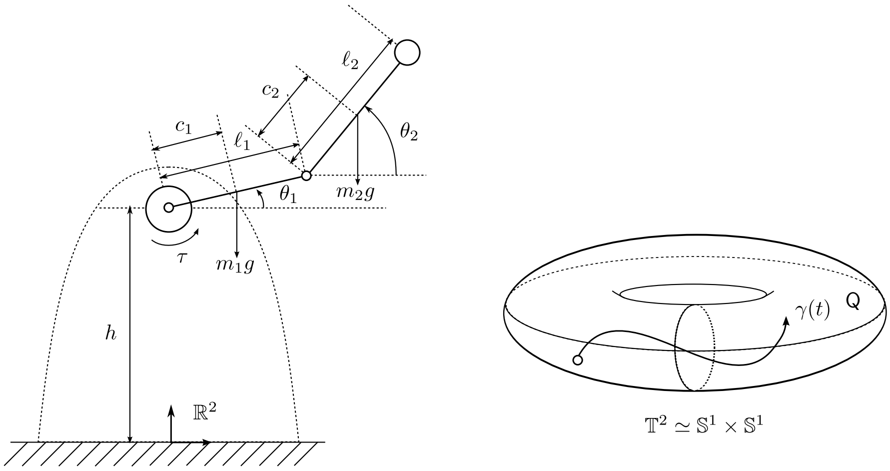


Figure 3-25: *Schematic overview of our setup together with a visual interpretation of the configuration manifold* Q*. Note that only the lower link is actuated!*

3-4-5-1 Equilibria and the Linearized Model

From a physical point of view, our system has four equilibrium points. We focus on the upright position, in the coordinates from Figure 3-25 this is $x^\star_{\text{up,up}} = (\pi/2, \pi/2, 0, 0)$. Then the linearized[47] model around $x^\star_{\text{up,up}} = (\pi/2, \pi/2, 0, 0)$, $\tau = 0$ in state space form $\dot{x} = Ax + B\tau$ becomes

$$\dot{x} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ g\nu\beta/\alpha & -c_2^2 g\ell_1 m_2^2/\alpha & -k_1\nu/\alpha & c_2 k_2 \ell_1 m_2/\alpha \\ -c_2 g\ell_1 m_2\beta/\alpha & c_2 g m_2 \gamma/\alpha & c_2 k_1 \ell_1 m_2/\alpha & -k_2\gamma/\alpha \end{pmatrix} x + \begin{pmatrix} 0 \\ 0 \\ \nu/\alpha \\ -c_2\ell_1 m_2/\alpha \end{pmatrix} \tau. \tag{3-4.24}$$

for $x = (\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2)$ (sometimes $x = (x_1, x_2, x_3, x_4)$), inertias J_1, J_2, $\alpha := ((m_2c_2^2+J_2)(m_1c_1^2+m_2\ell_1^2+J_1) - c_2^2\ell_1^2 m_2^2)$, $\beta := (c_1 m_1 + \ell_1 m_2)$, $\gamma := (m_1 c_1^2 + m_2\ell_1^2 + J_1)$ and $\nu := (m_2 c_2^2 + J_2)$. We will work however with a voltage $u \in \mathcal{U}$, not with the torque $\tau \in \mathcal{T}$. The map from u to τ is approximated by a linear system:

$$\begin{aligned} &\dot{x}(t) = A_\tau x(t) + B_\tau u(t), \quad x = \begin{pmatrix} x_5 \\ x_6 \end{pmatrix}, A_\tau = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \in \mathbb{R}^{2\times 2}, B_\tau = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \in \mathbb{R}^{1\times 2} \\ &\tau(t) = C_\tau x(t), \quad C_\tau = \begin{pmatrix} C_1 & C_2 \end{pmatrix} \in \mathbb{R}^{2\times 1}. \end{aligned}$$

Including this model in (3-4.24) yields for $x^e = (x_1, x_2, x_3, x_4, x_5, x_6)$

$$\dot{x}^e(t) = A^e x^e(t) + B^e u(t) = \begin{pmatrix} A & BC_\tau \\ 0 & A_\tau \end{pmatrix} x^e(t) + \begin{pmatrix} 0 \\ B_\tau \end{pmatrix} u(t). \tag{3-4.25}$$

Furthermore, as we want to control the upright position we can add an integrator to the first link.

[47] See [Lew02, BL04] and references therein for formal arguments why can consider a local linear model.

3-4-5-2 Experimental results

Let the first two states be outputs and use a zero order hold (ZOH) to discretize the system. Furthermore, let $(\alpha, Q, R, \Sigma_v, \Sigma_0)$ be given, exact values are not important for now. The reader should only understand that the emphasis is on controlling θ_2 (the underactuated upper-link) via θ_1. Now, to show when our robust framework might be useful lets give the designer an optimistic model in the following sense. For an inverted pendulum, the further the center of gravity is from the point of rotation, the easier the stabilization, especially for a large mass. So we *over-estimate* the mass and center of gravity from the first link, which all act through the matrix A^e[48]. Then the real system is in some sense, a *worst-case* model, since the actual link is harder to control. So we expect that after applying the robust control paradigm with a sufficiently large γ, we anticipate on this mismatch, and handle the uncertainty. This is exactly what is shown in Figure 3-26, for $D = I_7$[49]. For $\gamma = 0$ we cannot handle the model mismatch, but once we have increased $\gamma \to 1$, we stabilize the inverted pendula[50]. It is also interesting that this approach worked via Kalman-filter based state estimation.

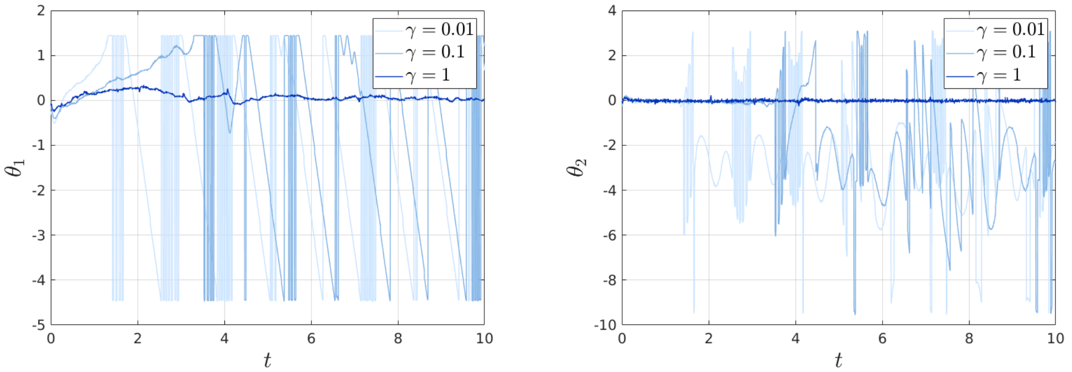


Figure 3-26: For $\gamma = 0$ the controller (always) fails to stabilize the upright position. Then it was found that be increasing γ, taking more "*uncertainty*" into account, we eventually stabilize the pendulum at $\gamma = 1$.

To clarify Figure 3-26 we can look at the first index of $K^\star(\gamma)$:

$$K^\star(\gamma)|_{\gamma=0} = [0.4\ldots],\ K^\star(\gamma)|_{\gamma=10^{-2}} = [0.5\ldots],\ K^\star(\gamma)|_{\gamma=0.1} = [0.7\ldots],\ K^\star(\gamma)|_{\gamma=1} = [1.4\ldots].$$

Based on our understanding of inertia, this is precisely what one would expect. The optimistic model assumes too much *traagheid*[51] Although this experiment is contrived, we did not tune A^e to work for our robust framework, we merely used mechanical intuition, which perfectly illustrates when and how our framework might be of use.

[48] To be specific we have overestimated m_2, c_2, J_2.

[49] We used $D = I_7$, but especially mechanical problems have a clearly structured matrix A, *e.g.* due to several orders of derivatives. However, this is often lost after discretization.

[50] For videos see: LQR: `https://www.youtube.com/watch?v=sglAreUnVvM&list=UU6B3_-BbUJRvEbZTMR9Ytqg&index=2` RLQR: `https://www.youtube.com/watch?v=_CnvX0aLUvQ&list=UU6B3_-BbUJRvEbZTMR9Ytqg`

[51] The dutch word for inertia is much more suitable in this case.

3-4-6 Playing a Game with Wasserstein

At last, one may wonder what can be said about uncertainties in the stochastic part of our model (v_k). As mentioned before, there is a large body of work on the *relative entropy* approach (see for example [RPUS00] and references therein). We provide a slightly different point of view, which at the same time exemplifies the general applicability of our approach, *i.e.,* this example shows that we can consider not only simple "*diagonal*" games like (3-2.16). Namely, based on the work by Yang [Yan18], we can extend those diagonal results to the realm of Wasserstein Distributionally Robust Control.

Consider the dynamical system $\Sigma : \{x_{k+1} = Ax_k + Bu_k + D\xi_k$ where $\xi_k \in \Xi \subseteq \mathbb{R}^{\xi}$ is a random variable ($\xi_k \sim \mathcal{P}_{\xi} \in \mathcal{P}(\Xi)$). In this case the distribution $\mathcal{P}_{\xi}$ is unknown, we only have access to N-samples $\{\widehat{\xi}_1, \dots, \widehat{\xi}_N\} =: \Xi_N$. Given such a dataset Ξ_N, then denote the empirical distribution by $\widehat{\mathbb{P}}_N = \frac{1}{N}\sum_{i=1}^N \delta_{\widehat{\xi}_i}$ and let a closed $\mathcal{W}_2$-ball (*e.g.,* see (3-4.33)) centered at $\widehat{\mathbb{P}}_N$ with radius ρ be denoted by: $B_{\rho}^{\mathcal{W}_2}(\widehat{\mathbb{P}}_N) := \{\mathbb{Q} \in \mathcal{P}(\Xi) \ : \ \mathcal{W}_2(\widehat{\mathbb{P}}_N, \mathbb{Q}) \leq \rho\}$.

Now, consider for some $\delta \in \mathbb{R}_{\geq 0}$ the dynamic game, where for this occasion player 2 pays via a Wasserstein term penalizing the deviation from the empirical distribution:

$$\begin{aligned} \inf_{\{\mu_k\}_{k=0}^{\infty}} \sup_{\{\mathbb{Q}_k \in \mathcal{P}(\Xi)\}_{k=0}^{\infty}} \quad & \mathbb{E}_{x_0,\xi}\left[\sum_{k=0}^{\infty} \alpha^k \left(x_k^\top Q x_k + u_k^\top R u_k - \delta^{-1}\mathcal{W}_2^2(\widehat{\mathbb{P}}_N, \mathbb{Q}_k)\right)\right], \\ \text{s.t.} \quad & x_{k+1} = Ax_k + Bu_k + D\xi_k, \quad u_k = \mu_k(x_k), \\ & \xi_k \sim \mathbb{Q}_k, \quad x_0 \sim \mathcal{P}(0, \Sigma_0). \end{aligned} \tag{3-4.26}$$

This game should be compared with (3-2.16) where we had a simple penalizing term of the form $\delta^{-1}w_k^\top w_k$. This example shows that indeed, we can consider more involved terms, leading to more involved uncertainty sets. Let $\widehat{\Sigma}_{\xi} := \frac{1}{N}\sum_{i=1}^N \widehat{\xi}_i \widehat{\xi}_i^\top$ be the empirical covariance and assume that the empirical mean $\widehat{m}_{\xi} := \frac{1}{N}\sum_{i=1}^N \widehat{\xi}_i$ is 0. Then we briely state Theorem 5 from [Yan18], but in an algebraically simpler form. This follows from applying the matrix inversion lemma[52] twice and indeed yields the familiar equation from deterministic dynamic game theory (*cf.* [BB95] or appendix B-3):

Lemma 3-4.3 (Zero sum game with Wasserstein, Theorem 5 [Yan18])**.** *Let us be given a symmetric minimal positive semi-definite solution P to the algebraic equation*

$$P = Q + \alpha A^\top P \Lambda^{-1} A, \quad \Lambda = (I_n + \alpha(BR^{-1}B^\top - \delta DD^\top)P).$$

Then, when $(\delta^{-1}I - \alpha D^\top PD) \succ 0$ the optimal policy and worst-case distribution in (3-4.26) are given by $u_k^\star = K^\star x_k$ and $\mathbb{Q}_k^\star(x_k) = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_i^\star(x_k)}$, respectively, for:

$$K^\star = -\alpha R^{-1}B^\top P\Lambda^{-1}A, \tag{3-4.27}$$

$$\xi_i^\star(x_k) = \alpha\delta D^\top P\Lambda^{-1}Ax_k + \delta^{-1}(\delta^{-1}I_d - \alpha D^\top PD)^{-1}\widehat{\xi}_i \quad i = 1, \dots, N. \tag{3-4.28}$$

It turns out that spotting the resemblance between program (3-4.26) and a standard dynamic game makes the analysis a lot simpler. Decompose (3-4.28) via the matrices

$$L^\star := \alpha\delta D^\top P\Lambda^{-1}A, \quad \widehat{L}^\star := \delta^{-1}(\delta^{-1}I_d - \alpha D^\top PD)^{-1}, \tag{3-4.29}$$

[52] Specifically, $(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$.

note that $\widehat{L}^{\star}$ is symmetric. Then instead of using $\mathbb{Q}_k^{\star}(x_k)$ from Lemma 3-4.3 we can equivalently write our worst-case dynamical system $\Sigma^{\star}$ as

$$\Sigma^{\star} : \{x_{k+1} = (A + BK^{\star} + DL^{\star})x_k + D\widehat{L}^{\star}\xi_k, \quad \xi_k \sim \widehat{\mathbb{P}}_N. \tag{3-4.30}$$

Of course this description is not unique, we can also write $\Sigma^{\star} : \{x_{k+1} = (A + BK^{\star} + DL^{\star})x_k + D\xi_k, \xi_k \sim \frac{1}{N}\sum_{i=1}^{N}\delta_{\widehat{L}^{\star}\widehat{\xi}_i}$. Note that the difference is essentially the covariance, $\widehat{\Sigma}_{\xi}$ vs. $\widehat{L}^{\star}\widehat{\Sigma}_{\xi}(\widehat{L}^{\star})^{\top}$. These system formulations help in understanding the closed-loop behaviour and describing the robustness. As Yang also hints at, increasing the multiplier δ, *i.e.*, decreasing the penalty on a non-zero Wasserstein distance, allows for more uncertainty (larger $B_{\rho}^{\mathcal{W}_2}(\mathbb{P}_N)$ balls), which is formalized below:

Lemma 3-4.4 (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 of Lemma 3-4.4. Since $\delta_1 \geq \delta_2 > 0$ we have $P(\delta_1) \succeq P(\delta_2)$ such that

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

$\square$

This Lemma nicely shows that a larger "*radius*" indeed allows for a larger amplification of the noise. Then, to see why a Wasserstein penalty can be interpreted via our framework we need one key observation: Since $\widehat{m}_{\xi} = 0$ we obtain:

$$\begin{aligned} h_{\mathcal{W}}(\delta) &:= \mathop{\mathbb{E}}_{x_0,\xi}\left[\sum_{k=0}^{\infty}\alpha^k \mathcal{W}_2^2(\widehat{\mathbb{P}}_N, \mathbb{Q}_k^{\star}(x_k))\right] = \mathop{\mathbb{E}}_{x_0,\xi}\left[\sum_{k=0}^{\infty}\alpha^k \frac{1}{N}\sum_{i=1}^{N}\|\widehat{\xi}_i - \xi_i^{\star}(x_k)\|_2^2\right] \\ &= \mathop{\mathbb{E}}_{x_0,\xi}\left[\sum_{k=0}^{\infty}\alpha^k\left(x_k^{\top}L^{\top}Lx_k + \left\langle (I-\widehat{L})^{\top}(I-\widehat{L}), \widehat{\Sigma}_{\xi}\right\rangle\right)\right] \\ &= \langle L^{\top}L, \Sigma_x\rangle + (1-\alpha)^{-1}\left\langle (I-\widehat{L})^{\top}(I-\widehat{L}), \widehat{\Sigma}_{\xi}\right\rangle. \end{aligned} \tag{3-4.31}$$

However, now it seems like we do not have a nice algebraic expression for Σ_x due to the distribution $\mathbb{Q}_k^{\star}(x_k)$ being state-dependent. Nevertheless, by using (3-4.30) we see that we *do* have an algebraic expression. Let $A_{c\ell} := A + BK + DL$ and $M := D\widehat{L}$, then we end up with a standard Discrete Lyapunov equation:

$$\Sigma_x = \alpha A_{c\ell}\Sigma_x A_{c\ell}^{\top} + \alpha(1-\alpha)^{-1}M\widehat{\Sigma}_{\xi}M^{\top} + \Sigma_0. \tag{3-4.32}$$

The big advantage of this rewriting is that we have a very clear interpretation of what is going on in (3-4.26), or differently put, a new uncertainty set we can handle and study.

Definition 3-4.5 (A Distributional extension of Definition 3-1.1)**.** *Given the tuple $(\widehat{A}, D, \Sigma_0, \alpha)$, some $\gamma \in \mathbb{R}_{\geq 0}$ and an empirical distribution $\widehat{\mathbb{P}}_N$ with empirical mean* 0 *and $\widehat{\Sigma}_{\xi} := \frac{1}{N}\sum_{i=1}^{N}\widehat{\xi}_i\widehat{\xi}_i^{\top}$ then define a set of system- and covariance matrices in $\mathbb{R}^{n\times n} \times \mathcal{S}_+^n$ by:*

$$\mathcal{D}_{\gamma}(\widehat{A}, \widehat{\mathbb{P}}_N) = \left\{ (A_{c\ell}, \Sigma_{\xi}) \quad : \begin{array}{c} A_{c\ell} = \widehat{A} + D\Delta_A, \quad \Sigma_{\xi} = \Delta_{\xi}\widehat{\Sigma}_{\xi}\Delta_{\xi}^{\top}, \\ \Sigma_x = \alpha A_{c\ell}\Sigma_x A_{c\ell}^{\top} + \Sigma_0 + \alpha(1-\alpha)^{-1}D\Sigma_{\xi}D^{\top}, \quad \Sigma_x \succ 0, \\ \left\langle \Delta_A^{\top}\Delta_A, \Sigma_x\right\rangle + (1-\alpha)^{-1}\left\langle (I_d-\Delta_{\xi})^{\top}(I_d-\Delta_{\xi}), \widehat{\Sigma}_{\xi}\right\rangle \leq \gamma \end{array} \right\}.$$

Thus in contrast to definition 3-1.1, we also allow for multiplicative uncertainties in the covariance of ξ. This covariance part has a clear Wasserstein interpretation. Consider two distributions, $\mathbb{P}_1 := \mathcal{N}(0, \widehat{\Sigma}_\xi)$ and $\mathbb{P}_2 := \mathcal{N}(0, \widehat{L}^\star \widehat{\Sigma}_\xi (\widehat{L}^\star)^\top)$ for $\widehat{L}^\star$ from (3-4.29). As derived in [Sho85], the $\mathcal{W}_2$-distance between two Gaussians $\mathbb{Q}_i := \mathcal{N}(\mu_i, Q_i)$ is given by

$$\mathcal{W}_2(\mathbb{Q}_1, \mathbb{Q}_2) = \sqrt{\|\mu_1 - \mu_2\|_2^2 + \mathrm{Tr}(Q_1) + \mathrm{Tr}(Q_2) - 2\mathrm{Tr}\left(Q_2^{1/2} Q_1 Q_2^{1/2}\right)^{1/2}}. \tag{3-4.33}$$

Since $\widehat{L}^\star$ is symmetric, the right-most term under the square-root in (3-4.33) can be easily factored such that we obtain $\mathcal{W}_2^2(\mathbb{P}_1, \mathbb{P}_2) = \mathrm{Tr}((I_d - \widehat{L}^\star)^2 \widehat{\Sigma}_\xi) = \left\langle (I_d - \widehat{L}^\star)^\top (I_d - \widehat{L}^\star), \widehat{\Sigma}_\xi \right\rangle$ indeed. Hence, one could interpret definition 3-4.5 as an extension to definition 3-1.1, where a zero-mean $\mathcal{W}_2^2$-term is added to penalize deviation from some nominal (empirical) covariance matrix. Although this final example showed further potential of our approach it must be remarked that the additional uncertainty has *no* influence on the control law. This observation makes the case for LEQR (3-4.17) again.

3-5 In Conclusion

In this chapter we saw that our Game Theoretic Robust LQ regulators possess considerable structure, which is especially interesting from a dynamical systems point of view (*cf.* Corollary 3-3.6). However, with our statistical motivation from section 1-1 in mind, we have to conclude differently, at least for unbiased estimators.

3-5-1 Game Theoretic Robust LQ Regulators are Almost Surely Conservative

In section 3-3 we saw that ellipsoidal uncertainty sets, *e.g.,* resulting from Least-Squares identification, must be *inscribed* in our set, promoting conservatism. Moreover, regardless from the used identification method or tuning of the cost-matrices, Lemma 3-3.10 told us that $\mathbb{P}\{A^\star(\gamma) = A\} = 0$. On top of that, Lemma 3-3.3.(iv) and Figure 3-7 tell us that unbiased estimators, like standard linear Least-Squares most likely (more than half the time) give rise to $\widehat{A}$ such that $A^\star(\gamma)$ is even *further* away from the real A. Hence, little improvement can be expected, on average. Then, empirical evidence from section 3-4-4 strengthens precisely this conclusion, since $\mathbb{P}\{\gamma^\star = 0\} > \frac{1}{2}$. This conclusion holds for unbiased estimators, when we *do* add for example regularization, nominal performance can be improved, as suggested by Figure 3-23c. The question is, to what extend are *biased* estimators appreciated in practice?

3-5-2 Systems Identified Under ℓ_2-Regularization Benefit from Game Theoretic Controllers

As Appendix B-5-1 explains, 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 $\|\widehat{A}|_{\lambda=0}\ \widehat{B}|_{\lambda=0}\|_F \geq \|\widehat{A}|_{\lambda>0}\ \widehat{B}|_{\lambda>0}\|_F$. Thus, we would like to select some control law

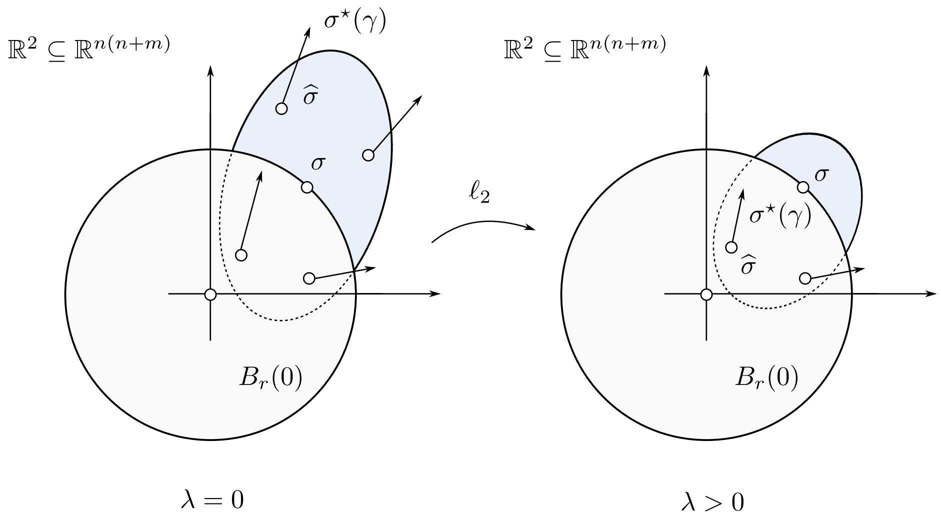


Figure 3-27: Let $\sigma := \mathrm{vec}(A\ B)$ be unknown. Using Least-Squares ($\lambda = 0$) obtain a set (ellipsoid) of estimates ($\widehat{\sigma}$) around this point. From Figure 3-7 and Lemma 3-3.3.(iv) we know that the worst-case models $\sigma^\star(\gamma)$, growing from some estimate $\widehat{\sigma}$, 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 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 $\gamma \in (0, \overline{\gamma})$.

which anticipates on this statistical under-estimation of the Frobenius-norm. Using Lemma 3-3.3.(iv), 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 3-27 (see Figure 5-1 for a remark on the *direction* of the arrows).

And indeed, in Figure 3-23c we observe 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[53], as suggested by the theory and explained in Figure 3-27.

Hence, when the pair (A, B) is identified using ℓ_2-regularized linear Least-Squares, which *is* common practice (see [HKvWV10] 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. Of course, an interesting open problem is to fully formalize this. We provide new motivation to study ℓ_2-regularized identification, which can be compared with the OFU principle, we select some $\widehat{A}$ which does well on the data *and* is small in Frobenius-norm, which happens to relate to our control framework. Of course, regularization should not be simply added to make our control law work. Hence, appropriate classes of problems must be studied, *e.g.*, unstable systems in the small data regime, plus obtaining a relation between the tuple $(\lambda, N, \Sigma, Q, R, \alpha)$ and γ would be interesting.

[53]In [MTR19] 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 3-4-4.

Chapter 4

Policy Iteration, an Involved Proof of Connectedness

Since [FGKM18] was a big inspiration all along, it was natural to ask if that work could be extended to the realms of dynamic game theory. In this chapter we provide tools and theory towards this goal and incidentally show that our uncertainty from chapter 3 is path-connected, which was otherwise a missing result. It must be remarked that during this process we were not the only ones trying to achieve this extension and indeed during the summer 0f 2019 several closely related papers appeared, most notably [ZYB19].

4-1 Introduction

In [FGKM18] it was shown that the LQR cost satisfies a gradient domination property with respect to K, which lead to a variety of model-free convergent gradient algorithms. In this chapter we will show that this property can be generalized to LQ Dynamic Games. However, practically speaking, just like in [FGKM18], this is more a matter of feedback improvement, since the initial conditions must be feasible (stable) already.

Besides that, we investigate when one should use a "*Jacobi*" or "*Gauss-Seidel*"-like update rule. To be specific, when given the task to find the extremizers in

$$\inf_{x\in\mathcal{X}}\sup_{y\in\mathcal{Y}} f(x,y),$$

one could employ some form of a gradient ascent/descent algorithm. Arguably, the most basic form is given by

$$\begin{aligned} x_{k+1} =& x_k - \eta\nabla_x f(x_k,y_k),\\ y_{k+1} =& y_k + \mu\nabla_y f(x_k,y_k), \end{aligned} \tag{4-1.1}$$

for some stepsizes η,μ, which is indeed referred to as the *Jacobi*-like update rule. Another option is to use the updated x already:

$$\begin{aligned} x_{k+1} =& x_k - \eta\nabla_x f(x_k,y_k),\\ y_{k+1} =& y_k + \mu\nabla_y f(x_{k+1},y_k), \end{aligned} \tag{4-1.2}$$

Such an algorithm is referred to as the *Gauss-Seidel*-like algorithm. Both of the names are in correspondence with standard Numerical Linear Algebra terminology (see sec 11.2 [GL13]). Empirical evidence suggests that (4-1.2) performs better (faster convergence) in our case, although our provided convergence rate cannot differentiate between the two types. This is an important open problem.

4-2 Gradient-Based Analysis of LQ Games

To keep it simple, we will consider for some $\gamma \in \mathbb{R}_{\geq 0}$[1], $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $F \in \mathbb{F}^{n\times f}$, $Q \in \mathcal{S}^n_+$, $R \in \mathcal{S}^m_{++}$ and $\Sigma_0 \in \mathcal{S}^n_{++}$ a feasible, deterministic – up to x_0 – infinite horizon game

$$\begin{aligned} \inf_{\{u_k\}_{k\in\mathbb{N}}} \sup_{\{v_k\}_{k\in\mathbb{N}}} \quad & \mathop{\mathbb{E}}_{x_0}\left[\sum_{k=0}^{\infty} x_k^\top Q x_k + u_k^\top R u_k - \gamma^2 v_k^\top v_k\right], \\ \text{subject to} \quad & x_{k+1} = Ax_k + Bu_k + Fv_k, \quad x_0 \sim \mathcal{P}(0,\Sigma_0). \end{aligned} \tag{4-2.1}$$

To emphasize the difference with our stochastic game (3-2.16), we use the notation of (F, v) instead of (D, w). Regardless of how we rewrite the expressions for the optimal strategies we know from Lemma (3-2.9) that they are both linear in x_k such that we can define them to be $u_k^\star := K^\star x_k$, $v_k^\star := L^\star x_k$, for some $K^\star$ and $L^\star$ depending on a (generalized) Riccati equation. Now let the optimal cost be given by

$$\mathcal{J}(K^\star, L^\star) : \begin{cases} & \mathop{\mathbb{E}}_{x_0}\left[\sum_{k=0}^{\infty} x_k^\top \left(Q + (K^\star)^\top R K^\star - \gamma^2 (L^\star)^\top L^\star\right) x_k\right] \\ \text{s.t.} & x_{k+1} = (A + BK^\star + FL^\star)x_k, \quad x_0 \sim \mathcal{P}(0,\Sigma_0). \end{cases}$$

Where we know that $\mathcal{J}(K, L) = \langle P_{KL}, \Sigma_0\rangle$ for

$$P_{KL} = Q + K^\top RK - \gamma^2 L^\top L + (A + BK + FL)^\top P_{KL}(A + BK + FL), \tag{4-2.2}$$

such that for $\Sigma_{KL} := \mathop{\mathbb{E}}_{x_0}\left[\sum_{k=0}^{\infty} x_k x_k^\top\right]$, under (K, L), we can write:

$$\nabla_K \mathcal{J}(K, L) = 2((R + B^\top P_{KL}B)K + B^\top P_{KL}(A + FL))\Sigma_{KL}, \tag{4-2.3}$$
$$\nabla_L \mathcal{J}(K, L) = 2((-\gamma^2 I + F^\top P_{KL}F)L + F^\top P_{KL}(A + BK))\Sigma_{KL}. \tag{4-2.4}$$

This can be derived using the exact same tools as in [FGKM18]. As a sidenote, recall that (4-2.2) is the Lyapunov version (implicit in K and L) of the GARE (Generalized Algebraic Riccati Equation seen for example in $\mathcal{H}_\infty$-control). The explicit version is given by

$$P_{KL} = Q + A^\top P_{KL}\left(I_n + \left(BR^{-1}B^\top - \gamma^{-2}FF^\top\right) P_{KL}\right)^{-1} A \tag{4-2.5}$$

This is usually written more compactly as $P_{KL} = Q + A^\top P_{KL}\Lambda_{KL}^{-1}A$ such that $K^\star = -R^{-1}B^\top P_{KL}\Lambda_{KL}^{-1}A$, $L^\star = \gamma^{-2}F^\top P_{KL}\Lambda_{KL}^{-1}A$. At last, we define some sets which come in very useful. They are essentially sets of feasible initial conditions.

[1] Without loss of generality we take the non-negative real-numbers instead of the entire line.

Definition 4-2.1 (Feasible set). *Consider the game theoretic cost function $\mathcal{J}(K,L)$ in the feedback gains K and L. Then define a set of feasible policies as $\mathcal{KL} := \{(K,L) \,:\, 0 \leq \mathcal{J}(K,L) < \infty\}$. Moreover, define $\mathcal{KL}|_{K'} := \{L \,:\, 0 \leq \mathcal{J}(K',L) < \infty\}$ and similarly $\mathcal{KL}|_{L'} := \{K \,:\, 0 \leq \mathcal{J}(K,L') < \infty\}$.*

It turns out that Definition 4-2.1 plays a key role providing sufficient conditions for the convergence of our algorithms. Specifically, to show that the upper- and lower-bound to the saddle-point cost remain bounded.

4-2-1 First-Order Properties of the Cost

Since we have for the LQ cost no convexity in K and no concavity in L the next best thing is to see if we have gradient domination like in [FGKM18]. As in their work, define:

$$\begin{aligned}
V_{KL}(x) &= x^\top P_{KL} x,\\
Q_{KL}(x,u,v) &= x^\top Q x + u^\top R u - \gamma^2 v^\top v + V_{KL}(Ax + Bu + Fv),\\
A_{KL}(x,u,v) &= Q_{KL}(x,u,v) - V_{KL}(x),
\end{aligned}$$

and two matrices related to the gradients:

$$\begin{aligned}
E^k_{KL} =& (R + B^\top P_{KL} B)K + B^\top P_{KL}(A + FL),\\
E^\ell_{KL} =& -(\gamma^2 I_f - F^\top P_{KL} F)L + F^\top P_{KL}(A + BK),
\end{aligned}$$

such that $\nabla_K \mathcal{J}(K,L) = 2E^k_{KL}\Sigma_{KL}$, $\nabla_L \mathcal{J}(K,L) = 2E^\ell_{KL}\Sigma_{KL}$ for $\Sigma_{KL} := \underset{x_0}{\mathbb{E}}\left[\sum_{k=0}^{\infty} x_k x_k^\top\right]$ and $\Sigma_0 = \underset{x_0}{\mathbb{E}}[x_0 x_0^\top]$. From there it can be observed that when $\Sigma_0 \succ 0$ then both gradients are 0 if and only if $E^k_{KL} = 0$ and $E^\ell_{KL} = 0$, of course, under the assumption that $(K,L) \in \mathcal{KL}$. Moreover, as also pointed in [BMFM19], $\Sigma_0 \succ 0$ helps in formally relating bounded cost to stability, *i.e.,* detectability is not enough[2]. It can be derived that[3] , under the assumption of (K,L) and (K',L') being members of $\mathcal{KL}$,

$$V_{K'L'}(x) - V_{KL}(x) = \sum_{k=0}^{\infty} A_{KL}(x'_k, u'_k, v'_k). \tag{4-2.6}$$

Their result thus implies that for any pair of matrices (K_i, L_j), inducing a finite cost, we have $V_{K'L'}(x) - V_{K_iL_j}(x) = \sum_{k=0}^{\infty} A_{K_iL_j}(x'_k, u'_k, v'_k)$. Moreover, we can rewrite the expression for

[2] Think of some $x'_0 \in \operatorname{Ker}(A)$, then if our oracle declares the cost to be finite this is not per se due to closed-loop stability. Hence we cannot infer a whole lot information regarding stability.

[3] As was done in a note by Tyler Summers and Peyman Mohajerin Esfahani.

A_{KL} quite heavily:

$$\begin{aligned}A_{KL}(x,K'x,L'x) =& x^\top Qx + x^\top K'^\top RK'x - \gamma^2 x'^\top L'^\top L'x \\ &+ x^\top (A+BK'+FL')^\top P_{KL}(A+BK'+FL') - x^\top P_{KL}x \\ =& x^\top((K'-K)^\top R(K'-K) + 2(K'-K)^\top RK)x \\ &+ x^\top(-\gamma^2\left((L'-L)^\top(L'-L)+2(L'-L)^\top L\right))x \\ &+ x^\top\left((B(K'-K)+F(L'-L))^\top P_{KL}(B(K'-K)+F(L'-L))\right)x \\ &+ 2x^\top\left((B(K'-K)+F(L'-L))^\top P_{KL}(A+BK+FL)\right) \\ =& x^\top\Big[(K'-K)^\top(R+B^\top P_{KL}B)(K'-K) \\ &+ (L'-L)^\top(-\gamma^2 I + F^\top P_{KL}F)(L'-L) \\ &+ 2(K'-K)^\top(R+B^\top P_{KL}B) + 2(K'-K)^\top B^\top(A+FL) \\ &+ 2(L'-L)^\top(-\gamma^2 I_f + F^\top P_{KL}F)L + 2(L'-L)^\top F^\top P_{KL}(A+BK) \\ &+ 2(K'-K)^\top B^\top P_{KL}F(L'-L)\Big]x.\end{aligned}$$

Using the expressions for E^k_{KL} and E^ℓ_{KL} this can be simplified to

$$\begin{aligned}&A_{KL}(x,K'x,L'x) = \\ &x^\top\Big[(K'-K)^\top(R+B^\top P_{KL}B)(K'-K) + (L'-L)^\top(-\gamma^2 I_f + F^\top P_{KL}F)(L'-L) \\ &+ 2(K'-K)^\top E^k_{KL} + 2(L'-L)^\top E^\ell_{KL} \\ &+ 2(K'-K)^\top B^\top P_{KL}F(L'-L)\Big]x,\end{aligned} \tag{4-2.7}$$

or in matrix form

$$A_{KL}(x,K'x,L'x) = x^\top \begin{pmatrix} K'-K \\ L'-L \\ I_n \end{pmatrix}^\top \begin{pmatrix} R+B^\top P_{KL}B & B^\top P_{KL}F & E^k_{KL} \\ F^\top P_{KL}B & -\gamma^2 I_f + F^\top P_{KL}F & E^\ell_{KL} \\ (E^k_{KL})^\top & (E^\ell_{KL})^\top & 0 \end{pmatrix} \begin{pmatrix} K'-K \\ L'-L \\ I_n \end{pmatrix} x.$$

So far this matrix form is nice, but will not be used. Now rewrite the expression again to end up with

$$\begin{aligned}&A_{KL}(x,K'x,L'x) = \\ &x^\top\Big[(K'-K)^\top(R+B^\top P_{KL}B)(K'-K) + (L'-L)^\top(-\gamma^2 I_f + F^\top P_{KL}F)(L'-L) \\ &+ (K'-K)^\top(2E^k_{KL} + B^\top P_{KL}F(L'-L)) + (L'-L)^\top(2E^\ell_{KL} + F^\top P_{KL}B(K'-K))\Big]x.\end{aligned} \tag{4-2.8}$$

Then to continue, with these expressions for A_{KL} in mind we can show that ("a form of") gradient domination holds for both pairs $(K,L^\star)$ and $(K^\star,L)$.

Lemma 4-2.2 (A form of Gradient Domination (extension of Lemma 3 [FGKM18])). *Suppose that $(K,L^\star)$ and $(K^\star,L)$ induce a finite cost,* i.e. $K \in \mathcal{KL}|_{L^\star}$, $L \in \mathcal{KL}|_{K^\star}$, *plus, assume that L is chosen such that θ defined by*

$$\theta := \left\|\gamma^2 I_f - F^\top P_{K^\star L}F\right\|_2. \tag{4-2.9}$$

is strictly larger than 0*. Then, we have for any feasible game satisfying the assumptions:*

$$\mathcal{J}(K,L^\star)-\mathcal{J}(K^\star,L^\star)\leq\frac{\|\Sigma_{K^\star L^\star}\|_2}{\mu^2\sigma_{\min}(R)}\|\nabla_K\mathcal{J}(K,L^\star)\|_F^2, \tag{4-2.10}$$

$$\mathcal{J}(K^\star,L^\star)-\mathcal{J}(K^\star,L)\leq\frac{\|\Sigma_{K^\star L^\star}\|_2}{\mu^2\theta}\|\nabla_L\mathcal{J}(K^\star,L)\|_F^2, \tag{4-2.11}$$

where $\mu:=\sigma_{\min}(\Sigma_0)$*.*

Proof. It would be nice if we could complete the square like in [FGKM18]. To do so we have to recall that $(\gamma^2 I_f-F^\top PF)\succ 0$. To keep the expression short, define

$$\begin{aligned}U^k_{KL'L}&:=E^k_{KL}+\frac{1}{2}B^\top P_{KL}F(L'-L),\\ U^\ell_{LK'K}&:=E^\ell_{KL}+\frac{1}{2}F^\top P_{KL}B(K'-K),\end{aligned}$$

then we can write (4-2.8) into

$$\begin{aligned}&A_{KL}(x,K'x,L'x)=\\&x^\top\Big[\Big(K'-K+(R+B^\top P_{KL}B)^{-1}U^k_{KL'L}\Big)^\top(R+B^\top P_{KL}B)\cdots\\&\cdots\Big(K'-K+(R+B^\top P_{KL}B)^{-1}U^k_{LK'K}\Big)-(U^k_{KL'L})^\top(R+B^\top P_{KL}B)^{-1}U^k_{KL'L}\Big]x\\&+x^\top\Big[\Big(L'-L+(-\gamma^2 I_f+F^\top P_{KL}F)^{-1}U^\ell_{LK'K}\Big)^\top(-\gamma^2 I_f+F^\top P_{KL}F)\cdots\\&\cdots\Big(L'-L+(-\gamma^2 I_f+F^\top P_{KL}F)^{-1}U^\ell_{LK'K}\Big)-(U^\ell_{LK'K})^\top(-\gamma^2 I_f+F^\top P_{KL}F)^{-1}U^\ell_{LK'K}\Big]x.\end{aligned}$$

Spot that $(R+B^\top PB)\succ 0$ while $(-\gamma^2 I_f+F^\top PF)\prec 0$, which of course relates to the saddle. To get similar inequalities as equation (12) of the supplementary material from [FGKM18], we need to plug in the optimal (or any other) strategies:

$$\begin{aligned}&A_{K^\star L}(x^\star,K^\star x^\star,L^\star x^\star)=\\&(x^\star)^\top\Big[\Big(L^\star-L+(-\gamma^2 I_f+F^\top P_{K^\star L}F)^{-1}E^\ell_{K^\star L}\Big)^\top(-\gamma^2 I_f+F^\top P_{K^\star L}F)\cdots\\&\cdots\Big(L^\star-L+(-\gamma^2 I_f+F^\top P_{K^\star L}F)^{-1}E^\ell_{K^\star L}\Big)-(E^\ell_{K^\star L})^\top(-\gamma^2 I_f+F^\top P_{K^\star L}F)^{-1}E^\ell_{K^\star L}\Big]x^\star\\&\leq-(x^\star)^\top\Big[(E^\ell_{K^\star L})^\top(-\gamma^2 I_f+F^\top P_{K^\star L}F)^{-1}E^\ell_{K^\star L}\Big]x^\star\end{aligned} \tag{4-2.12}$$

$$\begin{aligned}&A_{KL^\star}(x^\star,K^\star x^\star,L^\star x^\star)=\\&(x^\star)^\top\Big[\Big(K^\star-K+(R+B^\top P_{KL^\star}B)^{-1}E^k_{KL^\star}\Big)^\top(R+B^\top P_{KL^\star}B)\cdots\\&\cdots\Big(K^\star-K+(R+B^\top P_{KL^\star}B)^{-1}E^k_{KL^\star}\Big)-(E^k_{KL^\star})^\top(R+B^\top P_{KL^\star}B)^{-1}E^k_{KL^\star}\Big]x^\star.\\&\geq-(x^\star)^\top\Big[(E^k_{KL^\star})^\top(R+B^\top P_{KL^\star}B)^{-1}E^k_{KL^\star}\Big]x^\star\end{aligned} \tag{4-2.13}$$

Observe that both (4-2.13) and (4-2.12) do not depend on *optimality*, but on *feasibility*. This last condition is precisely the result from [FGKM18], so indeed, we have

$$\mathcal{J}(K,L^\star)-\mathcal{J}(K^\star,L^\star)\leq\frac{\|\Sigma_{K^\star L^\star}\|_2}{\mu^2\sigma_{\min}(R)}\|\nabla_K\mathcal{J}(K,L^\star)\|_F^2.$$

The second part can be derived similarly, we just show the main steps

$$\begin{aligned}\mathcal{J}(K^\star, L^\star) - \mathcal{J}(K^\star, L) &= \mathop{\mathbb{E}}_{x_0}\left[\sum_{k=0}^{\infty} A_{K^\star L}(x_k^\star, u_k^\star, v_k^\star)\right] \\ &\leq \mathop{\mathbb{E}}_{x_0}\left[\sum_{k=0}^{\infty} \operatorname{Tr}\left(x^\star (x^\star)^\top \left(E^\ell_{K^\star L}\right)^\top (\gamma^2 I_f - F^\top P_{K^\star L} F)^{-1} E^\ell_{K^\star L}\right)\right] \\ &\leq \frac{\|\Sigma_{K^\star L^\star}\|_2}{\mu^2 \theta} \|\nabla_L \mathcal{J}(K^\star, L)\|_F^2.\end{aligned} \tag{4-2.14}$$

This concludes showing that if one fixes either $K = K^\star$ or $L = L^\star$, then we can speak of gradient domination for the *other* player.

Note, we had to use θ instead of simply γ^2 due to minus sign in $\gamma^2 I_f - F^\top P_{K^\star L} F$. Indeed, in the next section on algorithms we have to assume that $\theta > 0$, which we do implicitly.

Moreover, we had to add the assumption that $(K^\star, L)$ and $(K, L^\star)$ induce a finite cost, just like was remarked in [ZYB19, p.15] or in the updated version of [FGKM18, p.20-22]. Otherwise the inequalities make little sense (formally speaking, no saddle-point). □

4-3 Gradient-Based Algorithms

In this section we extend the Gauss-Newton- and the Natural-Gradient algorithm from [FGKM18] and provide sufficient conditions for convergence. There do remain many open problems as further explained in section 4-4.

Before we state the algorithms we make a remark regarding existence and uniqueness of a stabilizing saddle-point $(K^\star, L^\star)$. We intend to converge to a stabilizing saddle-point, but is there only one? From [SW94] we know that any stabilizing solution P_{KL} to the GARE (4-2.5) is unique. Moreover, using tools from [BB95] (and references therein) we can easily assert existence (standard minimal realization requirements).

4-3-1 Gauss-Newton Algorithm

It turns out that the "*Gauss-Newton*"-update rule from [FGKM18] can be extended by using the same tools leading up to Lemma 4-2.2[4]. However, formally speaking, we only show the *existence* of these algorithms since the stepsize parameters cannot be fixed under our current assumptions. We summarize these findings below, providing two update rules.

Lemma 4-3.1 (Gauss-Newton, gradient based algorithm). *Let $(K_0, L_0) \in \mathcal{KL}$ for some (potentially partially unknown, yet feasible) 7-tuple $(A, B, F, Q, R, \gamma, \Sigma_0)$, assume that $K^\star \in \mathcal{KL}|_{L_0}$, $L^\star \in \mathcal{KL}|_{K_0}$ and let $\Sigma_0 \succ 0$. Then we state two types of algorithms. First, the* "Gauss-Seidel"*-like update rule*

$$\begin{aligned}K_{k+1} =& K_k - \eta_k (R + B^\top P_{K_k L_k} B)^{-1} E^k_{K_k L_k}, \\ L_{k+1} =& L_k + \mu_k (\gamma^2 I_f - F^\top P_{K_{k+1} L_k} F)^{-1} E^\ell_{K_{k+1} L_k}.\end{aligned} \tag{4-3.1}$$

[4]The fact that one can apply policy iteration to find the optimal strategies is however not new, *cf.* [ATLAK07].

*Secondly, the "*Jacobi*"-like update rule*

$$\begin{aligned} K_{k+1} =& K_k - \eta_k (R + B^\top P_{K_k L_k} B)^{-1} E^k_{K_k L_k}, \\ L_{k+1} =& L_k + \mu_k (\gamma^2 I_f - F^\top P_{K_k L_k} F)^{-1} E^\ell_{K_k L_k}. \end{aligned} \tag{4-3.2}$$

Both of these algorithms converge to $(K^\star, L^\star)$ at a linear rate and along a feasible path ($\mathcal{J}(K_k, L_k) < \infty$) for some sufficiently small *choice of $\eta_k, \mu_k \in (0,1]$ $\forall k$. To be explicit, there is a $\beta \in (0,1)$*[5] *defined by $\beta := \min_k\{\eta_k \|\Sigma_{K_k L^\star}\|_2^{-1}, \mu_k \|\Sigma_{K^\star L_k}\|_2^{-1}\} \cdot \sigma_{\min}(\Sigma_0)$. Furthermore, assume that γ is chosen such that $(\gamma^2 I_f - F^\top P_{K_{k+1} L_k} F) \succ 0$ $\forall k$ in the case of (4-3.1), or $(\gamma^2 I_f - F^\top P_{K_k L_k} F) \succ 0$ $\forall k$ in the case of (4-3.2). Then, if we run the algorithm for N or more steps, where $N \in \mathbb{N}$ satisfies*

$$N \geq \frac{1}{\beta} \log_e \left(\frac{\mathcal{J}(K_0, L^\star) - \mathcal{J}(K^\star, L_0)}{\varepsilon} \right) \tag{4-3.3}$$

for some $\varepsilon \geq 0$, we are ε-close to the optimizers, i.e., *$\mathcal{J}(K_N, L^\star) - \mathcal{J}(K^\star, L_N) \leq \varepsilon$.*

We do the proof in several parts, first showing that the basic update rules contract. Then we show that the overall algorithms contract and eventually we bound the corresponding rate.

Proof. First consider for some $\eta \in (0,1]$ the basic update-rule for K:

$$K' = f(K) = K - \eta (R + B^\top P_{KL} B)^{-1} E^k_{KL}. \tag{4-3.4}$$

One can think of (4-3.4) as the standard update $K_{k+1} = f|_L(K_k)$, we do not use the subscripts right now to avoid clutter in notation. Using this rule and the expression (4-2.13), we obtain the following series of inequalities:

$$\begin{aligned} \mathcal{J}(K', L) - \mathcal{J}(K, L) &= \mathbb{E}_{x_0} \sum_{k=0}^{\infty} A_{KL}(x_k, K' x_k, L x_k) \\ &= -2\eta \mathrm{Tr}(\Sigma_{K'L} (E^k_{KL})^\top (R + B^\top P_{KL} B)^{-1} E^k_{KL}) \\ &\quad + \eta^2 \mathrm{Tr}(\Sigma_{K'L} (E^k_{KL})^\top (R + B^\top P_{KL} B)^{-1} E^k_{KL}) \\ &\leq -\eta \mathrm{Tr}(\Sigma_{K'L} (E^k_{KL})^\top (R + B^\top P_{KL} B)^{-1} E^k_{KL}) \\ &\leq -\eta \sigma_{\min}(\Sigma_0) \mathrm{Tr}((E^k_{KL})^\top (R + B^\top P_{KL} B)^{-1} E^k_{KL}) \\ &\leq -\frac{\eta \sigma_{\min}(\Sigma_0)}{\|\Sigma_{K^\star L}\|_2} (\mathcal{J}(K, L) - \mathcal{J}(K^\star, L)). \end{aligned}$$

The last inequality follows from the observation that (4-2.13) does not depend on optimality, such that we can generalize the gradient domination result. However, we implicitly assumed that $(K', L) \in \mathcal{KL}$ in the first step, of course, when $(K, L) \in \mathcal{KL}$ then for *sufficiently small* $\eta \in (0,1]$ this holds (by continuity[6]). Note, that stepsize need not be constant. Indeed, this comment relates to the one in [FGKM18, p.22]. To proceed, these inequalities imply that

$$\mathcal{J}(K', L) - \mathcal{J}(K^\star, L) \leq \left(1 - \frac{\eta \sigma_{\min}(\Sigma_0)}{\|\Sigma_{K^\star L}\|_2}\right) (\mathcal{J}(K, L) - \mathcal{J}(K^\star, L)), \tag{4-3.5}$$

[5] β can only be 1 when there are effectively no dynamics.

[6] To be more specific, since $\mathcal{KL}|_{K_0}$ is open (in the standard topology), together with the initial condition that both L_0 and $L^\star$ are elements of $\mathcal{KL}|_{K_0}$, implies that there is a neighbourhood U around K_0 such that for all $K' \in U$ we have $\mathcal{J}(K', L_0) < \infty$, $\mathcal{J}(K', L^\star) < \infty$.

which contracts at least for $\eta \leq 1$ – hence any selected $\eta \in (0,1]$ – since $\Sigma_{KL} \succeq \Sigma_0 \succ 0$ for any $(K,L) \in \mathcal{KL}$. We cannot employ the Banach fixed point theorem yet since we lack a global bound (L does not remain fixed). We have to show that there is some $\beta_K \in (0,1]$ such that $\beta_K \leq \eta\sigma_{\min}(\Sigma_0)\|\Sigma_{K^\star L}\|_2^{-1}$ for any L generated by either (4-3.1) or (4-3.2).

Suppose for the moment that such a β_K exists, then we can rewrite (4-3.5) into

$$\mathcal{J}(K',L) - \mathcal{J}(K^\star,L) \leq (1-\beta_K)\left(\mathcal{J}(K,L) - \mathcal{J}(K^\star,L)\right), \tag{4-3.6}$$

Now, since $(1-\beta_K) \in [0,1)$, (4-3.6) shows that employing (4-3.4) iteratively – starting from indicated feasible conditions – results in a $K_\infty := \lim_{N\to\infty} f^N(K)$ such that $\mathcal{J}(K_\infty,L) = \mathcal{J}(K^\star,L)$ for some L. If this $L \in \mathcal{KL}|_{K^\star}$ then we even know that $\mathcal{J}(K_\infty,L) < \infty$.

Similarly, consider for some $\mu \in (0,1]$ the update-rule

$$\begin{aligned} L' = g(L) =& L + \mu(\gamma^2 I_f - F^\top P_{KL}F)^{-1}E^\ell_{KL} \\ & L - \mu(-\gamma^2 I_f + F^\top P_{KL}F)^{-1}E^\ell_{KL}. \end{aligned} \tag{4-3.7}$$

Then we obtain

$$\begin{aligned} \mathcal{J}(K,L') - \mathcal{J}(K,L) &= \mathop{\mathbb{E}}_{x_0} \sum_{k=0}^{\infty} A_{KL}(x_k, Kx_k, L'x_k) \\ &= -2\mu \mathrm{Tr}(\Sigma_{KL'}(E^\ell_{KL})^\top(-\gamma^2 I_f + F^\top P_{KL}F)^{-1}E^\ell_{KL}) \\ &\quad + \mu^2 \mathrm{Tr}(\Sigma_{KL'}(E^\ell_{KL})^\top(-\gamma^2 I_f + F^\top P_{KL}F)^{-1}E^\ell_{KL}) \\ &\geq -\mu \mathrm{Tr}(\Sigma_{KL'}(E^\ell_{KL})^\top(-\gamma^2 I_f + F^\top P_{KL}F)^{-1}E^\ell_{KL}) \\ &\geq -\mu\sigma_{\min}(\Sigma_0)\mathrm{Tr}((E^\ell_{KL})^\top(-\gamma^2 I_f + F^\top P_{KL}F)^{-1}E^\ell_{KL}) \\ &= \mu\sigma_{\min}(\Sigma_0)\mathrm{Tr}((E^\ell_{KL})^\top(\gamma^2 I_f - F^\top P_{KL}F)^{-1}E^\ell_{KL}) \\ &\geq \frac{\mu\sigma_{\min}(\Sigma_0)}{\|\Sigma_{KL^\star}\|}(\mathcal{J}(K,L^\star) - \mathcal{J}(K,L)). \end{aligned} \tag{4-3.8}$$

Thus, again:

$$\mathcal{J}(K,L^\star) - \mathcal{J}(K,L') \leq (1-\beta_L)\left(\mathcal{J}(K,L^\star) - \mathcal{J}(K,L)\right). \tag{4-3.9}$$

if we assume that there is some $\beta_L \in (0,1]$ such that $\beta_L \leq \eta\sigma_{\min}(\Sigma_0)\|\Sigma_{KL^\star}\|_2^{-1}$ for any K generated by either (4-3.1) or (4-3.2).

And again, this contracts for $\mu \leq 1$. Note that this contraction depends on $(\gamma^2 I_f - F^\top P_{KL}F) \succ 0$, therefore the assumption in the Lemma.

Both (4-3.6) and (4-3.9) are however individual contractions.

Next we show that the overall Gauss-Seidel algorithm (4-3.1) converges. Under the assumption that $K^\star \in \mathcal{KL}|_{L_0}$, $L^\star \in \mathcal{KL}|_{K_0}$, iterating the algorithms (4-3.4) and (4-3.7) consecutively, in any order, keeps the cost bounded. To see this, let $(K_0,L_0) \in \mathcal{KL}$. Then $K_N := f^N(K_0) \in \mathcal{KL}|_{L_0}$ since the right-handside of (4-3.6) is bounded. Now, think of (K_N,L_0) as a new initial condition such that $L_M := g^M(L_0) \in \mathcal{KL}|_{K_N}$ since $\mathcal{J}(K_N,L^\star) < \infty$ due to (4-3.6) (plug in $L^\star$) and the initial conditions. Thus, for such an algorithm we have for any integers N,M that $\mathcal{J}(K_N,L_M) < \infty$. This shows the existence of β_K and β_L and thus of some $\beta \in (0,1] := \min\{\beta_K, \beta_L\}$. At this point we could show that first iterating over K (infinitely long) and then over L would yield the solution (similar to [ZYB19]). However, by using (4-3.9) we see that $\mathcal{J}(K^\star, L_M) < \infty$ so that you can in fact alternate between updating K and L. This does however not show that we converge to something meaningful.

Now, due to previous observations we can write:

$$\begin{aligned} \mathcal{J}(K^\star, L^\star) - \mathcal{J}(K^\star, L_M) &\leq (1-\beta)\big(\mathcal{J}(K^\star, L^\star) - \mathcal{J}(K^\star, L_0)\big), \\ \mathcal{J}(K_N, L^\star) - \mathcal{J}(K^\star, L^\star) &\leq (1-\beta)\big(\mathcal{J}(K_0, L^\star) - \mathcal{J}(K^\star, L^\star)\big), \end{aligned} \tag{4-3.10}$$

which is indeed well-defined under the assumption that $K^\star \in \mathcal{KL}|_{L_0}$, $L^\star \in \mathcal{KL}|_{K_0}$. Add these terms and obtain

$$\mathcal{J}(K_N, L^\star) - \mathcal{J}(K^\star, L_M) \leq (1-\beta)\big(\mathcal{J}(K_0, L^\star) - \mathcal{J}(K^\star, L_0)\big), \quad N, M \in \mathbb{Z}_{>0}. \tag{4-3.11}$$

This implies that for any consecutive application of (4-3.4) and (4-3.7) we indeed contract. Since we work with a saddle-point[7] $\mathcal{J}(K^\star, L^\star)$:

$$\mathcal{J}(K^\star, L) \leq \mathcal{J}(K^\star, L^\star) \leq \mathcal{J}(K, L^\star),$$

we have our desired result, $K' \to K^\star$, $L' \to L^\star$ at a linear rate while maintaining a bounded cost throughout. In other words, the upper- and lower-bound converge to each other at a *linear* rate.

Now we can be brief on the proof for the Jacobi-like algorithm (4-3.2). The individual update-rules (4-3.4) and (4-3.7) induce contractions, but we cannot conclude that they induce an overall contraction as well. We do not even know if $\mathcal{J}(K_1, L_1)$ is bounded. To prove boundedness recall again that for the Jacobi-algorithm we have

$$\mathcal{J}(K_0, L^\star) - \mathcal{J}(K_0, L_1) \leq \left(1 - \frac{\mu \sigma_{\min}(\Sigma_0)}{\|\Sigma_{K_0 L^\star}\|_2}\right)\big(\mathcal{J}(K_0, L^\star) - \mathcal{J}(K_0, L_0)\big), \tag{4-3.12}$$

$$\mathcal{J}(K_1, L_0) - \mathcal{J}(K^\star, L_0) \leq \left(1 - \frac{\eta \sigma_{\min}(\Sigma_0)}{\|\Sigma_{K^\star L_0}\|_2}\right)\big(\mathcal{J}(K_0, L_0) - \mathcal{J}(K^\star, L_0)\big). \tag{4-3.13}$$

Since (4-3.12) must hold for any feasible K_0, we can plug in $K^\star$. Based on the initial conditions this implies that $\mathcal{J}(K^\star, L_1) < \infty$. Similarly, $\mathcal{J}(K_1, L^\star) < \infty$. Then plug K_1 into (4-3.12), which now implies the desired result: $\mathcal{J}(K_1, L_1) < \infty$ since $\mathcal{J}(K_1, L_0) < \infty$ follows from (4-3.13). Then the results follows from a standard induction argument and applying the same β idea to the Jacobi-case.

At last, regarding the bound on N via (4-3.3), assume we are given some scalar dynamical system $x_N = (1-\beta)^N x_0$, $x_0 > 0$, $\beta \in (0,1)$ with a desired bound $x_N \leq \varepsilon$. Then set $(1-\beta)^N x_0 \leq \varepsilon$ and obtain a first bound on N: $N \geq \log_{1-\beta}(\varepsilon/x_0) = -\log_e(x_0/\varepsilon)[\log_e(1-\beta)]^{-1}$. Using the identity[8]: $\beta^{-1} \geq -[\log_e(1-\beta)]^{-1}\ \forall\, \beta \in (0,1)$, this previous bound can be simplified to:

$$N \geq \frac{1}{\beta}\log_e\left(\frac{x_0}{\varepsilon}\right) \implies x_N \leq \varepsilon. \tag{4-3.14}$$

Then the result follows after using (4-3.14) in the context of (4-3.11). □

[7] Recall, a saddle, but not over any arbitary Euclidean space for the pair (K, L), see [Mag76]

[8] Observed it, probably known but could not find a source. To (partially) see it, consider the equivalent formulation $1 \geq -\beta\big(\log_e(1-\beta)\big)^{-1} =: g(\beta)$ and see that indeed $\lim_{\beta \uparrow 1} g(\beta) = 0$ while $\lim_{\beta \downarrow 0} g(\beta) = 1$.

What is interesting about this proof is that we converge along a feasible path. Recall that in standard linear optimal control theory, *e.g.,* using Riccati equations, iterative algorithms usually only yield formally stable controllers asymptotically. Moreover, it should be clear that the stepsize is the critical parameter regarding convergence, which we can unfortunately not properly fix at this point.

Corollary 4-3.2. *Let $\mathcal{J}$ correspond to the cost of a feasible game. Then feasible set $\mathcal{F}$ of initial conditions defined by*

$$\mathcal{F} := \{(K_0, L_0) \ : \ (K_0, L_0) \in \mathcal{KL}, \ K^\star \in \mathcal{KL}|_{L_0}, \ L^\star \in \mathcal{KL}|_{K_0}\} \tag{4-3.15}$$

is path-connected.

Proof. Follows from Lemma 4-3.1. □

This means that if $(K_{\mathrm{LQR}}, 0) \in \mathcal{F}$, then there is a feasible path from the worst-case game theoretic uncertainty to the unperturbed model under its optimal control law. Moreover, we can finally show the the finite union of connected sets from Proposition 3-2.1 are actually one set. If this would not have been the case, then inscribed sets are possibly even more conservative.

Corollary 4-3.3. *The set $\mathcal{A}_\gamma$ is path-connected.*

Proof. Consider the initial condition $(K^\star, L_0) \in \mathcal{KL}$, then this L_0 is necessarily a perturbation in $\mathbb{A}_\gamma$. Employing the fixed point algorithm $L_{k+1} = g|_{K^\star}(L_k)$ from Lemma 4-3.1 yields a path from L_0 to $L^\star$, all within $\mathbb{A}_\gamma$. Since this holds for any feasible L_0, the set $\mathbb{A}_\gamma$, and thereby $\mathcal{A}_\gamma$, must be path-connected. □

We can already make one remark regaring the comparison between (4-1.1) and (4-1.2). As is generally the case (see [GL13]), providing a sharp convergence *rate* is more difficult for the Gauss-Seidel algorithm. In some sense is the provided convergence rate a lower-bound in that all the variations of these Gauss-Newton algorithms converge *at least* at this rate (See the recent work [BMFM19] for sharper LQR convergence rates).

As may become clear from the proof, Lemma 4-3.1 can be greatly generalized, *i.e.,* first update K k_i-times and then L ℓ_j-times, for what it is worth.

Also, to see why we speak of a *gradient* based algorithm, recall that we can write (4-3.1) as

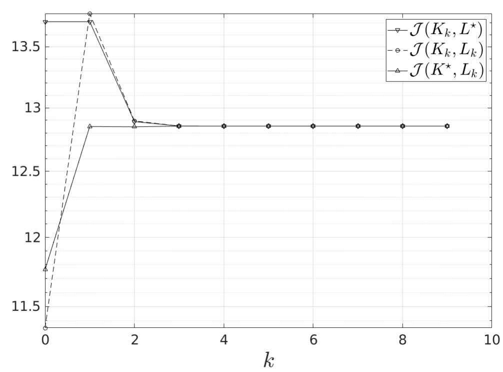
$$\begin{aligned} K_{k+1} =& K_k - \frac{1}{2}\eta_k (R + B^\top P_{K_k L_k} B)^{-1} \nabla_K \mathcal{J}(K_k, L_k) \Sigma^{-1}_{K_k L_k} \\ L_{k+1} =& L_k + \frac{1}{2}\mu_k (\gamma^2 I_f - F^\top P_{K_{k+1} L_k} F)^{-1} \nabla_L \mathcal{J}(K_{k+1}, L_k) \Sigma^{-1}_{K_{k+1} L_k}. \end{aligned} \tag{4-3.16}$$

Although we need a very knowledgeable oracle for algorithms like (4-3.16) to work, we can do a simple simulation. Here we will also compare several alternating schemes, showing the differences between for example (4-3.2) and (4-3.1). More schemes are possible, but these are the most realistic from a practical point of view.

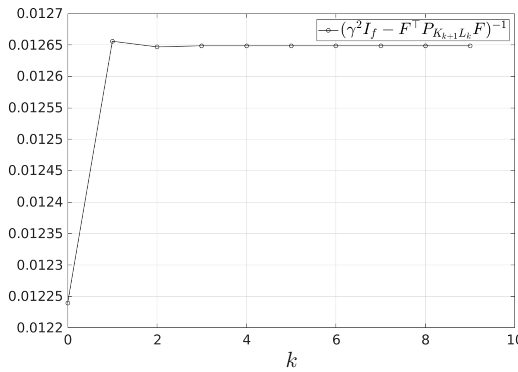
Example 4-3.4 (Gauss-Newton method, $\eta = 1$, $\mu = 1$)**.** *Let the model be parametrized by the controllable pair* (A, B) *and the adversarial input matrix* F *defined as*

$$A = \begin{pmatrix} 1.2 & 0.5 \\ 0 & 0.2 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Furthermore let $Q = I_2$*,* $R = 1$ *and* $\Sigma_0 = I_2$*. Throughout we set* $\gamma := 10$ *and fix both stepsizes to* 1*. Now we simulate* (4-3.16) *for some initial* (K_0, L_0)*, which needs to correspond to well-defined cost* $\mathcal{J}(K_0, L_0)$*. The most simple choice is* $K_0 = K_{\text{LQR}}$ *and* $L_0 = 0$*, for which thus only* $L^\star \in \mathcal{KL}|_{K_0}$ *is questionable. As shown in Figure 4-1 the algorithm converges as predicted, with indeed a bounded cost throughout. Note however that* $\mathcal{J}(K_k, L_k)$ *is by no means sandwiched between* $\mathcal{J}(K_k, L^\star)$ *and* $\mathcal{J}(K^\star, L_k)$*, which exemplifies perfectly the difficulty in bounding the induced cost.*



(a) The upper- and lower-bound converge to each other.



(b) γ is sufficiently large such that the analysis holds throughout

Figure 4-1: For $\gamma = 10$ the algorithm (4-3.1) converges.

If instead of the "Gauss-Seidel"*-like algorithm* (4-3.1)*, we do the exact same simulation, but then under the* "Jacobi"*-like algorithm* (4-3.2)*, the behaviour changes, see Figure 4-2. We observe less contractive and more oscillatory behaviour.*

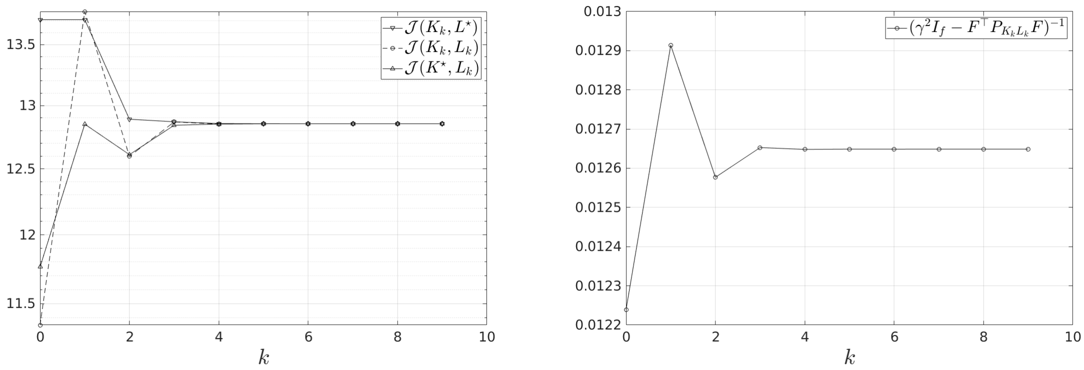
Now if we let γ *to be* 6.5*, which still corresponds to a feasible infinite-horizon game, then we do not converge under* (4-3.1) *as shown in Figure 4-3a. There is an easy explanation,* γ *is not sufficiently large, the factor* $(\gamma^2 I_f - F^\top P_{K_{k+1}L_k} F)$ *fails to be positive-definite throughout (compare Figure 4-1b and Figure 4-3b).*

Again, we do the same simulation, but now for the algorithm as given by (4-3.2)*, see Figure 4-4. It can be observed that* $(\gamma^2 I_f - F^\top P_{K_k L_k} F) \succ 0$ *holds, but still the algorithm fails to converge. The explanation is that in this case we fail to satisfy the assumption* $L^\star \in \mathcal{KL}|_{K_0}$*. In other words,* K_0 *cannot stabilize* $L^\star$*, which is used in the proof of Lemma 4-3.1.*

While performing simulations like in Example 4-3.4, it can be observed that the rule

$$L_{k+1} = L_k + (\gamma^2 I_f + F^\top P_{K_{k+1}L_k} F)^{-1} E^\ell_{K_{k+1}L_k},$$

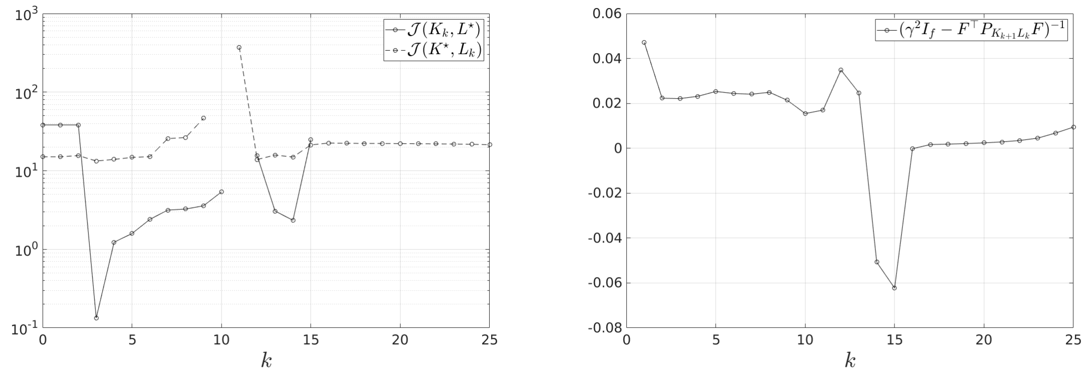
works way too well. This can be explained. Implicitly we need to assume that γ^2 is sufficiently large such that $(\gamma I_f - F^\top P_{K_{k+1}L_k} F) \succ 0 \ \forall k \in \mathcal{I}_k \subset \mathbb{N}$. Here, $\mathcal{I}_k$ denotes the set of



(a) The upper- and lower-bound still converge to each other.

(b) γ is sufficiently large such that the analysis remain correct throughout

Figure 4-2: For $\gamma = 10$ the algorithm (4-3.2) still converges, but slower than (4-3.1)



(a) The upper- and lower-bound fail to converge to each other.

(b) The scaling factor fails to remain positive throughout.

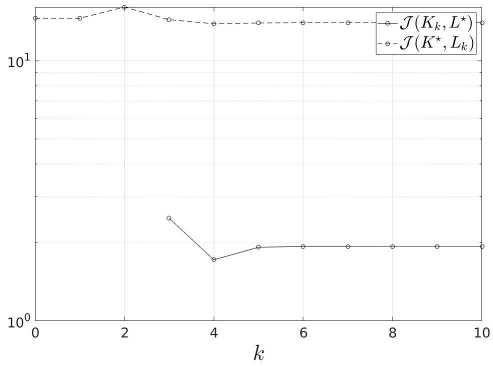
Figure 4-3: For $\gamma = 6.5$ the algorithm (4-3.1) fails to converge.

steps to converge to desired precision. If we do not assume that this inequality holds then the inequalities in (4-3.8) break down. Of course, flipping the sign makes the expression more well-conditioned and as it turns out, can already make the algorithm converge, but not provably.

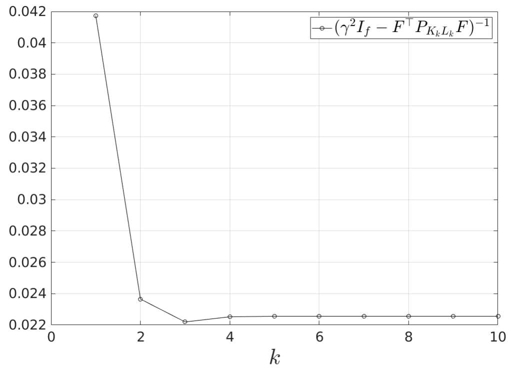
Also, we can construct examples where $\eta = 1$, $\mu = 1$ do not suffice and for example fail to make $\mathcal{J}(K_1, L^\star) < \infty$. In the next example we show how non-trivial the solution might be.

Example 4-3.5 (Ill-conditioned Gauss-Newton)**.** *We use the Gauss-Seidel algorithm* (4-3.1) *on the same parameters as in Example 4-3.4, but now we use a different, still feasible, initial condition, namely*

$$K_0 = \begin{pmatrix} -0.5059 & 0.3294 \end{pmatrix}, \quad L_0 = \begin{pmatrix} -0.0198 & 0.0077 \end{pmatrix}.$$



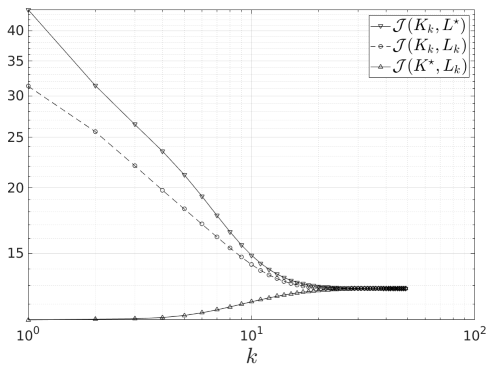
(a) The upper- and lower-bound fail to converge to each other.



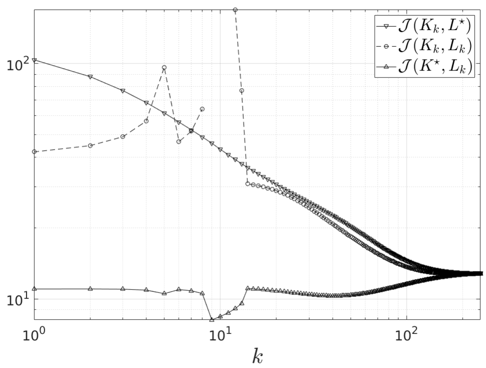
(b) The scaling factor does remain positive throughout.

Figure 4-4: For $\gamma = 6.5$ the algorithm (4-3.2) fails to converge, which could be contributed to $\mathcal{J}(K_0, L^\star) = \infty$.

Then we observe precisely that $\mathcal{J}(K_1, L^\star) = \infty$ under the stepsizes $\eta = \mu = 1$. However if we decrease them to (1) $\eta = \mu = 10^{-1}$ or (2) $\eta = \mu = 10^{-2}$ some typical non-convex behaviour is displayed in Figure 4-5. Stepsize option (1) works, but we cannot simply decrease it since (2) fails. Intuition might be that option (1) jumps over the unstable points, while (2) does not. This further clarifies why a fixed stepsize might not be ideal.



(a) The upper- and lower-bound converge to each other for $\eta = \mu = 10^{-1}$.



(b) The upper- and lower-bound fail to converge to each other for $\eta = \mu = 10^{-2}$.

Figure 4-5: Convergence of the algorithm is not preserved under shrinking stepsizes.

For the sake of experiment we do some simulations while performing the most basic form of line-search to find η_k and μ_k. Note, here we use a rather clever oracle since we assert at each step that the current cost *and* the bounds are finite. The point is to show that $\mu = \eta = 1$ usually works, but not always! Also, by allowing for a varying stepsize we can convergence much more quickly.

4-3-2 Natural Gradient Algorithm

The Gauss-Newton method has a few drawbacks. Most notably, we need a rather clever oracle to update the strategies. Here we investigate a slightly simpler algorithm. As before, it turns out that the results from [FGKM18] extend *naturally*.

Consider for some non-negative η and μ the update rules

$$\begin{aligned} K' &= f|_L(K) = K - \eta\frac{1}{2}\nabla_K \mathcal{J}(K,L)\Sigma_{KL}^{-1} = K - \eta E_{KL}^k, \\ L' &= g|_K(L) = L + \mu\frac{1}{2}\nabla_L \mathcal{J}(K,L)\Sigma_{KL}^{-1} = L + \mu E_{KL}^\ell. \end{aligned}$$

Then first for K, under $\eta \leq \|R + B^\top P_{KL}B\|^{-1}$ we obtain:

$$\begin{aligned} \mathcal{J}(K',L) - \mathcal{J}(K,L) =& \mathbb{E}_{x_0}\sum_{k=0}^{\infty} A_{KL}(x_k, K'x_k, Lx_k) \\ =& -2\eta \mathrm{Tr}(\Sigma_{K'L}(E_{KL}^k)^\top E_{KL}^k) \\ &+ \eta^2 \mathrm{Tr}(\Sigma_{K'L}(E_{KL}^k)^\top (R + B^\top P_{KL}B)E_{KL}^k) \\ \leq& -2\eta \mathrm{Tr}(\Sigma_{K'L}(E_{KL}^k)^\top E_{KL}^k) \\ &+ \|(R + B^\top P_{KL}B)\|\eta^2 \mathrm{Tr}(\Sigma_{K'L}(E_{KL}^k)^\top E_{KL}^k) \\ \leq& -\eta \mathrm{Tr}(\Sigma_{K'L}(E_{KL}^k)^\top E_{KL}^k) \\ \leq& -\eta\sigma_{\min}(\Sigma_0)\mathrm{Tr}((E_{KL}^k)^\top E_{KL}^k) \\ \leq& -\frac{\eta\sigma_{\min}(\Sigma_0)\sigma_{\min}(R)}{\|\Sigma_{K^\star L}\|}(\mathcal{J}(K,L) - \mathcal{J}(K^\star,L)). \end{aligned} \tag{4-3.17}$$

Again, we have to make the remark that η must be sufficiently small, yet upper-bounded by $\|R + B^\top P_{KL}B\|^{-1}$. Similarly for L, under $\mu \leq \|\gamma^2 I_f - F^\top P_{KL}F\|^{-1}$ we get

$$\begin{aligned} \mathcal{J}(K,L') - \mathcal{J}(K,L) =& \mathbb{E}_{x_0}\sum_{k=0}^{\infty} A_{KL}(x_k, Kx_k, L'x_k) \\ =& 2\mu \mathrm{Tr}(\Sigma_{KL'}(E_{KL}^\ell)^\top E_{KL}^\ell) \\ &+ \mu^2 \mathrm{Tr}(\Sigma_{KL'}(E_{KL}^\ell)^\top (-\gamma^2 I_f + F^\top P_{KL}F)E_{KL}^\ell) \\ \geq& 2\mu \mathrm{Tr}(\Sigma_{KL'}(E_{KL}^\ell)^\top E_{KL}^\ell) \\ &- \|(-\gamma^2 I_f + F^\top P_{KL}F)\|\mu^2 \mathrm{Tr}(\Sigma_{KL'}(E_{KL}^\ell)^\top E_{KL}^\ell) \\ \geq& \mu \mathrm{Tr}(\Sigma_{KL'}(E_{KL}^\ell)^\top E_{KL}^\ell) \\ \geq& \mu\sigma_{\min}(\Sigma_0)\mathrm{Tr}((E_{KL}^\ell)^\top E_{KL}^\ell) \\ \geq& \frac{\mu\sigma_{\min}(\Sigma_0)\theta}{\|\Sigma_{KL^\star}\|}(\mathcal{J}(K,L^\star) - \mathcal{J}(K,L)). \end{aligned} \tag{4-3.18}$$

To remove the explicit dependency on P_{KL} in stepsize bounds, we use the same bound as in [FGKM18], namely $\mathcal{J}(K,L) = \langle P_{KL}, \Sigma_0\rangle \geq \|P_{KL}\|_2\sigma_{\min}(\Sigma_0)$, since $P \in \mathcal{S}_+^n$. Then we can constrain the stepsizes using

$$\eta \leq \frac{1}{\|R\| + \|B\|^2\mathcal{J}(K,L)\sigma_{\min}(\Sigma_0)^{-1}}, \quad \mu \leq \frac{1}{\theta + \|F\|^2\mathcal{J}(K,L)\sigma_{\min}(\Sigma_0)^{-1}}.$$

Instead of removing P_{KL}, we rather removed the dependency on A, which is indeed often the most "unknown" part of a dynamical system. Like before, by construction of the inequalities (4-3.17),(4-3.18) we have:

$$\mathcal{J}(K',L^\star)-\mathcal{J}(K',L') \leq \left(1-\frac{\mu\sigma_{\min}(\Sigma_0)\theta}{\|\Sigma_{K'L^\star}\|}\right)(\mathcal{J}(K',L^\star)-\mathcal{J}(K',L)),$$
$$\mathcal{J}(K',L')-\mathcal{J}(K^\star,L') \leq \left(1-\frac{\eta\sigma_{\min}(\Sigma_0)\sigma_{\min}(R)}{\|\Sigma_{K^\star L'}\|}\right)(\mathcal{J}(K,L')-\mathcal{J}(K^\star,L'))'.$$

Now we can plug in the stepsize bounds and obtain the contractions:

$$\begin{aligned}\mathcal{J}(K',L^\star)-\mathcal{J}(K',L') &\leq \left(1-\frac{\theta}{\theta+\|F\|^2\mathcal{J}(K',L')\sigma_{\min}(\Sigma_0)^{-1}}\frac{\sigma_{\min}(\Sigma_0)}{\|\Sigma_{K'L^\star}\|}\right)\cdots\\ &\quad\cdots(\mathcal{J}(K',L^\star)-\mathcal{J}(K',L)),\\ \mathcal{J}(K',L')-\mathcal{J}(K^\star,L') &\leq \left(1-\frac{\sigma_{\min}(R)}{\|R\|+\|B\|^2\mathcal{J}(K',L')\sigma_{\min}(\Sigma_0)^{-1}}\frac{\sigma_{\min}(\Sigma_0)}{\|\Sigma_{K^\star L'}\|}\right)\cdots\\ &\quad\cdots(\mathcal{J}(K,L')-\mathcal{J}(K^\star,L')).\end{aligned}$$

Convergence is proved as before. To summarize the "*Gauss-Seidel*"-version of the algorithm, we use

$$\begin{aligned}K_{k+1}&=K_k-\eta_k\frac{1}{2}\nabla_K\mathcal{J}(K_k,L_k)\Sigma^{-1}_{K_kL_k}=K_k-\eta_kE^k_{K_kL_k}\\ L_{k+1}&=L_k+\mu_k\frac{1}{2}\nabla_L\mathcal{J}(K_{k+1},L_k)\Sigma^{-1}_{K_{k+1}L_k}=L_k+\mu_kE^\ell_{K_{k+1}L_k}.\end{aligned} \tag{4-3.19}$$

where the sequences of stepsizes $\{\eta_k\}_k$, $\{\mu_k\}_k$ can, at least, be constrained by

$$\eta_k \leq \frac{1}{\|R\|+\|B\|^2\mathcal{J}(K_k,L_k)\sigma_{\min}(\Sigma_0)^{-1}}, \quad \mu_k \leq \frac{1}{\gamma^2+\|F\|^2\mathcal{J}(K_{k+1},L_k)\sigma_{\min}(\Sigma_0)^{-1}}. \tag{4-3.20}$$

Note that we changed θ into γ^2, which is just a more conservative, yet simple stepsize.

To be complete, we again summarize these findings in a compact Lemma without proof (see Lemma 4-3.1 and the construction above).

Lemma 4-3.6 (Natural gradient algorithm). *Let $(K_0,L_0)\in\mathcal{KL}$ for some (potentially partially unknown) 7-tuple $(A,B,F,Q,R,\gamma,\Sigma_0)$ and assume that $K^\star\in\mathcal{KL}|_{L_0}$, $L^\star\in\mathcal{KL}|_{K_0}$. Now, consider the algorithm (4-3.19) under the stepsize constraints (4-3.20) and select sufficiently small $\{\eta_k\}_k$, $\{\mu_k\}_k$, at least satisfying the constraint. Moreover, assume that γ is chosen such that $(\gamma^2I_f-F^\top P_{K_{k+1}L_k}F)\succ 0\ \forall k$. Then, there is a $\beta\in(0,1)$ (as derived from above) such that if we let $N\in\mathbb{N}$ satisfy*

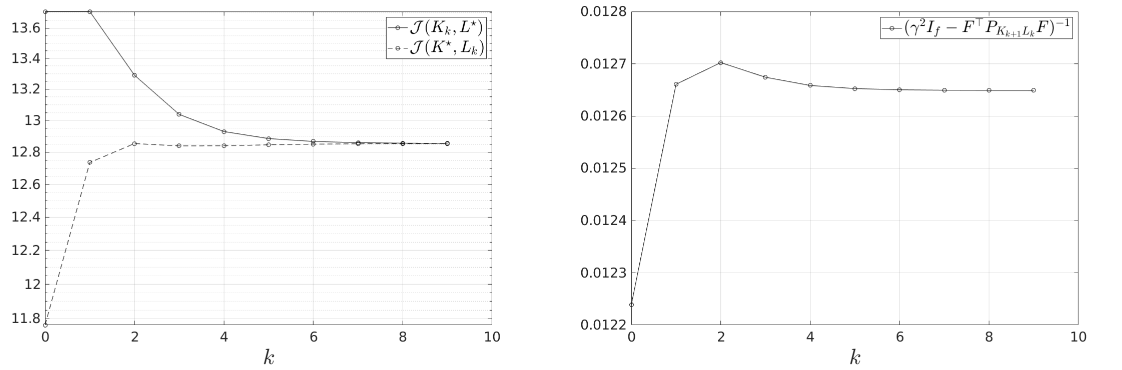
$$N\geq\frac{1}{\beta}\log_e\left(\frac{\mathcal{J}(K_0,L^\star)-\mathcal{J}(K^\star,L_0)}{\varepsilon}\right)$$

for some $\varepsilon\geq 0$ and run the algorithm for N or more steps, we obtain:

$$\mathcal{J}(K_N,L^\star)-\mathcal{J}(K^\star,L_N)\leq\varepsilon.$$

Remark 4-3.7 (Stronger demands than dynamic game feasibility). *From the definition and construction of β in Lemma 4-3.6 we see that to converge, we need $R\succ 0$ and $\Sigma_0\succ 0$, while this is by no means a demand for a dynamic game to be well-defined. Especially the constraint on Σ_0 cannot be relaxed.*

Example 4-3.8 (Natural gradient method)**.** *Consider the exact same model as in Example 4-3.4 with $\gamma = 10$ and let $\Sigma_0 = I_2$. Now we simulate the algorithm (4-3.19) under the upper-bounds of the stepsizes. As shown in Figure 4-6 the algorithm converges again as predicted. When compared with the Gauss-Newton method we observe slightly slower convergence (see Figure 4-1a and Figure 4-6a), which was to be expected with a gradient- and newton-step in mind.*



(a) The upper- and lower-bound converge to each other.

(b) γ is sufficiently large such that the analysis holds throughout

Figure 4-6: For $\gamma = 10$ the algorithm converges, just like the Gauss-Newton method.

Now if we let γ to be 6.5, it seems to still work, in contrast to the Gauss-Newton method (see Figure 4-7). The explanation can be that this natural gradient method is less ill-conditioned, we have no "Hessian" anymore. However, since we took the same initial condition as before, we still have $L^\star \notin \mathcal{KL}|_{K_0}$ and therefore we cannot conclude on convergence to the optimal pair $(K^\star, L^\star)$. Nevertheless, we do converge to the correct saddle-point at the cost of having the upper-bound being undefined (∞) for the first few iterations. Note that our theory so far has been sufficient*, therefore failures of upper-bounds etc. do not prohibit convergence.*

4-3-3 Simple Setting

In this last section we will consider a slightly less general setting and see that the results from [FGKM18, BMFM19] extend without any effort.

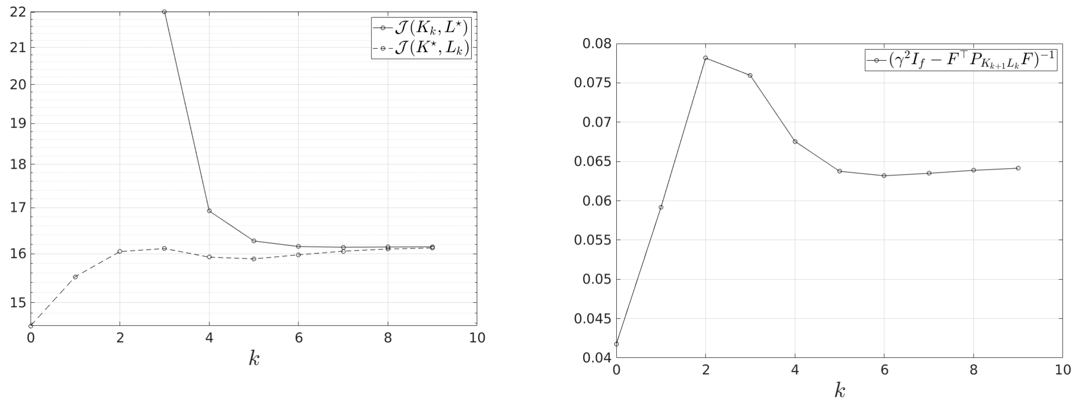
To that end, consider for some $\delta_K, \delta_L \in \mathbb{R}_{\geq 0}$ the following game

$$\begin{aligned} \inf_{\{u_k\}_{k\in\mathbb{N}}} \sup_{\{v_k\}_{k\in\mathbb{N}}} \quad & \mathbb{E}_{x_0}\left[\sum_{k=0}^{\infty} x_k^\top Q x_k + \delta_K^{-1} u_k^\top u_k - \delta_L^{-1} v_k^\top v_k\right], \\ \text{subject to} \quad & x_{k+1} = Ax_k + B(u_k + v_k), \quad x_0 \sim \mathcal{P}(0, \Sigma_0). \end{aligned} \tag{4-3.21}$$

Note that we have set $B = F$. Then we know that $L^\star = -\delta_L \delta_K^{-1} K^\star$, which might greatly simplify the analysis, while still being a practically relevant problem. Later on we will again generalize this.

Now we know that $\mathcal{J}(K) = \langle P_K, \Sigma_0 \rangle$ for

$$P_K = Q + \left(1 - \frac{\delta_L}{\delta_K}\right) \frac{1}{\delta_K} K^\top K + \left(A + B\left(1 - \frac{\delta_K}{\delta_L}\right) K\right)^\top P_K \left(A + B\left(1 - \frac{\delta_L}{\delta_K}\right) K\right),$$



(a) The upper- and lower-bound still converge to each other.

(b) The scaling factor remains positive throughout.

Figure 4-7: For $\gamma = 6.5$ the algorithm converges, in contrast to what we saw in Figure 4-3. However, theoretically, we fail to satisfy the conditions from Lemma 4-3.6 and indeed the upper-bound is not well-defined for all k.

such that for Σ_K, the usual symmetric positive semi-definite solution to the Discrete Lyapunov equation, we can write:

$$\nabla_K \mathcal{J}(K) = 2\left(1 - \frac{\delta_L}{\delta_K}\right)\left(\left[\delta_K^{-1} I_m + B^\top P_K B\left(1 - \frac{\delta_L}{\delta_K}\right)\right] K + B^\top P_K A\right)\Sigma_K.$$

Under these observations, the Quasi-Newton algorithm would be of the form

$$K_{k+1} = K_k - \frac{1}{2}\eta_k\left(\delta_K^{-1}(1 - \delta_L\delta_K^{-1})I_m + B^\top P_{K_k} B(1 - \delta_L\delta_K^{-1})^2\right)^{-1}\nabla_K \mathcal{J}(K_k)\Sigma_{K_k}^{-1}.$$

We could go through all the analysis again, but there is shortcut. Under the transformations given by $\widetilde{B} = \left(1 - \delta_L\delta_K^{-1}\right)B$, $\widetilde{R} = \left(1 - \delta_L\delta_K^{-1}\right)R$, we can just solve a standard LQR problem and indeed apply all the theory from the aforementioned papers. Of course, we need some basic conditions on (δ_K, δ_L) (for now $R = \delta_K^{-1}$). Thus, this section is more or less a corollary to [FGKM18, BMFM19].

For example, we can simply apply section 5.1 from [BMFM19], such that using

$$K_{k+1} = K_k - \eta_k \nabla_K \mathcal{J}_K(K_k)\Sigma_{K_k}^{-1}, \quad K_0 \in \mathcal{S}(A, \widetilde{B}) \tag{4-3.22}$$

with stepsize $\eta_k = 0.5/\lambda_{\max}\left(\widetilde{R} + \widetilde{B}^\top P_{K_k}\widetilde{B}\right)$ provably converges, at a linear rate to $K^\star$, whereafter $L^\star$ can simply be retrieved via $L^\star = -\delta_L\delta_K^{-1}K^\star$. Here we mean by $\mathcal{S}(A, B)$ the set of controllers stabilizing (A, B). This approach is of course reminiscent of [ZYB19], but they do not *use* the simple relation between $K^\star$ and $L^\star$.

Note that so far we did the transformations for δ_K and δ_L both being simply scalars, but without any troubles we can apply the same idea to invertible pairs of cost- and even input matrices. In such a general case, where $B \neq F$, we might have things like $L^\star = -R_v^{-1}F^\top B^{-\top}RK^\star$. To be clear, this follows from $K^\star = -R^{-1}B^\top P_{KL}\Lambda_{KL}^{-1}A$, $L^\star = \gamma^{-2}F^\top P_{KL}\Lambda_{KL}^{-1}A$. Other authors usually rely on the lengthy expression, from where the relation between $K^\star$ and $L^\star$ is less obvious.

It can be argued that this approach has already plenty of applications. Moreover, the case where $B = F$, not per se being invertible, is possibly the most natural in practice. Even more so, although we get $L^\star$ for free, in practice one might be solely interested in the controller $K^\star$, which protects the system from $L^\star$.

We can turn these observations into a simple Lemma

Lemma 4-3.9. *Consider the game*

$$\begin{aligned} \inf_{K\in\mathbb{R}^{m\times n}} \sup_{L\in\mathbb{R}^{f\times n}} \quad & \mathbb{E}_{x_0}\left[\sum_{k=0}^{\infty} x^\top (Q + K^\top R_u K - L^\top R_w L) x_k\right], \\ \text{subject to} \quad & x_{k+1} = (A + BK + FL)x_k, \quad x_0 \sim \mathcal{P}(0, \Sigma_0). \end{aligned} \tag{4-3.23}$$

If (4-3.23) is feasible, $R_u, R_w \succ 0$ and there is a $Z \in \mathbb{R}^{m\times f}$ such that $F = BZ$, then (4-3.23) can be approached as a standard LQR problem under the transformations $B \leftarrow B(I - ZR_w^{-1}Z^\top R_u)$, $R_u \leftarrow R_u(I - ZR_w^{-1}Z^\top R_u)$, after which $L^\star = -R_w^{-1}Z^\top R_u K^\star$.

4-4 Preliminary Conclusions

Regarding practical applications of policy iteration in the context of LQ games, the simpler LQR setting seems to provide a sufficient arsenal of tools. Theoretically speaking, it remains an interesting problem with many open ends as further discussed below. It turns out that (part of) the paper [FGKM18] can be extended to our setting (although we miss explicit stepsize bounds, the generic gradient algorithm and the data-driven extension). The crux might be not to overly generalize optimal control problems, but to directly exploit properties of the actual problem at hand, in this case the quadratic cost. Substantial open problems remain. How to extend the pure gradient algorithm? How to properly include data? How to find sharper convergence rates ([BMFM19] proposed some ideas), especially to compare "*Jacobi*" and "*Gauss-Seidel*"-like algorihms? At this point we *bound* them with the same rate (see Lemma 4-3.1). At last, it would be interesting to make the convergence conditions necessary and sufficient instead of just sufficient. As Example 4-3.8 shows, convergence is possible under less stringent conditions, but at the cost of having an upper-bound attaining ∞ for a few iterations. However, we would like to make another remark. Although this model-free approach is interesting and does provide new insights, *e.g.,* Corollary 4-3-1, it is amazingly data-inefficient. The hope is that this policy gradient method allows for controller synthesis under constraints since we can turn the gradient-algorithms into projected-gradient algorithms. However, it is not clear how the community benefits from this approach since as frequently mentioned before, the domain is generally non-convex, which will turn most constraints into impossible-to-verify assumptions.

Chapter 5

After the Ending

We set out to find if dynamic game theory would give rise to a structurally nice uncertainty set; a set which would lead to exact and tractable finite-sample control algorithms. We showed that even the most basic dynamic game has a rich structure which can be exploited in the study of robust linear-quadratic optimal control problems. Indeed, as remarked throughout chapter 3, our framework is not particularly useful in the context of unbiased estimators. The moral might be, to cope with perceived conservatism one could relax the minimax problem, but it is wiser to first compare the geometry of your data-driven uncertainty set and the *full* set considered by the robust optimizer. When your glass is however half-full, then it should be remarked that our game theoretic control law might be an appropriate substitution of the nominal LQ regulator in the context of ℓ_2-regularized linear Least-Squares identification.

5-1 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.* [Tu19])? 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 provide expressions for inscribed norm-balls in section 3-2-3, but note again that they are usually inherently small,*e.g.*, see Figure 3-15b, an inscribed ball around 0 is significantly smaller than the set itself. Section 3-4-2-4 hints at concentration possibilities for our set, but rigorous results are an open problem. In line with the remarks from sections 3-4-3-1-3-4-4, 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 3-4-4, especially regarding an uncertain B matrix, be further formalized? It is especially interesting to note that we sample from a line (path $p(\gamma)$) in $\mathbb{R}^{m\times n}$, not some ball around $K^{\star}(0)$ (see Figure 5-1). 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. This map might also shed some light on *why*

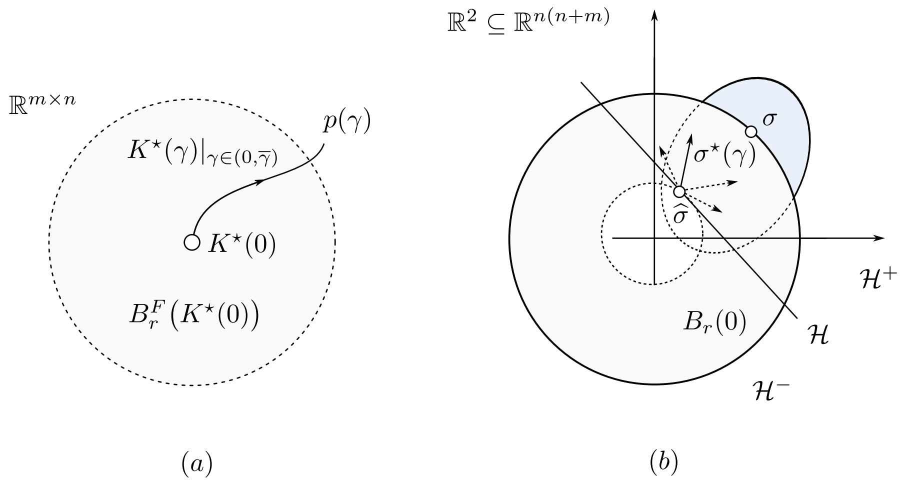


Figure 5-1: (a) Throughout the selection methods in section 3-4-4 we select $K^\star(\gamma)$ from a grid on $p(\gamma)$, not some ball $B_r^F(K^\star(0))$. Still, we can outperform $K^\star(0)$, which is in favour of the theoretically justified intuition skected out in Figure 3-27. (b) Copying Figure 3-27 $(\lambda > 0)$ into Figure 5-1 (b), we drew *a potential* $\widehat{\sigma}^\star(\gamma)$, but all we know from Lemma 3-3.3 is that this vector, for some arbitrarily small $\delta > 0$, could have been any of the other dashed arrows pointing in $\mathcal{H}^+ := \mathbb{R}^2 \setminus \{\mathcal{H} \cup \mathcal{H}^-\}$.

it works, Figures 3-7-3-27 are justified intuition, but it is hypothesized that the structural preservation we see throughout, *e.g.,* Corollary 3-3.6, also comes into play. Additionally, we need a better understanding of the *direction* of the worst-case path; see Figure 5-1 (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 $\mathcal{A}_\gamma(\widehat{A} + \widehat{B}K^\star(\gamma))$ be further formalized as a function of γ? Examples in section 3-4-2 hinted at this possibility, *e.g.,* is there a generic method to find γ_t, such that for all $\gamma < \gamma_t$ the nominal- and worst-case drift are topologically equivalent? It would be interesting to investigate *all* the possible model classes. For example, using the same ideas as in section 3.5.1 from [BB95] it can be shown that *cross-terms* in the cost can be interpreted as an offset in the uncertainty set, *e.g.,* we get $\left\langle (\Delta_A - \Delta_o)^\top (\Delta_A - \Delta_o), \Sigma_x \right\rangle \leq \gamma$ for some Δ_o. Moreover, in section 3-4-6 we briefly show that our framework can handle more than a simple "*diagonal*" game. Nevertheless, this example also showed the limitations of the current utility function, leading to certainty equivalence (CE), *i.e.,* $K^\star(\gamma)$ is optimal for any Σ_v, $v_k \overset{i.i.d.}{\sim} \mathcal{P}(0, \Sigma_v)$. To properly incorporate distributional uncertainties we indeed suggest a further exploration of non-trivial utility functions, *e.g.,* the LEQR (3-4.17) formulation. Furthermore, can we formalize *quantitative* properties? At which value of γ does our set become non-convex in $\mathbb{R}^{d\times n} \times \mathbb{R}^{d\times m}$? Even more so, 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 author believes further investigations are worthwhile, improving our understanding of how to efficiently link identification- and control algorithms towards safe data-driven control. Nevertheless, the main future challenge would be to bring the Optimization-, Control-, Dynamical Systems- and especially Statistics communities (even further) together.

Appendix A

Auxiliary Tools

The following Lemma, in its original form by Peyman Mohajerin Esfahani, is the key to bridge the RLQR problem (2-1.2) under uncertainty sets from Definition 3-1.1 to a dynamic game theory perspective.

Lemma A-0.1 (Exact constraint relaxation). *Let f, g be functions from $\mathcal{X}$ to $\mathbb{R} \cup \{\infty\}$. Given a parameter $\gamma \geq 0$, we define the optimization programs*

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

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

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

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

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

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

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

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). □

Next, we provide two trace inequalities.

Lemma A-0.2. *Given some $n \times n$-dimensional symmetric matrices $A \succ 0$ and $B \succ 0$, then for $X \succ 0$*

$$\mathrm{Tr}(XA) \geq \mathrm{Tr}(XB) \implies \kappa(X)\mathrm{Tr}(A) \geq \|B\|_2.$$

Proof. First, let $V\Sigma V^\top$ be the SVD of X, then

$$\mathrm{Tr}(XA) \geq \mathrm{Tr}(XB) \implies \sigma_{\max}(X)\mathrm{Tr}(VV^\top A) \geq \mathrm{Tr}(XB).$$

Then since $X \succ 0$ we know $\mathrm{Tr}(XB) \geq \|B\|_2 \sigma_{\min}(X)$ such that

$$\frac{\sigma_{\max}(X)}{\sigma_{\min}(X)}\mathrm{Tr}(A) \geq \|B\|_2.$$

□

Lemma A-0.3. *Given some $X, Y, Z \in \mathcal{S}^n_{++}$, Let $Y \succ Z \succ 0$, and $[X, Y] = 0$, $[X, Z] = 0$, for $[Q, R] := QR - RQ$. Then,* $\mathrm{Tr}(XY) > \mathrm{Tr}(XZ)$.

Proof. Let $A := Y - Z \succ 0$, such that by linearity of $\mathrm{Tr}(\cdot)$ and $[X, Y] = 0$, $[X, Z] = 0$ we get $\mathrm{Tr}(X^{1/2} A X^{1/2}) > 0$, hence the result. □

Appendix B

Further Background Information

In this appendix we provide further explanations regarding topics assumed to be (partially) known within the main part.

B-1 Degenerate Quadratic Programs

We start with a very useful result which shows up everywhere once you work with a quadratic cost. Consider the optimization problem for $C \succ 0$:

$$\inf_u f(x,u) := \inf_u \begin{pmatrix} x \\ u \end{pmatrix}^\top \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}.$$

Then the solution and cost are given by $u^\star = -C^{-1}B^\top x$, $f(x,u^\star) = x^\top(A - BC^{-1}B^\top)x$. Now what happens when $C \succeq 0$ instead? We can without loss of generality consider the QP given by

$$\inf_u \frac{1}{2}u^\top Hu + x^\top u, \quad H \succeq 0.$$

If a solution exists, it must satisfy $Hu^\star = -x$. To see this, let $H = Q\Lambda Q^\top$ (H is without loss of generality symmetric) and consider the equivalent program

$$\inf_{z=Q^\top u} = \frac{1}{2}z^\top \Lambda z + r^\top z = \inf_z \sum_{i=1}^{n} \frac{1}{2}z_i^2\lambda_i + z_i r_i,$$

for $r = Q^\top x$. This effectively decouples the problem and yields a set of scalar problems. We know that $\lambda_i \geq 0\ \forall i$, but also that optimizing z_i must satisfy $z_i^\star \lambda_i = -r_i$. When this relation is transformed back we get the relation $Hu^\star = -x$. In other words, if $r_i \neq 0$ while $\lambda_i = 0$, the problem is unbounded, which is rather intuitive. This simple observation is used in section 3-4-2-2.

B-2 Introduction to Discrete Topological Dynamical Systems

Here we give a brief introduction into topological equivalence of dynamical systems[1].

It is widely known that we call two topological spaces X and Y *topologically* equivalent if there is some homeomorphism $\varphi \in C^r(X, Y)$, $r \geq 0$. Now consider the subspaces $c_1 \subseteq X$, $c_2 \subseteq Y$, which can indeed be some trajectories (curves). Then we call c_1 and c_2 topologically equivalent if $\psi = \iota_2^{-1} \circ \varphi \circ \iota_1$ is a homeomorphism, *i.e.,* the diagram

$$\begin{tikzcd} c_1 \arrow[r, "\psi"] \arrow[d, "\iota_1"'] & c_2 \arrow[d, "\iota_2"] \\ X \arrow[r, "\varphi"] & Y \end{tikzcd}$$

commutes. Here, ι is an inclusion map, *e.g.* $\iota_1 : c_1 \hookrightarrow X$. For an illustrative example, see Figure B-1, or see [BB00] for a beautiful example in control theory. To extend these

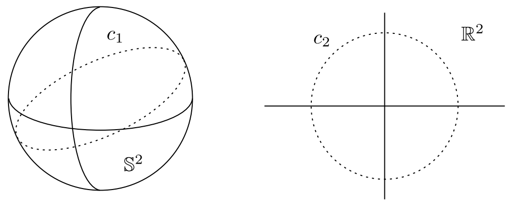


Figure B-1: Although c_1 and c_2 might appear to be "equivalent", they are not, topologically speaking. This because the spaces these curves are embedded in are not topologically equivalent.

ideas to dynamical systems, introduce the semigroup[2] G_i with the actions $g \in G_i$ defined by $g \cdot x = A_i^k x$, $k \in \mathbb{Z}_{\geq 0}$. Then, define two orbits, $O_1(x)$ and $O_2(x)$ by $O_i(x) := \{g \cdot x \ : \ g \in G_i\}$. These orbits allow for (some) partitions $X = \bigsqcup_x O_1(x)$, $Y = \bigsqcup_y O_2(y)$, note that the (disjoint) unions are not per se over all x and y (consider a limit cycle). Now we call the dynamical systems defined by G_1 (A_1) and G_2 (A_2) topologically equivalent if we can find a homeomorphism φ and bijective map h such that $\varphi(G_1(x)) = G_2(h(x))$ $\forall x \in X$. In other words, each orbit $O_1(x)$ is homeomorphic to some other orbit $O_2(y)$, where the selection is made by h. Here we impose a slight, but most natural, restriction on h, namely to be equivalent to φ. This leads us to definition 3-3.1. For example, to continue with the scalar system from section 2-2-4, there, the corresponding homeomorhisms follow from proposition 1.5 in [KR73] and are given by $\varphi(x) = x|x|^{c-1}$ with $c = \log_{|a_2|}(|a_1|)$ such that for example $a_1 x \circ \varphi = \varphi \circ a_2 x$, $\forall\, a_1, a_2 \in (0, 1)$.

In some sense, the goal of topological dynamical systems is to reduce the amount of effort, instead of a continuum of systems, study a (hopefully) finite set. For example, for scalar linear maps we have just 7 equivalence classes [KR73]. This should be contrasted with linear equivalence (similarity transformations), there we do have a continuum (since eigenvalues are preserved and they compromise $\mathbb{C}^n$). Take for example the scalar maps $a_1 x$ and $a_2 x$ which are only linearly equivalent if they are identical.

[1] Effectively elaborating on the superb https://www.encyclopediaofmath.org/index.php/Topological_equivalence.

[2] We do not demand $A \in \mathsf{GL}(n, \mathbb{R})$.

B-3 Compact Expressions in Dynamic Game Theory

In Lemma 3-2.9 we use the compact expression for $P, K^{\star}$ and $L^{\star}$, all hinging on this matrix Λ. Similar expressions can be found in [BB95], it was however not immediately clear how to arrive at these expressions. As it turns out, understanding these Linear Algebraic tricks is very useful to identify familiar expressions, like we did in section 3-4-6.

Reconsider the game (3-2.16), but for the sake of corresponding references in notation similar to [BB95], so let $\delta^{-1} \triangleq \gamma^2$. We will briefly highlight where the (compact) optimal policy expressions come from.

For a derivation of the deterministic results with $\alpha = 1$ we refer to theorem 6.4 from [BO99] or theorem 3.1 from [BB95]. We start by truncation of the horizon to $K < \infty$.

Assume that duality holds such that we can consider solving the robust Bellman (or Isaacs) equation

$$V_k = \max_{w_k} \min_{u_k} \left\{ x_k^\top Q x_k + u_k^\top R u_k - \gamma^2 w_k^\top w_k + \alpha \mathop{\mathbb{E}}_{x_0, v} [V_{k+1} | x_k] \right\}.$$

To solve this equation assume that $V_k = x_k^\top P_k x_k + q_k$, then we find

$$\begin{aligned} x_k P_k x_k + q_k = \max_{w_k} \min_{u_k} & \left\{ \begin{pmatrix} x_k \\ w_k \\ u_k \end{pmatrix}^\top \left[\begin{pmatrix} Q & 0 & 0 \\ 0 & -\gamma^2 I & 0 \\ 0 & 0 & R \end{pmatrix} \right. \right. \\ & \left. \left. + \alpha \begin{pmatrix} A^\top P_{k+1} A & A^\top P_{k+1} D & A^\top P_{k+1} B \\ D^\top P_{k+1} A & D^\top P_{k+1} D & D^\top P_{k+1} B \\ B^\top P_{k+1} A & B^\top P_{k+1} D & B^\top P_{k+1} B \end{pmatrix} \right] \begin{pmatrix} x_k \\ w_k \\ u_k \end{pmatrix} \right\} \\ & + \alpha (\mathrm{Tr}(P_{k+1} \Sigma_v) + q_{k+1}). \end{aligned} \tag{B-3.1}$$

Like before, Schur complements provide us with the tools to find the optimal strategies. First for u_k, when $R \succ 0$[3]:

$$u_k^\star = -\alpha (R + \alpha B^\top P_{k+1} B)^{-1} \begin{pmatrix} B^\top P_{k+1} A & B^\top P_{k+1} D \end{pmatrix} \begin{pmatrix} x_k \\ w_k \end{pmatrix}.$$

We can plug $u_k^\star$ back in (B-3.1) and get

$$\begin{aligned} & x_k^\top P_k x_k + q_k \\ = \max_{w_k} & \left\{ \begin{pmatrix} x_k \\ w_k \end{pmatrix}^\top \left[\begin{pmatrix} Q + \alpha A^\top P_{k+1} A & \alpha A^\top P_{k+1} D \\ \alpha D^\top P_{k+1} A & -\gamma^2 I + \alpha D^\top P_{k+1} D \end{pmatrix} \right. \right. \\ & \left. \left. - \alpha^2 \begin{pmatrix} A^\top P_{k+1} B \\ D^\top P_{k+1} B \end{pmatrix} (R + \alpha B^\top P_{k+1} B)^{-1} \begin{pmatrix} B^\top P_{k+1} A & B^\top P_{k+1} D \end{pmatrix} \right] \begin{pmatrix} x_k \\ w_k \end{pmatrix} \right\} \\ & + \alpha (\mathrm{Tr}(P_{k+1} \Sigma_v) + q_{k+1}). \end{aligned}$$

This max step is well-defined since we have $-\gamma^2 I + \alpha D^\top P_{k+1} D \prec 0$. Again, the Schur complements deliver the optimal policy and cost-to-go, nevertheless as rather long expressions.

[3] Or, when relaxed $(R + \alpha B^\top P_{k+1} B) \succ 0$.

Before writing it down like that, recall that for $A, C, \in \mathsf{GL}$, the matrix inversion lemma gives

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}. \tag{B-3.2}$$

This allows for writing $P - PB(R + B^\top PB)^{-1}B^\top P$ as $(P^{-1} + BR^{-1}B^\top)^{-1}$, if P is invertible of course. For P_{k+1} this cannot be assumed. What we can do is the following, write $P[I - \alpha B(R + \alpha B^\top PB)^{-1}B^\top P]$. Then the expression between the square brackets can be dealt with using the inversion lemma. We get

$$X := P - \alpha PB(R + \alpha B^\top PB)^{-1}B^\top P = P(I + \alpha BR^{-1}B^\top P)^{-1} =: PY, \tag{B-3.3}$$

where the definitions for X and Y will be useful later on. This allows for writing the rather long expression for P_k into

$$P_k = Q + \alpha A^\top P_{k+1}\left(I + \alpha\left(BR^{-1}B^\top - \frac{1}{\gamma^2}DD^\top\right)P_{k+1}\right)^{-1}A, \tag{B-3.4}$$

or as indeed often done: $P_k = Q + \alpha A^\top P_{k+1}\Lambda_k^{-1}A$.

Then for the optimal $w_k^\star$

$$\begin{aligned} w_k^\star = &-\left(-\gamma^2 I + \alpha D^\top P_{k+1}(I + \alpha BR^{-1}B^\top P_{k+1})^{-1}D\right)^{-1}\cdots \\ &\cdots \alpha D^\top P_{k+1}(I + \alpha BR^{-1}B^\top P_{k+1})^{-1}Ax_k =: L^\star x_k. \end{aligned}$$

This can however be simplified significantly, using X from (B-3.3) we can write

$$P_k = Q + \alpha A^\top XA - \alpha^2 A^\top XD(-\gamma^2 I + \alpha D^\top XD)^{-1}D^\top XA.$$

We know from before that $w^\star = -(-\gamma^2 I + \alpha D^\top XD)^{-1}\alpha D^\top XAx$. Now consider the following expression for $\widetilde{L}$:

$$\begin{aligned} \widetilde{L} =& \frac{1}{\gamma^2}\alpha D^\top P_{k+1}\Lambda_k^{-1}A, \\ =& \frac{1}{\gamma^2}\left(\alpha D^\top XA - \alpha^2 D^\top XD(-\gamma^2 I + \alpha D^\top XD)^{-1}D^\top XA\right) \\ =& \frac{1}{\gamma^2}\left(\alpha D^\top XA - \alpha\frac{1}{\gamma^2}D^\top XD(-I + \alpha\frac{1}{\gamma^2}D^\top XD)^{-1}\alpha D^\top XA\right), \\ =& \frac{1}{\gamma^2}\left(\left(I - \left(-I + \alpha\frac{1}{\gamma^2}D^\top XD\right)^{-1}\frac{1}{\gamma^2}D^\top XD\right)\alpha D^\top XA\right) \\ =& -(-\gamma^2 I + \alpha D^\top XD)^{-1}\alpha D^\top XA \\ =& L^\star. \end{aligned}$$

Where the first three steps are just rewriting, then we use $P(1 + QP)^{-1} = (1 + PQ)^{-1}P$. Finally to go from step 4 to 5 we use $(-I + P)^{-1} = -I + (-I + P)^{-1}P$[4]. Hence, we obtain the compact expression:

$$w_k^\star = \alpha\gamma^{-2}D^\top P_{k+1}\Lambda_k^{-1}Ax_k.$$

[4] $(-I + P)^{-1} = -(-I + P)^{-1}(-I + P - P) = -I + (-I + P)^{-1}P$

This can be plugged into $u_k^\star$, or use the symmetries in the problem to obtain:

$$u_k^\star = -\alpha R^{-1} B^\top P_{k+1} \Lambda_k^{-1} A x_k =: K^\star x_k.$$

At last, it can be checked by direct computation that

$$P_k = Q + (K^\star)^\top R K^\star - \gamma^2 (L^\star)^\top L^\star + \alpha A_{\text{cl}}^\top P_{k+1} A_{\text{cl}} \tag{B-3.5}$$

for $A_{\text{cl}} := \Lambda_k^{-1} A$. With LQR derivations in mind, this form is very useful and concludes our digression into where the game theoretic equations come from.

Note that we do not say anything about stability, we merely provide an explanation of why the equations can be written in this convenient compact form.

Extending the previous results to the infinite horizon case seems as obvious as done in the LQR case. The author of these notes is very much guilty of that thought. However, as indicated in [BB95, BO99, BLW91], for the infinite horizon case the resulting strategies are *not* per se saddle points. It turns out that we can effectively extended the practical part of the finite horizon theory (*cf.* section 3.4 from [BB95]), but not the terminology. This was first pointed out by [Mag76] whereafter [Jac77] showed that the saddle equilibrium still holds, but over a more restricted set of strategies. We effectively show this in chapter 4 as well.

B-4 Lyapunov Equations

Throughout we frequently use the so-called discrete-time Lyapunov equation of the form (B-4.1). We briefly mention the most simple method to solve this algebraic equation. Consider for $A \in \mathbb{R}^{n\times n}$ and some $W \succ 0$ the equation

$$\Sigma = A\Sigma A^\top + W. \tag{B-4.1}$$

Using the kronecker product identity $\text{vec}(ABC) = (C^\top \otimes A)\text{vec}(B)$ this can be rewritten into $(I_{n^2} - A\otimes A)\text{vec}(\Sigma) = \text{vec}(W)$. This allows for solving for Σ but note that since the dimensions of this linear system grow quadratically with n it might be computationally inefficient. Of course, to have a valid and meaningful solution we demand A to be stable.

Regarding α-Lyapunov equations: $\Sigma = \alpha A\Sigma A^\top + W$ the solution extends to $\text{vec}(\Sigma) = (I_{n^2} - \alpha A \otimes A)^{-1}\text{vec}(W)$. From here it especially clear that problems occur when the spectrum of A approaches $\alpha^{-1/2}$, indeed, then $\Sigma \to \infty$. Especially in the covariance context this has a clear interpretation, with too little contraction, cost accumulates quickly.

B-5 On Linear Least-Squares System Identification

The vast majority of linear system identification methods use some form of Least-Squares. In this section we discuss some of its properties and work towards clarifying the ellipsoidal shape from Figure 3-16b. We do assume some familiarity with the topic.

B-5-1 Regularization in Linear Least-Squares System Identification

Initially, regularization was introduced to tame ill-conditioned problems. Nowadays, there is clear dynamical systems interpretation in the context of Mean Square Error (MSE) estimation, to quote [LC13] "*To minimize the MSE is a trade off in constraining the model: a flexible model gives small bias (easier to describe complex behaviour) and large variance (with a flexible model it is easier to get fooled by the noise).*" In this section we shed some light on the use of ℓ_2-regularization in the context of Linear Least-Squares System Identification.

Given some unknown system $\Sigma : \{x_{k+1} = Ax_k + Bu_k$ parametrized by $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times n}$. Then, the most common method to obtain an estimate of (A, B) is to gather data $\{(x_k, u_k)\}_{k=0}^N$ and let

$$(\widehat{A}_N, \widehat{B}_N) = \operatorname*{argmin}_{(A,B)} \sum_{k=0}^{N-1} \|x_{k+1} - Ax_k - Bu_k\|_2^2. \tag{B-5.1}$$

However, is this problem always well-defined? The answer is clearly no. Towards finding conditions such that $\lim_{N\to\infty}(\widehat{A}_N, \widehat{B}_N) = (A, B)$, and perhaps more importantly, towards the existence of some $M \in \mathbb{N}$ such that $(\widehat{A}_N, \widehat{B}_N)$ is well-defined $\forall N \geq M$, we can look at what is called *persistence of excitation*. Following [GM86] (there are still multiple definitions around),

Definition B-5.1 (Persistent Excitation). *A sequence $\{z_i\}_{i=1}^{\ell}$, with $z_i \in \mathbb{R}^z$, is* ***persistently exciting*** *(PE) when*

$$\liminf_{\ell\to\infty} \frac{1}{\ell}\sum_{i=1}^{\ell} z_i z_i^\top \succ 0, \tag{B-5.2}$$

or locally, when there is a $C > 0$ such that $\sum_{i=k+\ell}^{k+\ell} z_i z_i^\top \succ CI_z$.

The main take away is that the PE condition (B-5.2) is not easily verifiable, hence conditions on the input sequence and/or dynamical system are sought. For example, the most heavily used trick [Moo87] is to make the input (partially) stochastic, like we did in section 3-4-4 and clarify below (section B-5-3). For simplicity of notation, assume that we have an autonomous system, hence only A is unknown. Then, using Linear Least-Squares, assuming that $\{x_k\}_k$ is (locally) PE, we obtain

$$\widehat{A}_N^\top = \left(\sum_{k=0}^{N-1} x_k x_k^\top\right)^{-1}\left(\sum_{k=0}^{N-1} x_k x_{k+1}^\top\right). \tag{B-5.3}$$

This expression immediately clarifies definition B-5.1, without a PE assumption the inverse in (B-5.3) would not be defined.

Now, one can study PE conditions, but in practice it is much simpler to simply add regularization to (B-5.1), in our case we use ℓ_2-regularization (sometimes called *ridge regression*). Recall that adding ℓ_2-regularization corresponds to turning $\inf_{x\in\mathbb{R}^n} \|Ax - b\|_2^2$ into $\inf_{x\in\mathbb{R}^n} \|Ax - b\|_2^2 + \lambda\|x\|_2^2$ for some $\lambda \in \mathbb{R}_{>0}$, which is again equivalent to $\inf_{x\in\mathbb{R}^n} \|Ax - b\|_2^2$ subject to $\|x\|_2 \leq \varepsilon_x$ for some ε_x related to λ.

To study the effects of ℓ_2-regularization, following [VV07, ch.4], let $y = Fx + v$, $v \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_v^2 I_p)$, where we can measure $y \in \mathbb{R}^p$ and would like to find an estimate of $x \in \mathbb{R}^n$. The

Linear Least-Squares solution is given by $\widehat{x} = (F^\top F)^{-1}F^\top y$ whereas the regularized solution becomes $\widehat{x}_\lambda = (F^\top F + \lambda I_n)^{-1}F^\top y$. Compare this to (B-5.3), indeed, $\lambda > 0$ introduces numerical stability. However at some cost, under ℓ_2-regularization we introduce a bias, since $\mathbb{E}_v[\widehat{x}_\lambda] = (F^\top F + \lambda I_n)^{-1}F^\top F x \neq x = \mathbb{E}_v[\widehat{x}]$. On the other hand we reduce variance. To see this, recall that $\mathrm{Var}(\widehat{x}) = \sigma_v^2(F^\top F)^{-1}$, which can be well above 0 when the problem is ill-conditioned (more on that later). Now, it can be found that $\widehat{x}_\lambda = X\widehat{x} = (F^\top F+\lambda I_n)^{-1}F^\top F\widehat{x}$. Hence, $\mathrm{Var}(\widehat{x}_\lambda) = X\mathrm{Var}(\widehat{x})X^\top$. In the scalar case we immediately see

$$\mathrm{Var}(\widehat{x}_\lambda) = \frac{f^4}{(f^2+\lambda)}\mathrm{Var}(\widehat{x}),$$

such that $\mathrm{Var}(\widehat{x}_\lambda) < \mathrm{Var}(\widehat{x})\ \forall \lambda \in \mathbb{R}_{>0}$. Similarly, in the vector-valued case:

$$\begin{aligned}\mathrm{Var}(\widehat{x}) - \mathrm{Var}(\widehat{x}_\lambda) =& \sigma_v^2\left((F^\top F)^{-1} - X(F^\top F)^{-1}X^\top\right),\\ =& \sigma_v^2 X\left(2\lambda(F^\top F)^{-2} + \lambda^2(F^\top F)^{-3}\right)X^\top,\end{aligned}$$

which follows from solving $XVX^\top = (F^\top F)^{-1}$ for V and plugging it back in. The conclusion still stands since the part between the brackets is strictly positive for all $\lambda > 0$.

The crux is the following, there are circumstances where an unbiased estimator has that much variance such that some biased estimator (with smaller variance) has a smaller MSE indeed, think of $F : \kappa(F^\top F) \gg 1$. In system identification this relates to barely satisfied PE conditions. Hence, the introduction of λ is not *just* to make the problem well-defined, the statistical performance of the corresponding estimator $\widehat{x}_\lambda$ might be preferred over that of $\widehat{x}$.

It must be remarked that "*regularization is a transient phenomenon* [LCM19], *e.g.,* asymptotic PE conditions are easily satisfied and we can take for example $\lambda(N) = \lambda_0/\sqrt{N}$. Thus, if $\widehat{A}_\lambda$ outperforms $\widehat{A}$, then this should occur in the finite-data regime, in the real world.

B-5-2 On the Asymptotic Normality of Linear Least-Squares Identification

This section is in part based on notes by Ping Yu[5]. As discussed before, under mild assumptions it can be shown that a linear Least-Squares estimator is *unbiased*. We will be however mostly concerned with the full distribution of the estimation error. Note that since we consider a dynamical system with $x_0 \sim \mathcal{N}(0, \Sigma_0)$, $\Sigma_0 \in \mathcal{S}^n_{++}$, the data points from a single trajectory are not independent, in contrast to the data resulting from some noisy linear map $y = Ax + v$.

Now, let $\xrightarrow{d}$ denote convergence in *distribution*, which is sometimes referred to as *weak* convergence. Then, following [Cai18, p.292] we use the following definition.

Definition B-5.2 (Asymptotically Normal Estimator). *An estimator $\widehat{\theta}_N$ of $\theta^\star$ is asymptotically normal if $\sqrt{N}(\widehat{\theta}_N - \theta^\star) \xrightarrow{d} \mathcal{N}(0, \Sigma)$ for $N \to \infty$.*

Consider the linear discrete-time system for some yet unknown $A \in \mathbb{R}^{n\times n}$:

$$x_{k+1} = Ax_k + v_k, \quad v_k \overset{i.i.d}{\sim} \mathcal{N}(0, \Sigma_v). \tag{B-5.4}$$

[5] http://web.hku.hk/~pingyu/6005/6005.htm.

To start, introduce a computationally simpler form (avoiding tensors)

$$x_{k+1} = (I_n \otimes x_k^\top)\mathrm{vec}(A^\top) + v_k = Y_k\theta + v_k. \tag{B-5.5}$$

To recover $\theta \in \mathbb{R}^{n^2}$, and thus A, let the system perform Z experiments and let all of them run for N timesteps. Now, take only the two last data points from each experiment: $\{x_N^{(z)}, x_{N-1}^{(z)}\}_{z=1}^{Z}$. This approach greatly simplifies the analysis and is used in practice, where in Reinforcement-learning language one might say we do just 1 experiment with Z episodes (or rollouts), all of length N.

Then, we obtain the following expression for our Least-Squares estimator of $\theta \in \mathbb{R}^{n^2}$:

$$\begin{aligned}\widehat{\theta}_Z &= \operatorname*{argmin}_{\theta\in\mathbb{R}^{n^2}} \frac{1}{Z}\sum_{z=1}^{Z} \|x_N^{(z)} - Y_{N-1}^{(z)}\theta\|_2^2 \\ &= \left(\frac{1}{Z}\sum_{z=1}^{Z} (Y_{N-1}^{(z)})^\top Y_{N-1}^{(z)}\right)^{-1}\left(\frac{1}{Z}\sum_{z=1}^{Z} (Y_{N-1}^{(z)})^\top x_N^{(z)}\right).\end{aligned}$$

Hence, the error is given by

$$\sqrt{Z}(\widehat{\theta}_Z - \theta^\star) = \left(\frac{1}{Z}\sum_{z=1}^{Z} (Y_{N-1}^{(z)})^\top Y_{N-1}^{(z)}\right)^{-1}\left(\frac{1}{\sqrt{Z}}\sum_{z=1}^{Z} (Y_{N-1}^{(z)})^\top v_{N-1}^{(z)}\right).$$

Now, assume that $\mathbb{E}\left[(Y_{N-1}^{(z)})^\top v_{N-1}^{(z)}\right] = 0$, $\mathbb{E}\left[\|v\|_2^4\right] < \infty$ and $\mathbb{E}\left[\|x\|_2^4\right] < \infty$.

To use a Central Limit Theorem (CLT), we need to bound the second moment, which can be established by bounding the expected norm of the elements:

$$\begin{aligned}\mathbb{E}\left[\left\|(Y_{N-1}^{(z)})^\top v_{N-1}^{(z)}(v_{N-1}^{(z)})^\top Y_{N-1}^{(z)}\right\|_F\right] &\leq \mathbb{E}\left[\left\|(Y_{N-1}^{(z)})^\top v_{N-1}^{(z)}(v_{N-1}^{(z)})^\top Y_{N-1}^{(z)}\right\|_F^2\right]^{1/2} \\ &\leq n^2\mathbb{E}\left[\left\|v_{N-1}^{(z)}\right\|_2^4\right]^{1/2}\mathbb{E}\left[\left\|x_{N-1}^{(z)}\right\|_2^4\right]^{1/2} < \infty.\end{aligned}$$

Where the inequalities follow from trace properties and Jensen's inequality, *i.e.,* let f be a convex function, then $\mathbb{E}[f(x)] \geq f(\mathbb{E}[x])$ such that from $f(x) = x^2$ it follows that $\mathbb{E}[X^2]^{1/2} \geq \mathbb{E}[X]$.

Therefore, by the CLT we have that

$$\left(\frac{1}{\sqrt{Z}}\sum_{z=1}^{Z} (Y_{N-1}^{(z)})^\top v_{N-1}^{(z)}\right) \xrightarrow{d} \mathcal{N}(0,\Sigma_1), \quad \Sigma_1 = \mathbb{E}\left[(Y_{N-1}^{(z)})^\top v_{N-1}^{(z)}(v_{N-1}^{(z)})^\top Y_{N-1}^{(z)}\right].$$

when $Z \to \infty$. Note that this step critically hinges on the trajectory independence.

Similarly, by the Law of Large numbers

$$\frac{1}{Z}\sum_{z=1}^{Z} (Y_{N-1}^{(z)})^\top Y_{N-1}^{(z)} \xrightarrow{p} \mathbb{E}\left[(Y_{N-1}^{(z)})^\top Y_{N-1}^{(z)}\right] =: \Sigma_2$$

for $Z \to \infty$. Then under the assumption that Σ_2 is invertible (usually a Persistent Excitation (PE) condition is imposed) the result follows from Slutsky's theorem.

Hence, under the aforementioned scheme

$$\sqrt{Z}(\widehat{\theta}_z - \theta^\star) \xrightarrow{d} \mathcal{N}(0, \Sigma_2^{-1}\Sigma_1\Sigma_2^{-1}) \tag{B-5.6}$$

for $Z \to \infty$.

The point is, indeed, for a sufficiently large Z, the Least-Squares estimator $\widehat{A}_Z$ is Normally distributed around the real A. Thus, after an embedding of A into $\mathbb{R}^{n^2}$ we observe these ellipsoidal sublevel-sets of estimates.

Like in [DMM+17], we turned to a simplified estimator with the main advantage being the immediate independence of the data. However, as pointed out in [MT19], although this "multiple rollout" scheme is easier to analyze, it is not data-efficient. Instead, one want to understand single-trajectory properties. This is still an active area of research, especially since the predominant solution method is to just add a load of noise here and there. Why noise might help is briefly outlined below.

B-5-3 Identification using Single-Trajectories

Consider the deterministic n-dimensional system $x_{k+1} = Ax_k$, applying standard linear least-squares identification using the data $\{x_k\}_{k=0}^N$ results in the estimator given by:

$$\widehat{A}_N^\top = \left(\sum_{k=0}^{N-1} A^k x_0 x_0^\top (A^k)^\top\right)^{-1} \left(\sum_{k=0}^{N-1} A^k x_0 x_0^\top (A^k)^\top\right) A^\top. \tag{B-5.7}$$

Indeed, since $x_0 x_0^\top$ is a rank 1 matrix, the estimator in (B-5.7) is not immediately well-defined. However, we can recognize the ***controllability grammian*** within, *e.g.*, for (A, B), let $W_i := \sum_{k=0}^{m-1} A^k BB^\top (A^k)^\top$. Indeed, $W_{m\geq n} \succ 0$ when the pair (A, B) is controllable, hence when (A, x_0) is controllable, then the estimator is unbiased, better yet, after $N \geq n$, we recover the solution: A.

Now, what about $x_{k+1} = Ax_k + Mv_k$, $v_k \overset{i.i.d.}{\sim} \mathcal{N}(0, \Sigma_v)$ and $x_0 = 0$? Since $x_k = \sum_{i=1}^k A^{k-i} M v_{i-1}$ we get

$$\begin{aligned}\widehat{A}_N^\top = &\left(\frac{1}{N}\sum_{k=0}^{N-1}\sum_{i=1}^{k} A^{k-1} M v_{i-1} v_{i-1}^\top M^\top (A^{k-i})^\top\right)^{-1} \left(\frac{1}{N}\sum_{k=0}^{N-1}\sum_{i=1}^{k} A^{k-1} M v_{i-1} v_{i-1}^\top M^\top (A^{k-i})^\top\right) A^\top \\ &+ \left(\frac{1}{N}\sum_{k=0}^{N-1}\sum_{i=1}^{k} A^{k-1} M v_{i-1} v_{i-1}^\top M^\top (A^{k-i})^\top\right)^{-1} \frac{1}{N}\sum_{k=0}^{N-1}\sum_{i=1}^{k} A^{k-1} M v_{i-1} v_k^\top.\end{aligned}$$

Now, given a sufficiently large $N \in \mathbb{N}$, is $\mathbb{E}_{v_0,\dots,v_{N-1}}[\widehat{A}_N] = A$? This is unfortunately a non-trivial evaluation, but the structure, resembling (B-5.7) is clear. If we additionally assume that v_k is *ergodic*, *i.e.*, $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N v_i = \mathbb{E}[v_k]$, we can provide an asymptotic result. Note, this is a natural assumption if v is zero-mean white-noise (see [Hay96, p.89]). Assume that $(A, M\Sigma_v^{1/2})$ is a controllable pair, then

$$\lim_{N\to\infty} \mathop{\mathbb{E}}_{v_0,\dots,v_{N-1}} [\widehat{A}_N^\top] = (W_n + W_t)^{-1}(W_n + W_t)A^\top + 0 = A^\top,$$

for $W_n = \sum_{k=0}^{n-1} A^k M \Sigma_v M^\top (A^k)^\top \succ 0$ and some $W_t \succeq 0$. Hence, for sufficiently large N and non-degenerate noise, least-squares presents us with an unbiased estimator.

For state of the art probabilistic finite-sample properties, see for example [SMT+18].

Bibliography

[Ack02] J. Ackermann. *Robust Control, the Parameter Space Approach.* Communications and Control Engineering. Springer, 2002.

[AL06] Krishna B. Athreya and Soumendra N. Lahiri. *Measure Theory and Probability Theory.* Springer Texts In Statistics. Springer, 2006.

[AL18] Marc Abeille and Alessandro Lazaric. Improved regret bounds for thompson sampling in linear quadratic control problems. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1–9. PMLR, 2018.

[AMR88] R Abraham, J.E. Marsden, and T Ratiu. *Manifolds, Tensor analysis, and Applications, Second Edition.* Applied Mathematical Sciences. Springer, 1988.

[ATLAK07] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf. Model-free q-learning designs for discrete-time zero-sum games with application to h-infinity control. In *2007 European Control Conference (ECC)*, pages 1668–1675, 2007.

[AW73] K.J. Astrom and B. Wittenmark. On self tuning regulators. *Automatica*, 9:185–199, 1973.

[AYPS11] Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 2312–2320. 2011.

[AYS11] Yasin Abbasi-Yadkori and Csaba Szepesvári. Regret bounds for the adaptive control of linear quadratic systems. *JMLR: Workshop and Conference Proceedings 19, 24th Annual Conference on Learning Theory*, 2011.

[Bar83] B. Barmish. Stabilization of uncertain systems via linear control. *IEEE Transactions on Automatic Control*, 28(8):848–850, August 1983.

[Bar94] B. Ross Barmish. *New Tools for Robustness of Linear Systems.* Macmillan, 1994.

[Bar02] Alexander Barvinok. *A Course in Convexity.* Graduate Studies in Mathematics. American Mathematical Society, 2002.

[BB95] Tamer Basar and Pierre Bernhard. *H_∞-Optimal Control and Related Minimax Design Problems A Dynamic Game Approach.* Birkhauser, 1995.

[BB00] Sanjay P. Bhat and Dennis S. Bernstein. A topological obstruction to continuous global stabilizatio of rotational motion and the unwinding phenomenon. *Systems & Control Letters*, 39:63–70, 2000.

[Ber76] Dimitri P. Bertsekas. *Dynamic Programming and Stochastic Control.* Academic Press, 1976.

[Ber99] Dimitri P. Bertsekas. *Nonlinear Programming.* Second Edition. Athena Scientific, 1999.

[Ber05] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control.* Volume 1. Athena Scientific, 2005.

[Ber07] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*. Volume 2, third edition. Athena Scientific, 2007.

[BGFB94] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

[BL04] Francesco Bullo and Andrew D. Lewis. *Geometric Control of Mechanical Systems*. Texts in Applied Mathematics. Springer, 2004.

[BLW91] Sergio Bittanti, Alan J. Laub, and Jan C. Willems, editors. *The Riccati Equation*. Springer-Verlag, Berlin, Heidelberg, 1991.

[BMFM19] Jingjing Bu, Afshin Mesbahi, Maryam Fazel, and Mehran Mesbahi. LQR through the Lens of First Order Methods: Discrete-time Case. *arXiv e-prints*, 2019.

[BO99] Tamer Basar and Geert Jan Olsder. *Dynamic Noncooperative Game Theory*. SIAM, 1999.

[BPR16] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry*. Springer, 2016.

[BPT12] G. Blekherman, P. Parrilo, and R. Thomas. *Semidefinite Optimization and Convex Algebraic Geometry*. Society for Industrial and Applied Mathematics, 2012.

[BYDM93] R. D. Braatz, P. M. Young, J. C. Doyle, and M. Morari. Computational complexity of μ calculation. In *1993 American Control Conference*, pages 1682–1683, 1993.

[Cai18] Peter E. Caines. *Linear Stochastic Systems*. Classics In Applied Mathematics. SIAM, 2018. Republication of the 1988 version.

[CDS$^+$09] C. Cobelli, C. Dalla Man, G. Sparacino, L. Magni, G. De Nicolao, and B. P. Kovatchev. Diabetes: Models, signals, and control. *IEEE Reviews in Biomedical Engineering*, 2:54–96, 2009.

[CK98] M. C. Campi and P. R. Kumar. Adaptive linear quadratic gaussian control: the cost-biased approach revisited. *SIAM J. CONTROL OPTIM.*, 36(6):1890–1907, 1998.

[CKM19] Alon Cohen, Tomer Koren, and Yishay 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*, pages 1300–1309, 2019.

[DB92] Garry Didinsky and Tamer Basar. Design of minimax controllers for linear systems with non-zero initial states under specified information structures. *International Journal of Robust and Nonlinear Control*, 2(1):1–30, 1992.

[Del84] David F. Delchamps. Analytic feedback control and the algebraic riccati equation. *IEEE Transactions on Automatic Control*, AC-29(11):1031–1033, 1984.

[DK99] J.J. Duistermaat and J.A.C. Kolk. *Lie Groups*. Springer, 1999.

[DMM$^+$17] Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. On the Sample Complexity of the Linear Quadratic Regulator. *arXiv e-prints*, 2017.

[DMM$^+$18] Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. Regret bounds for robust adaptive control of the linear quadratic regulator. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 4188–4197. Curran Associates, Inc., 2018.

[dOBG99] M.C. de Oliveira, J. Bernussou, and J.C. Geromel. A new discrete-time robust stability condition. *Systems & Control Letters*, 37(4):261 – 265, 1999.

[Doy78] John C. Doyle. Guaranteed margins for lqg regulators. *IEEE Transactions on Automatic Control*, AC-23(4):756–757, 1978.

[FB86] M. Fu and B. Barmish. Adaptive stabilization of linear systems via switching control. *IEEE Transactions on Automatic Control*, 31(12):1097–1103, 1986.

[FGKM18] Maryam Fazel, Rong Ge, Sham Kakade, and Mehran Mesbahi. Global convergence of policy gradient methods for the linear quadratic regulator. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1467–1476, Stockholmsmässan, Stockholm Sweden, 10–15 Jul 2018. PMLR.

[Fie97] Claude-Nicolas Fiechter. Pac adaptive control of linear systems. In *Proceedings of the Tenth Annual Conference on Computational Learning Theory*, COLT '97, pages 72–80, New York, NY, USA, 1997. ACM.

[Fol99] Gerald B. Folland. *Real Analysis*. Wiley-Interscience, 1999.

[GA10] M. S. Grewal and A. P. Andrews. Applications of kalman filtering in aerospace 1960 to the present [historical perspectives]. *IEEE Control Systems Magazine*, 30(3):69–78, 2010.

[GBA94] Germain Garcia, Jacques Bernussou, and Denis Arzelier. Robust stabilization of discrete-time linear systems with norm-bounded time-varying uncertainty. *Systems & Control Letters*, 22(5):327 – 339, 1994.

[Gev05] Michel Gevers. Identification for control: From the early achievements to the revival of experiment design*. *European Journal of Control*, 11(4):335 – 352, 2005.

[GL13] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The John Hopkins University Press, 2013. Fourth edition.

[GM86] Michael Green and John B. Moore. Persistence of excitation in linear systems. *Systems & Control Letters*, 7(5):351 – 360, 1986.

[GPM$^+$14] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks. *arXiv e-prints*, 2014.

[Hal70] Paul R. Halmos. *Measure theory*. Graduate Texts in Mathematics. Springer, 1970.

[Hay96] Monson H. Hayes. *Statistical Digital Signal Processing and Modeling*. Wiley, 1996.

[HJ90] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, 1990.

[HKvWV10] Ivo Houtzager, Balazs Kulcsar, Jan-Willem van Wingerden, and Michel Verhaegen. System identification for the control of wind turbine systems. In *The Science of Making Torque from Wind*, TORQUE 2010, pages 675–680, 2010.

[HLL96] Onésimo Hernández-Lerma and Jean Bernard Lasserre. *Discrete-Time Markov Control Processes, Basic Optimality Criteria*. Applications of Mathematics, Stochastic Modelling and Applied Probability. Springer, 1996.

[HLS09] D. Henrion, J. Lasserre, and C. Savorgnan. Approximate volume and integration for basic semialgebraic sets. *SIAM Review*, 51(4):722–743, 2009.

[HM77] Robert Hermann and Clyde F. Martin. Applications of algebraic geometry to systems theory - part i. *IEEE Transasctions on Automatic Control.*, 22(1):19–25, 1977.

[HS07] Lars Peter Hansen and Thomas J. Sargent. *Robustness*. Princeton University Press, 2007.

[Jac73] D. Jacobson. Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Transactions on Automatic Control*, 18(2):124–131, 1973.

[Jac77] D. Jacobson. On values and strategies for infinite-time linear quadratic games. *IEEE Transactions on Automatic Control*, 22(3):490–491, 1977.

[JSM19] Wouter Jongeneel, Tyler Summers, and Peyman Mohajerin Esfahani. Robust linear quadratic regulator: Exact tractable reformulation. In *2019 IEEE Conference on Decision and Control (CDC)*, 2019.

[Kak03] Sham Machandranath Kakade. *On the Sample Complexity of Reinforcement Learning*. PhD thesis. University College London, 2003.

[Kal60] R.E. Kalman. Contributions to the theory of optimal control. *Boletin de la Sociedad Matematica Mexicana*, pages 102–119, 1960.

[KP03] Steven Q. Krantz and Harold R. Parks. *The Implicit Function Theorem. History, Theory, and Applications*. Birkhauser, 2003.

[KPC93] M.M. Konstantinov, P.H. Petkov, and N.D. Cliristov. Perturbation analysis of the discrete riccati equation. *Kybernetika*, 29(1):18–29, 1993.

[KPZ90] 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*, 35(3):356–361, 1990.

[KR73] N.H. Kuiper and J.W. Robbin. Topological classification of linear endomorphisms. *Inventiones math. Springer Verlag*, 19:83–106, 1973.

[KS72] Huibert Kwakernaak and Raphael Sivan. *Linear Optimal Control Systems.* Wiley, 1972.

[Las09] Jean Bernard Lasserre. *Moments, Positive Polynomials and Their Applications.* Imperial College Press, 2009.

[LC13] Lennart Ljung and Tianshi Chen. What can regularization offer for estimation of dynamical systems. *IFAC Proceedings Volumes*, 46(11):1 – 8, 2013.

[LCM19] Lennart Ljung, Tianshi Chen, and Biqiang Mu. A shift in paradigm for system identification. *International Journal of Control*, 0:1 – 8, 2019.

[Lew92] F.L. Lewis. *Applied Optimal Control and Estimation.* Prentice-Hall, 1992.

[Lew02] Andrew D. Lewis. A brief on controllability of nonlinear systems. 2002.

[LR95] Peter Lancaster and Leiba Rodman. *Algebraic Riccati Equations.* Oxford Science Publications. Oxford University Press, 1995.

[Luc76] Robert E. Lucas. Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy*, 1:19 – 46, 1976.

[Lue69] David G. Luenberger. *Optimization by Vector Space Methods.* Wiley Professional Paperback Series. John Wiley & Sons, 1969.

[Mag76] E. Mageirou. Values and strategies for infinite time linear quadratic games. *IEEE Transactions on Automatic Control*, 21(4):547–550, 1976.

[Moo87] J.B. Moore. A universality advantage of stochastic excitation signals for adaptive control. *Systems & Control Letters*, 9(1):55 – 58, 1987.

[MR18] Eric Mazumdar and Lillian J. Ratliff. On the convergence of competitive, multi-agent gradient-based learning. 2018.

[MRJS19] Eric Mazumdar, Lillian J. Ratliff, Michael I. Jordan, and S. Shankar Sastry. Policy-gradient algorithms have no guarantees of convergence in continuous action and state multi-agent settings. 2019.

[MT19] Nikolai Matni and Stephen Tu. A Tutorial on Concentration Bounds for System Identification. *arXiv e-prints*, 2019.

[MTR19] Horia Mania, Stephen Tu, and Benjamin Recht. Certainty equivalence is efficient for linear quadratic control, 2019.

[Pet87] Ian R Petersen. A stabilization algorithm for a class of uncertain linear systems. *Systems & Control Letters*, 8(4):351 – 357, 1987.

[PH86] Ian R. Petersen and Christopher V. Hollot. A riccati equation approach to the stabilization of uncertain linear systems. *Automatica*, 22(4):397 – 411, 1986.

[Pol86a] J.W. Polderman. A note on the structure of two subsets of the parameter space in adaptive control problems. *Systems & Control Letters*, (7):25–34, 1986.

[Pol86b] J.W. Polderman. On the necessity of identifying the true parameter in adaptive lq control. *Systems & Control Letters.*, (8):87–91, 1986.

[Pol87] Jan Willem Polderman. *Adaptive Control & Identification: Conflict or Conflux.* Phd thesis. 1987.

[PR93] Svatopluk Poljak and Jiří Rohn. Checking robust nonsingularity is np-hard. *Mathematics of Control, Signals and Systems*, 6(1):1–9, 1993.

[Pri10] J. A. Primbs. Lqr and receding horizon approaches to multi-dimensional option hedging under transaction costs. In *Proceedings of the 2010 American Control Conference*, pages 6891–6896, 2010.

[PW98] Jan Willem Polderman and Jan C. Willems. *Introduction to mathematical systems theory*, volume 26 of *Texts in Applied Mathematics.* Springer-Verlag, New York, 1998. A behavioral approach.

[Rob72] J. W. Robbin. Topological conjugacy and structural stability for discrete dynamical systems. *Bull. Amer. Math. Soc.*, 78(6):923–952, 11 1972.

[RPUS00] Ian R. Petersen, V.A. Ugrinovskii, and Andrey Savkin. *Robust Control Design Using H^∞ Methods*. 2000.

[Sha86] U. Shaked. Guaranteed stability margins for the discrete-time linear quadratic optimal regulator. *IEEE Transactions on Automatic Control*, 31(2):162–165, 1986.

[Sho85] Clark R. Givens Rae Michael Shortt. A class of wasserstein metrics for probability distributions. *Michigan Math. J.*, 31:231–240, 1985.

[SMT+18] Max Simchowitz, Horia Mania, Stephen Tu, Michael I. Jordan, and Benjamin Recht. Learning without mixing: Towards a sharp analysis of linear system identification. In Sébastien Bubeck, Vianney Perchet, and Philippe Rigollet, editors, *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 439–473. PMLR, 06–09 Jul 2018.

[Sun98] Ji-Guang Sun. Perturbation theory for algebraic riccati equations. *SIAM J. MATRIX ANAL. APPL.*, 19(1):39–65, 1998.

[SW94] 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*, 39(3):686–691, 1994.

[SW97] H. J. Sussmann and J. C. Willems. 300 years of optimal control: from the brachystochrone to the maximum principle. *IEEE Control Systems Magazine*, 17(3):32–44, 1997.

[Tu19] Stephen Tu. *Sample Complexity Bounds for the Linear Quadratic Regulator*. PhD thesis, EECS Department, University of California, Berkeley, May 2019.

[US18] Jack Umenberger and Thomas B. Schön. Learning convex bounds for linear quadratic control policy synthesis. In *Neural Information Processing Systems 2018 :*, volume 31 of *Advances in Neural Information Processing Systems*, 2018.

[VF95] C. Verde and J. Flores. Nominal model selection for robust control design. In *Proceedings of 1995 American Control Conference - ACC'95*, volume 6, pages 4096–4100, 1995.

[vS94] J.H. van Schuppen. Tuning of gaussian stochastic control systems. *IEEE Transactions on Automatic Control*, 39(11):2178–2190, 1994.

[vS18] Jan H. van Schuppen. *Mathematical Control and System Theory of Discrete-Time Stochastic Systems*. 2018.

[VV07] Michel Verhaegen and Vincent Verdult. *Filtering and System Identification*. Cambridge University Press, 2007.

[Whi90] Peter Whittle. *Risk-sensitive Optimal Control*. Wiley, 1990.

[Yan18] Insoon Yang. Wasserstein Distributionally Robust Stochastic Control: A Data-Driven Approach. *arXiv e-prints*, 2018.

[YUP02] 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.*, volume 1, pages 604–609 vol.1, 2002.

[Zhu89] S.Q. Zhu. *Robustness of Feedback Stabilization: A Topological Approach*. PhD thesis, Technische Universiteit Einhoven. 1989.

[ZYB19] Kaiqing Zhang, Zhuoran Yang, and Tamer Başar. Policy Optimization Provably Converges to Nash Equilibria in Zero-Sum Linear Quadratic Games. *arXiv e-prints*, 2019.