Minimax Adaptive Control and Estimation

Kjellqvist, Olle

2024

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Kjellqvist, O. (2024). *Minimax Adaptive Control and Estimation*. [Doctoral Thesis (compilation), Department of Automatic Control]. Department of Automatic Control, Lund University.

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Minimax Adaptive Control and Estimation

OLLE KJELLQVIST
DEPARTMENT OF AUTOMATIC CONTROL | LUND UNIVERSITY

Printed by Media-Tryck, Lund 2024 NORDIC SWAN ECOLABEL 3041 0903

Department of Automatic Control
P.O. Box 118, 221 00 Lund, Sweden
www.control.lth.se

PhD Thesis TFRT-1145
ISBN 978-91-8104-167-5
ISSN 0280–5316

Minimax Adaptive Control and Estimation

Olle Kjellqvist

LUND UNIVERSITY

Department of Automatic Control

PhD Thesis TFRT-1145
ISBN 978-91-8104-167-5 (print)
ISBN 978-91-8104-168-2 (web)
ISSN 0280–5316

Department of Automatic Control
Lund University
Box 118
SE-221 00 LUND
Sweden

Printed in Sweden by Media-Tryck.
Lund 2024

Abstract

This thesis presents five papers on minimax adaptive control and estimation. Minimax adaptive estimation is a framework for output prediction and state estimation that provides a priori computable performance bounds for estimators. Minimax adaptive controllers ensure that the closed loop has finite gain, maintaining stability and performance under model class uncertainty.

The contributions of these papers are as follows: Paper I: Presents a minimax optimal output prediction algorithm for linear systems with parameter uncertainty. Paper II: Proposes an algorithm to compute performance bounds for minimax adaptive estimators. Paper III: Develops a minimax suboptimal adaptive controller for scalar linear systems with noisy measurements. Paper IV: Introduces a class of nonlinear systems for which minimax dual control admits a finite-dimensional sufficient statistic, builds dynamic programming theory around this class, and designs an adaptive controller for stabilizing an integrator from absolute-value measurements. Paper V: Provides a unified framework for state-feedback and output-feedback minimax adaptive control and methods for synthesizing suboptimal controllers. Complementing these theoretical contributions are two software artifacts: one for adaptive control and the other for adaptive estimation.

The contributions apply to simple systems that represent components of larger systems, marking a step towards automating controller synthesis and maintenance for critical infrastructures.

Acknowledgements

This thesis would not have been possible without emotional, technical, and financial support. I am privileged to call Ingrid Isacsson my wife, and I am grateful for her unwavering support and encouragement during this five-year research commitment. I am also grateful to my family for their support and encouragement.

My supervisors, Anders Rantzer, Bo Bernhardsson, and Venkatraman Renganathan, have provided invaluable guidance and feedback throughout this journey. I am grateful for their mentorship and the opportunities they have provided. My colleagues at the Department of Automatic Control have made this journey enjoyable and rewarding.

I would especially like to thank Eva for our extra morning coffee breaks and for being a link to reality. Mika, Monika, and Cecilia, thank you for making the department a welcoming place and for making administrative issues disappear like magic. Not much would have been possible without Anders & Anders, who keep the computers and labs running.

My fellow PhD students have been a source of inspiration and support, and I am grateful for the friendships we have formed. I am also indebted to my friends outside academia for their support and encouragement.

Richard Murray graciously hosted me at Caltech, where I had the opportunity to meet and collaborate with Jing Yu, Lisa Li, Anish Sarma, Ather Gattami, and John Doyle. I am grateful for their hospitality and the opportunity to work with them.

Thanks also to Ylva Wahlquist, Anders Rantzer, Julia Adlercreutz and Venkatraman Renganathan for reading and providing feedback on the manuscript, and to Venkatraman Renganathan, Bo Bernhardsson and Richard Pates for their feedback on several of the papers.

Large-Language models

The author acknowledges the use of Large Language Models (LLMs) for soft editing. These tools were employed to enhance the clarity, coherence, and grammatical accuracy of the text.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 834142 (ScalableControl).

Contents

1

Introduction

This thesis is a collection of five papers on reliable learning, estimation, and control of physical systems. The types of systems we consider are simple but are typically parts of larger, more complex systems, like the critical infrastructure systems on which society relies. They supply us with water, electricity, heat, transportation, communication, and so on. Their performance and reliability are crucial for the well-being of society. However, inefficient operations and maintenance of these systems risk wasting resources and increasing the risk of failures.

To operate and maintain these systems, we employ sensors to monitor their state and actuators to control them based on the sensor measurements. That is, we use feedback loops to ensure that the system behaves as desired. Control theory provides methods to design and analyze such feedback loops. It has been instrumental in improving efficiency and reliability across different applications in the process industry, aviation, automotive industry, and robotics, to name a few.

However, accessing the most potent control methods requires a deep understanding of the system and accurate models. Obtaining and maintaining such models can be challenging and requires significant engineering effort. Hence, obtaining accurate models for modern infrastructure systems is often infeasible due to their scale and complexity, preventing control engineers from harnessing the true potential of control theory.

There is a growing interest in automating the design and analysis of control systems. Such automatic methods alleviate the need for manual model-based design and are a practical approach to larger systems. However, automatic methods trade the manual engineering effort for additional complexity in the control system. This complexity can be challenging to analyze and understand, leading to unexpected behaviors and system failures. For example, caution is not always the safe approach, as illustrated by the teddy bears in Figure 1.1.

Figure 1.1 A figure of teddy bears illustrating some difficulties of simultaneous learning and control. The left teddy bear is observing the pot and will only note a change in the potato when it starts to dissolve. The middle and right teddy bears take an active approach and will undoubtedly gain more information than the left teddy bear. However, the right teddy bear prioritizes the experiment over the result. Although the hammer is a powerful tool, and smashing the potatoes will reveal their consistency, the dish is ruined. The middle teddy bear is more careful and uses a knife to test one potato, revealing enough information to decide whether to remove the potatoes from the heat without negatively impacting the result. This thesis aims to be the middle teddy bear. It contains control methods that are bold enough to extract the necessary information yet careful enough to avoid compromising the end goal.

2

Background

All reasonable methods for designing control systems rely on some information about the system. This information can be a model, a set of measurements, or expert rules. Some methods include an explicit experiment to gather information about the system. However, we still require some initial knowledge, like whether it has an inverse response, the sign of the gain, signal ranges, and so forth, to apply the method.

Our prior knowledge about the system is imperfect, as with all engineering practices. This imperfection comes from many sources, like simplifying assumptions, non-exhaustive experiments, changing conditions, and so on. Feedback, as illustrated in Figure 2.1, is used because of its remarkable ability to deal with these imperfections. However, feedback is not a panacea and, if not appropriately designed, may introduce unwanted behaviors, like oscillations, instability, or slow responses.

If we have an accurate model of the system, using this model as a proxy for the actual system in a feedback loop typically results in a well-behaved

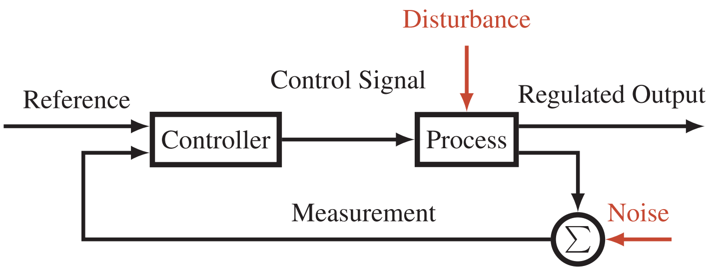


Figure 2.1 Block diagram of a feedback system. The controller generates a control signal and feeds it to the process. The process is the system that we want to control. It is affected by our control signal and exogenous disturbances. The regulated output is a system quantity that we wish to regulate and may or may not be directly measurable. The measurement is a typically noisy sensor reading of a system quantity, which we use to calculate our control signal.

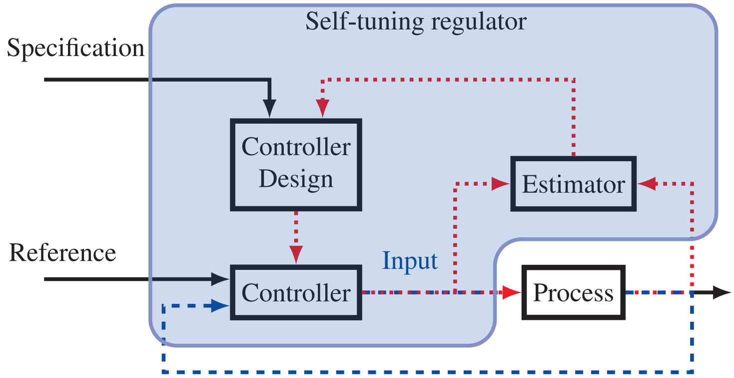


Figure 2.2 Block-diagram of a self-tuning regulator, recreated from [Åström and Wittenmark, 2008, Figure 1.19]. The system consists of two feedback loops, one for control signal computation (dashed, blue) and one for the parameter estimator (dotted, red). The shaded area represents the "controller" portion of the feedback loop. The estimator and control loop may run at different rates; in the extreme case, the analog controller may run continuously, while the estimator and the controller design procedure may rely on sampling and run in discrete time.

system. Unfortunately, as argued in the Introduction, we often do not have accurate models. Two umbrella terms cover the methods that explicitly deal with the model uncertainty: *Adaptive Control* and *Robust Control*.

Adaptive control concerns methods that adjust the controller based on the system's response. The controller is typically parameterized, and an additional feedback loop is used to adjust the parameters, like the self-tuning regulator in Figure 2.2. Although adaptive control methods can be compelling, the added complexity can make them challenging to analyze and understand. Both successful applications and disastrous failures have been reported [Anderson and Dehghani, 2008], and there is a need for methods that come with rigorous guarantees [Matni et al., 2019; Alleyne et al., 2023].

Robust control quantifies the initial uncertainty and poses controller design as a robust optimization problem. The uncertainty is often represented as a set of models, and the controller is designed to work well for all models in the set. As illustrated in Figure 2.3, the controller is designed to work *reasonably* well in reality, provided that model-to-reality mismatch belongs to the uncertainty set. The controller structure is typically fixed and simpler, most often linear, than in adaptive control. However, the methods have strong guarantees, like stability and performance bounds for the closed-loop system. Unfortunately, the simplicity of the controller structure comes at the cost of a lower performance ceiling compared to adaptive control methods.

In this thesis, we will build on the robust control framework but relax the assumption that the model and the controller are linear. This relaxation

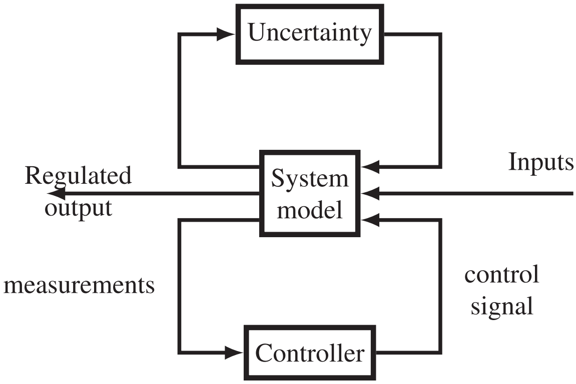


Figure 2.3 Representation of real systems as a nominal model in feedback with a controller and an uncertainty block typical of robust control. This representation is quite general. While the model and controllers are typically linear in the robust control literature, we will relax these assumptions. Usually, robust control methods come with guarantees of the following form: "If the uncertainty is within a certain set, then the closed-loop system is finite gain. " Finite gain means the closed-loop system is asymptotically stable and attenuates exogenous disturbances.

allows us to design and analyze adaptive controllers from a robust control perspective. The chapter proceeds with the story of uncertainty in control systems via a simple system, a tank process, in Section 2.1, followed by a discussion on robust control and its limitations in Section 2.2. Section 2.3 argues that the minimax adaptive control framework is a natural extension with the potential to overcome these limitations, leading up to our research questions in Section 2.4.

2.1 Tanks filled with uncertainty

In the introductory control course at Lund University, students experiment with a tank process, illustrated in Figure 2.4. We will use eight supposedly identical tank processes to discuss different aspects of uncertainty in control systems.

The physical system consists of a cylindrical tank with a hole in the bottom, a basin, a pump that pumps water from the basin back into the tank, a pressure transducer, and an I/O-box with AD/DA conversion connected to a computer.

The students are tasked with designing feedback controllers to regulate the tank's water level. Teaching assistants challenge informal robustness specifications by pouring additional water into the tanks, occasionally squeezing the tubes connected to the pressure sensor, and shaking the table.

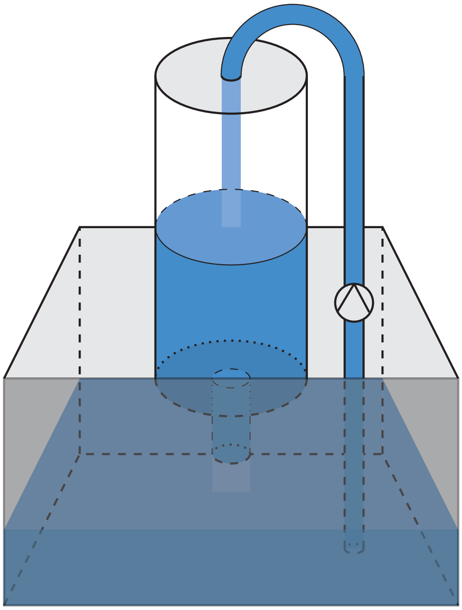


Figure 2.4 Drawing of the tank process. The pump pumps water from the basin back into the tank, and the pressure transducer measures the water level. The outflow depends on the water level and the hole's area. The pump's efficiency and the hole's area vary between tanks due to manufacturing tolerances and wear and tear, leading to uncertainty in the system.

A process model is straightforward to derive using *Toricelli's law*, relating the outflow speed v to the water level h of the tank and mass balance. The law states that the speed v equals that of a free body falling from h, i.e., $v = \sqrt{2gh}$.

Denote the inflow by Q_{in}, the outflow by Q_{out}, then conservation of mass means that their difference describes the volume's change, i.e.,

$$\dot{V} = Q_{\text{in}} - Q_{\text{out}}, \tag{2.1}$$

where V denotes that tank's volume.

Denoting the cylinder's cross-sectional area by A and the outflow by a, we get $V = Ah$ and $Q_{\text{out}} = av$. We linearize the pump with a fast, compared to the tank's dynamics, PI controller, so $Q_{\text{in}} \approx ku$, where u is the control signal of the (outer) tank controller. Substituting the above expressions and Toricelli's law into (2.1) results in

$$\dot{h} = -\frac{a}{A}\sqrt{2gh} + \frac{k}{A}u. \tag{2.2}$$

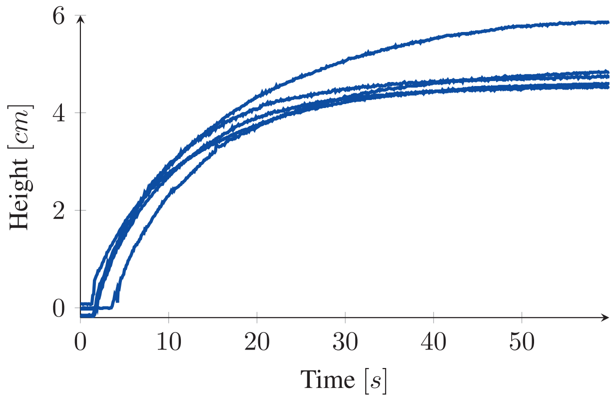


Figure 2.5 Step response of the tanks.

The control signal is limited to $0 \leq u \leq 10\,\mathrm{V}$ and $0 \leq h \leq 14\,\mathrm{cm}$. We will examine this process around the stationary point $u_0 = 5\,\mathrm{V}$, with

$$0 = -\frac{a}{A}\sqrt{2gh_0} + \frac{k}{A}u_0 \implies h_0 = \left(\frac{ku_0}{a}\right)^2 \frac{1}{2g}.$$

We would proceed with nominal values for the parameters and linearize around (u_0, h_0), arriving at a first-order model

$$\frac{\mathrm{d}\Delta h}{\mathrm{d}t} = a_0 \Delta h + a_0 k_0 \Delta u. \tag{2.3}$$

We will now investigate the tanks experimentally.

Experimental Evaluation

This section illustrates some differences between the tanks and discrepancies to the physics-inspired model above. We perform two experiments; we first apply $u(t) = 5$ for $t \geq 0$ and highlight differences in the stationary point. The second experiment is a chirp signal around $u_0 = 5$: a linear frequency sweep from 0.01Hz to 10 Hz with amplitude 3. The chirp experiment is repeated after one hour of continuous operation to illustrate some time-varying effects.

Figure 2.5 shows the initial response of the tanks. The stationary points vary between $4.4\,\mathrm{cm} \leq h_0 \leq 6.4\,\mathrm{cm}$, revealing a significant difference between the tanks. Figures 2.6 and 2.7 display Bode plots of the first and second chirp experiments. When compared to a first-order time-delay model (red), our empirical responses have a slightly higher degree of roll-off and are not as smooth as the first-order model. The pole is around 0.07 rad/s, the gain around 2 and the time-delay is around 2 s. There is also significant uncertainty in the low and high-frequency spectra inherent to the estimation method not shown here.

	Stationary Point h_0							
Experiment \ Tank	1	2	3	4	5	6	7	8
Cold	4.6	5.1	5.1	5.9	6.3	6.9	5.0	4.8
Hot	4.6	4.8	4.5	5.7	5.8	6.4	4.7	4.4

Table 2.1 Stationary points of the tanks, the cold-start experiment is performed directly after starting the pump, and the hot-start experiment is performed after one hour of continuous operation. Tank 1 was used for prototyping during the initial setup, and its cold start response was thus compromised.

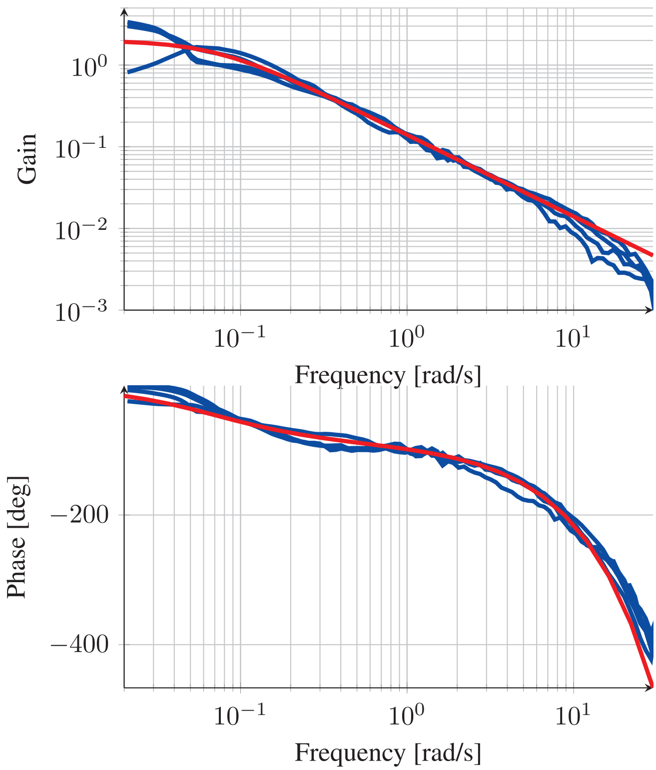


Figure 2.6 Empirical frequency response of the tanks, blue, from the cold-start chirp. A first-order time-delay model, red, is fitted to the data.

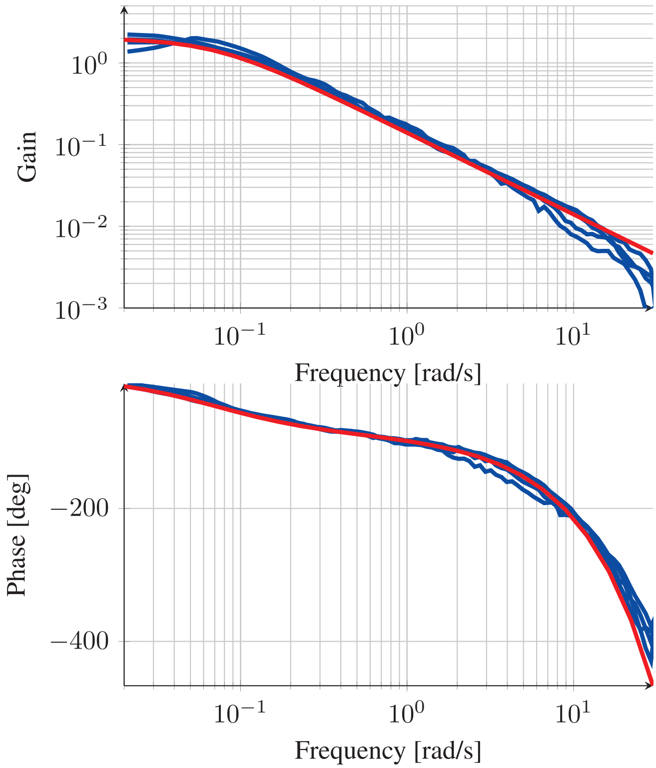


Figure 2.7 Empirical frequency response of the tanks, blue, from the hot-start chirp. A first-order time-delay model, red, is fitted to the data.

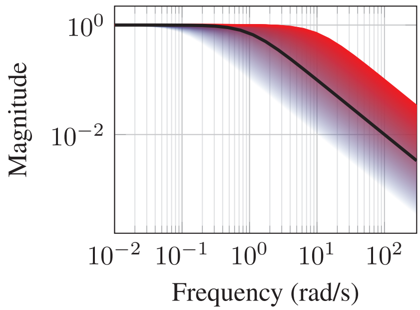
Uncertainty classes

When we talk about uncertainty, we typically discuss its origin, its effects, how to model it, and how to mitigate its effects. The discussion is complicated because even simple systems, such as tanks, are affected by uncertainty from many sources with different characteristics. Table 2.2 lists some common classes of uncertainty, their effects, and their origins.

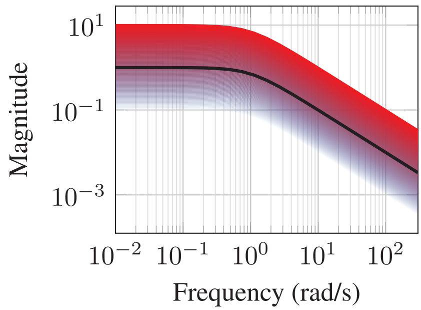
Consider the tanks' stationary points, h_0, in Table 2.1. The stationary points depend on the cross-sectional areas of the tank, the hole, and the pump's efficiency. Slight variations in the tank's area do not affect the stationary points much, but the hole's area and the pump's efficiency are more sensitive to perturbations. Comparing the cold and hot start experiments, we may hypothesize that the continuous water flow cleans the hole, increasing its area and, therefore, the outflow. The pump's efficiency may also change due to operating conditions. Performing the experiments, one can hear different pitches from the pumps, indicating that they are not operating at the same

Table 2.2 Summary of uncertainty properties in the tank process (and control systems in general).

Model Uncertainty	Affect	Origin
Parameters	Poles and Zeros, Stationary points, low frequencies	Tolerances, wear & tear operating conditions
Additive Noise	Measurements Stationary points	Sensor noise, load disturbances
Dynamic Uncertainty	High frequencies	unmodeled dynamics, time-varying effects, simplifications



(a) Pole, $a_0 \in \{0.01, 10\}$



(b) Gain, $k_0 \in \{0.01, 10\}$

Figure 2.8 Effects of parameter variations on the first-order model. The pole, a_0, affects the bandwidth, while the gain, k_0, affects the gain. The case $a_0, k_0 = 1$ is highlighted in black.

speed. This pump speed variation likely affects the pump efficiency constant k in (2.2), leading to differences in the stationary points.

For the behavior around the operating points, the parameter variations translate to uncertainty about a_0 and k_0 in (2.3). By varying a_0, we affect the bandwidth of the system, illustrated in Figure 2.8a. By varying k_0, we affect the gain, illustrated in Figure 2.8b. So, given the ability to experiment with each tank, we fit the parameters of the first-order model to the empirical data for one tank by adjusting a_0 and k_0, illustrated in Figure 2.9. The first-order model captures the low-frequency behavior well, but there are discrepancies in the high-frequency behavior. This discrepancy is likely due to unmodeled phenomena, the pump linearization, sampling, measurement noise, et cetera.

Toward quantifying the remaining uncertainty, we add a multiplicative

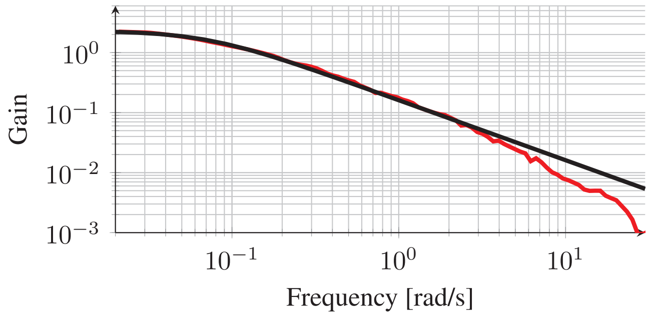


Figure 2.9 ...

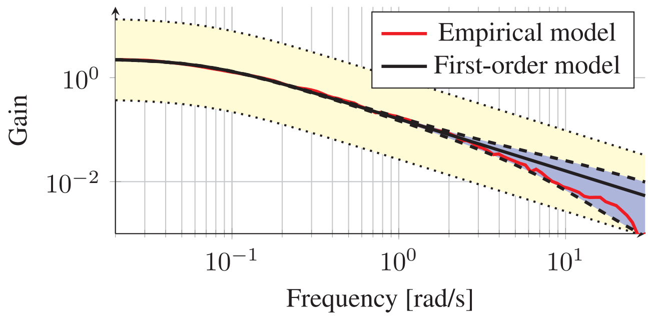


Figure 2.10 Uncertain first-order model with multiplicative uncertainty. The light-yellow area indicates the feasible realizations bounded by $\|\Delta\|_\infty \leq 1$ with constant weighting $W(s) = 6$. The light-blue area indicates the feasible realizations with frequency-dependent weighting $W(s) = \frac{0.9s}{s+12}$.

block, $\Delta(s)$,

$$P_\Delta(s) = (I + W(s)\Delta(s))P, \tag{2.4}$$

where W is a frequency-dependent weight, and Δ is a norm-bounded uncertainty; $\|\Delta\|_\infty \leq 1$. The shaded areas in Figure 2.10 indicate the feasible realizations of P_Δ , and the black line indicates the nominal model. The light-yellow area indicates the feasible realizations with constant weighting $W(s) = 6$, and the light-blue area indicates the feasible realizations with frequency-dependent weighting $W(s) = \frac{0.9s}{s+12}$. Both weights are chosen to include the high-frequency behavior of the empirical data. The constant weight, $W(s) = 6$, describes an uncertainty set too large for meaningful controller design. The frequency-dependent weight, $W(s) = \frac{0.9s}{s+12}$, describes a smaller uncertainty set that works well with robust control synthesis methods.

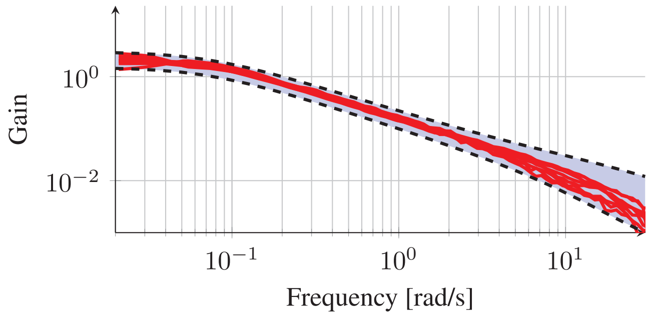


Figure 2.11 The eight responses of the tanks, red, and an uncertain first-order model with parametric and frequency-dependent uncertainty, shaded area.

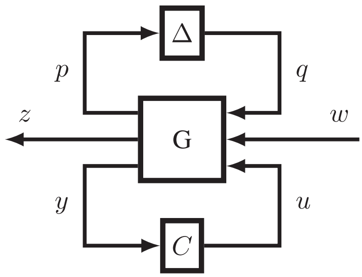


Figure 2.12 Generalized plant for robust control synthesis. The controller C is designed to work well for all possible uncertainty realizations Δ.

Conclusions

The tanks are simple systems, analyzed in an ideal setting, a laboratory, yet they illustrate many aspects of uncertainty in control systems. The experiments indicate the presence of both low and high-frequency uncertainty. Parametric uncertainty models low-frequency uncertainty well, while high-frequency uncertainty is due to higher-order dynamics and measurement noise. In particular, the tanks illustrate that high-frequency uncertainty remains even after carefully fitting the model parameters. This high-frequency uncertainty can be modeled as a norm-bounded, otherwise arbitrary, uncertainty, but the robust control design will be conservative unless paired with a suitable weighting.

2.2 Robust control

Robust control is a subfield of control theory that aims to analyze and design controllers for uncertain systems, like in Figure 2.11. The abstraction, Figure 2.12, captures a variety of uncertainty models and is the starting point for many robust control methods. The block diagram contains a *generalized plant* G that captures the interaction between the linear controller C, the nominal plant (P), and the uncertainty block Δ For example, consider multiplicative weighted uncertainty as in (2.4), assuming that the controller can access noisy measurements. Let

$$y = z + v = (I + W(s)\Delta(s))P(s)(u + d) + v$$

where u is the control signal, d are load disturbances, v is measurement noise, z the regulated output and y the measurements, then

$$G = \left[\begin{array}{c|c|c} G_{pq} & G_{pw} & G_{pu} \\ \hline G_{zq} & G_{zw} & G_{zu} \\ \hline G_{yq} & G_{yw} & G_{yu} \end{array}\right] = \left[\begin{array}{c|cc|c} 0 & P & 0 & P \\ \hline W(s) & P & 0 & P \\ \hline W(s) & P & 1 & P \end{array}\right]. \tag{2.5}$$

Here, G_{ab} denotes the open-loop transfer function from b to a.

For more complex uncertainty models, like the one illustrated in Figure 2.11, one proceeds similarly by laying out the inputs and outputs of each source of uncertainty. These inputs are concatenated into p and their outputs into q. This procedure is often called “pulling out the deltas” and always admits generalized plants such that Δ is block diagonal, and by absorbing the weights into the plant, we may assume $\|\Delta\| < 1$.

Unfortunately, the associated synthesis problems are computationally challenging. One typically minimizes a biconvex upper bound in two variables: the controller C (sometimes called K in the literature) and the scaling operator, often called D. For dynamic uncertainty, the relaxation takes the form

$$\min_{D \in \mathcal{D}} \|D^{-1} T_{pq}(C) D\|_\infty,$$

where $T_{pq}(C)$ is the closed-loop transfer function $q \mapsto p$ given controller C and $\mathcal{D}$ is the set of norm bounded operators that commute with Δ, i.e.,

$$\mathcal{D} = \{D \in \mathcal{H}_\infty : D \text{ is invertible, } D\Delta = \Delta D \text{ for all } \Delta\}.$$

An operator D is in $\mathcal{H}_\infty$ if its frequency response is a bounded holomorphic function outside the unit disc. For linear analysis preliminaries, see [Dullerud and Paganini, 2000, Chapter 3] or [Zhou and Doyle, 1998, Chatper 4]. Controller synthesis concerning this upper bound then becomes a joint minimization problem. Unfortunately, the minimization problem is not jointly convex in C and D but convex in C for fixed D and convex in D for fixed C. This coordinate-wise convexity has inspired the use of coordinate-wise descent, alternating between C and D, often called $D - K$ iteration.

The upper bound for mixed uncertainty is significantly more involved and conservative, leading to more challenging synthesis problems. The relationship between uncertainty models and analysis and synthesis methods is unsettling: the more we restrict the uncertainty, the more complex the computations and the more conservative the results become. We will next discuss the limitations of linear time-invariant controllers.

Controller classes and uncertainty types

It is well known that linear controllers achieve optimal robustness margins for unstructured uncertainty, where Δ is a full block and possibly nonlinear or time-varying dynamic uncertainty. However, when the uncertainty is parametric, nonlinear and time-varying controllers can do much better. Consider the systems

$$P(s) = \frac{B(s)}{A(s)} = \frac{b}{s}, \text{ where } b = \pm 1.$$

There is no linear time-invariant controller, $C(s) = \frac{R(s)}{Q(s)}$ that stabilizes the system for both b. To see this, examine the characteristic polynomials of the closed-loop systems

$$A(s)Q(s) + B(s)R(s) = q_n s^{n+1} + \ldots + q_0 s + b\left(r_n s^n + \ldots + r_0\right).$$

The systems are asymptotically stable only if all coefficients have the same sign, but this is impossible for both $b = \pm 1$.

Instead, consider sampling the system at one Hertz with zero-order hold. The sampled system is described by the difference equation

$$x_{k+1} = \phi x_k + b\gamma u_k,$$

for some constants ϕ and γ. The controller

$$u_k = (-1)^k \frac{\phi}{\gamma} x_k,$$

stabilizes the system for both $b = \pm 1$, illustrating that time-varying controllers can stabilize uncertain systems that linear time-invariant controllers cannot.

Conclusions

Robust control is a powerful framework with methods to model and mitigate uncertainty in control systems. A significant advantage is the ability to explicitly model how uncertainty affects the system and treat different types of uncertainty differently. The synthesis methods require output feedback controllers, since measuring the signals before and after the uncertainty block is typically impossible, leading to partial and corrupted measurements. However, if the uncertainty is parametric, the analysis and synthesis problems are computationally demanding and often conservative. Further, the restriction to linear time-invariant controllers significantly limits the achievable performance under parametric uncertainty.

Historical Notes

Robust control has roots in early work on input/output descriptions by Zames and Sandberg [Zames, 1966], the Lur'e problem of absolute stability [Liberzon, 2001] and Popov's hyperstability [Popov and Georgescu, 1973]. Much of the historical development of the theory is nicely described in the historical accounts [Safonov, 2012; Dorato, 1987] and the textbooks [Francis, 1987; Zhou and Doyle, 1998; Dullerud and Paganini, 2000].

George Zames initiated the $\mathcal{H}_\infty$-control problem when he argued for minimizing the induced norm of the weighted sensitivity function (formulated for Banach algebras but specialized to the induced $\mathcal{L}_2$-gain setting.) [Zames, 1981]. The primary motivation is that the $\mathcal{H}_\infty$-norm is an induced norm whose sub-multiplicative property can be used to guarantee stability for plants that deviate from a nominal model. $\mathcal{H}_\infty$ control theory has gone through three major stages: the early frequency-domain (functional analysis) approach [Francis, 1987; Feintuch, 1998], the Riccati equations approach [Zhou and Doyle, 1998] and the linear-matrix inequality approach [Dullerud and Paganini, 2000]. The Riccati equation approach strongly connects to the theory of dynamic differential games [Basar and Bernhard, 2008; Tadmor, 1993]. Game theoretic and passivity-based approaches can be extended to the nonlinear setting [James, 1995]. In the nonlinear setting, partial differential equations replace the Riccati equations.

Critics of $\mathcal{H}_\infty$-control claim it is overly conservative. One reason for conservativeness is that naive applications discard any structural or topological information about the nature of the perturbations entering the system. To remedy this conservativeness [Doyle, 1982] introduced the frequency-dependent *structured singular value* (μ) to analyze robustness against structured perturbations, and [Doyle et al., 1982] extended μ to robust performance. These results generalize earlier work [Safonov, 1978; Safonov, 1981]. In the mid-'80s, researchers were concerned with computing upper and lower bounds of μ for structured uncertainty where the perturbations are linear time-invariant systems. [Fan and Tits, 1986] reformulated the problem as a smooth, non-convex optimization problem. This reformulation is amenable to gradient-based optimization methods and always returns the correct value when the upper bound is tight. A power method for computing lower bounds was introduced in [Packard et al., 1988], and the case of robustness against static, mixed real, and complex uncertainties was considered in [Fan et al., 1988; Fan et al., 1991] with power methods for lower bounds in [Young and Doyle, 1990; Young et al., 1992] [Shamma, 1994] showed that the upper bound with constant D-scales is necessary and sufficient for LTV perturbations, and [Poola and Tikku, 1995] showed that the upper bound with frequency-weighted D-scales is necessary and sufficient for "arbitrarily slowly time-varying structured linear perturbations". Structured robustness can be further generalized and studied in the framework of integral-quadratic constraints [Megretski and Rantzer, 1997], which extends to nonlinear systems.

2.3 Minimax adaptive control

The $\mathcal{H}_\infty$-norm minimization problem of robust control can be approached by finding minimax equilibria in an associated two-player zero-sum game, sometimes called a *minimax control* problem. In the standard setup, the controller selects the control signal to minimize the performance measure, while the adversary selects the uncertain inputs to maximize the performance measure. In minimax adaptive control, the adversary is also empowered to select the parameters of the plant, and the resulting controller is adaptive—with bounded ℓ_2-gain for all possible parameter realizations. The rest of this section will provide a brief overview of minimax adaptive control.

Finite uncertainty sets

[Vinnicombe, 2004] studied the achievable closed-loop induced ℓ_2-gain for scalar systems,

$$x_{t+1} = ax_t + bu_t + w_t,$$

where $x_t \in \mathbb{R}, u_t \in \mathbb{R}$ and $w_t \in \mathbb{R}$ in two settings. In both settings, $a \in \mathbb{R}$ and $b \in \mathbb{R}$ are fixed. In the first setting, $a \in \{-a_0, a_0\}$ is unknown but $b = 1$ is known. In the second setting, a is known, but $b \in \{-1, 1\}$.

Vinnicombe found that a certainty-equivalence deadbeat controller achieves an induced ℓ_2 gain from $w \to x$ of $a + \sqrt{1 + a^2}$ and that, to within numerical accuracy, no controller can do better for the *second* problem. For the first problem, this is demonstratedly suboptimal. The parameter estimate for certainty equivalence is the least squares estimate of a and b, respectively.

[Rantzer, 2021] extends the previous results to uncertain linear systems of the form

$$x_{t+1} = Ax_t + Bu_t + w_t \qquad t \geq 0 \; u_t = \mu_t(x_0, \ldots, x_t, u_0, \ldots, u_{t-1}).$$

Here, (A, B) are unknown but belong to a finite set. The solution is based on a similar reformulation, using the current state x_t and sum-of-squares matrix,

$$Z_{t+1} = Z_t + \begin{bmatrix} v_t \\ x_t \\ u_t \end{bmatrix} \begin{bmatrix} v_t \\ x_t \\ u_t \end{bmatrix}^\mathsf{T}, \quad Z_0 = 0.$$

The proposed formal statement for an explicit expression of an adaptive controller satisfying a pre-specified bound on the induced ℓ_2 gain for finite model sets is the following:

PROPOSITION 1
Given $A_1, \ldots, A_N \in \mathbb{R}^{n_x \times n_x}$, $B_1, \ldots, B_N \in \mathbb{R}^{n_x \times n_u}$ *and positive definite* $Q \in \mathbb{R}^{n_x \times n_x}$, $R \in \mathbb{R}^{n_u \times n_u}$, *supposed there exists* $K_1, \ldots, K_N \in \mathbb{R}^{n_u \times n_x}$ *and* $P_{ij} \in \mathbb{R}^{n_x \times n_x}$ *with* $0 \prec P_{ij} \prec \gamma^2 I$ *and*

$$|x|^2_{P_{ik}} \geq |x|^2_Q + |K_k x|^2_R + |(A_i - B_i K_k + A_j - B_j K_k)x/2|^2_{(P_{ij}^{-1} - \gamma^2 I)^{-1}} - \gamma^2 |(A_i - B_i K_k - A_j + B_j K_k)x/2|^2 \quad (2.6)$$

for $x \in \mathbb{R}^{n_x}$ *and* $(i, j, k) \in \{1, \ldots, N\}$. *Then* $J_\mu(x_0) \leq \max_{ij} \left\{ |x_0|^2_{P_{ij}} \right\}$ *is valid for the control law* μ *defined by*

$$u_t = -K_{k_t} x_t,$$
$$k_t = \arg\min \sum_{\tau=0}^{t-1} |A_i x_\tau + B_i u_\tau - x_{\tau+1}|$$

Inequality (2.6) is generally not simultaneously convex in P_{ij} and K_k, so simultaneous synthesis is challenging. Rantzer remarked that one might take K_k as the γ-suboptimal $\mathcal{H}_\infty$ state-feedback gain associated with the model (A_k, B_k). [Cederberg et al., 2022] proposed a linearization technique around nominal gains K_k^0 and demonstrated significant improvement (lowering γ). Initial results show that the inequality admits a solution for pairs of systems that are individually stabilizable with a common quadratic Lyapunov function, [Bencherki and Rantzer, 2023].

Conclusions

The minimax adaptive control framework treats both low and high-frequency uncertainty: low-frequency uncertainty is treated via adaptation, and high-frequency uncertainty is treated via robustness. Rantzer's piecewise-quadratic approximation of the value functions allows for explicit controller synthesis for finite model sets, and the results are promising. The conditions under which (2.6) admits a solution still need to be fully understood. Output feedback extensions are so far unexplored.

Historical notes

Minimax control originated in the '60s [Witsenhausen, 1966]. We refer the reader to [Basar and Bernhard, 2008, Chapter 1.1] for the historical development up to the mid-'90s and connections to game theory and statistical decision theory. Nonlinear state-feedback analysis and control design was considered in the early '90s, see [Schaft, 1992] and the references therein. The generalization to the output-feedback problem is considerably more involved. Sufficient and necessary conditions were derived in [James and Baras, 1995], who brought in the concept of an information state (a sufficient statistic [Striebel, 1965]) and argued that it is typically infinite-dimensional. Although the optimal controller is computable, in principle, by solving nonlinear functional equations, there are only a few known practical cases where the information state has a finite-dimensional representation, and computations become tractable [James and Yuliar, 1995].

The term "minimax adaptive control" was introduced in [Didinsky and Basar, 1994]. The authors consider continuous-time dynamics that are linear in the unknown constant parameters but possibly nonlinear in the state and control signals, as well as a soft-constrained performance measure that is quadratic in the disturbances and unknown parameters but non-quadratic otherwise. The authors obtain a set of necessary and sufficient conditions and show that the minimax controller, if it exists, is a function of the least-squares estimate of the unknown parameters. The results were generalized to nonlinear SISO systems on "parametric strict-feedback form" in [Pan and Basar, 1998]. First-order discrete-time linear systems with sign-uncertainty were considered in [Vinnicombe, 2004] and [Megretski and Rantzer, 2003] provides a lower bound on the achievable ℓ_2-gain for first-order discrete-time linear systems where the uncertain pole belongs to an interval.

2.4 Research topics

This thesis extends the minimax adaptive control framework to partial and imperfect measurements, aiming for a reliable and practical framework for adaptive control in mind. In particular we study

RT 1 Estimation and prediction for uncertain linear systems with minimax objectives.

RT 2 Output feedback where the preimage of a measurement under the measurement function is a finite set.

RT 3 Output feedback for finite sets of linear systems, where the controller can only access noisy linear combinations of the state.

The first point is a natural starting point for extending the minimax adaptive control framework to partial and imperfect measurements. The second point considers systems where the uncertainty set may change over time and is constructed after each measurement $y_t = h(x_t)$ as $h^{-1}\{y_t\}$. The nature of the variation may be due to time-varying parameters, nonlinear phenomena, or exogenous signals. The last point means that the synthesis method can be applied to the μ-synthesis framework, allowing for much larger parameter uncertainty sets than previously possible, keeping the ability to incorporate frequency-dependent weighting functions.

3

Contributions

This chapter discusses the theoretical contributions of the five included papers and the author's specific contributions. The chapter ends with a short discussion of five excluded peer-reviewed papers.

3.1 Included papers

The first two papers concern estimation and prediction, research topic *RT 1* in Chapter 2. The insights from these papers, especially those about forward dynamic programming and its connection to observer design, are leveraged to tackle non-injective measurement functions in Paper IV (*RT 2*), and noisy measurements (*RT 3*) in Papers III and V. We proceed to discuss each paper in turn.

Paper I

O. Kjellqvist and A. Rantzer (2022c). "Minimax adaptive estimation for finite sets of linear systems". In: *Proc. 2022 IEEE Amer. Control Conf.* Pp. 260–265. DOI: 10.23919/ACC53348.2022.9867474, © 2022 IEEE. Reprinted, with permission.

Theoretical contributions This paper addresses output prediction in linear dynamical systems with uncertain dynamics, where the uncertainty belongs to a finite set. We provide a convex program that computes an estimate of the output at the next time step, ensuring that a constant, γ, bounds the gain from unmeasured disturbances to the output prediction error, provided that γ is a feasible gain bound. Additionally, we show how to evaluate online whether γ_N is a feasible gain bound.

Software contributions The paper includes a Julia implementation[1] of the algorithms, which is used to verify the theoretical results.

[1] https://github.com/kjellqvist/MinimaxEstimation.jl

Credits The paper is a joint work with Anders Rantzer, who suggested the initial problem formulation, provided guidance throughout the project, and pointed to [Basar and Bernhard, 2008] as a source of inspiration. The author implemented the algorithms, derived the theoretical results, and wrote the first draft of the paper. The author and Anders Rantzer jointly edited the paper, and the author revised it.

Paper II

O. Kjellqvist (2024c). "Minimax performance limits for multiple-model estimation". In: *Proc. 2024 Eur. Control Conf.* Pp. 2540–2546. DOI: 10.23919/ECC64448.2024.10590947, © 2024 IEEE. Reprinted, with permission.

This paper was presented at ECC2024 in Stochkolm, Sweden, and the abridged form is to be published in the conference proceedings. This thesis includes proofs that were previously omitted and are available on Arxiv:

O. Kjellqvist (2024d). *Minimax performance limits for multiple-model estimation.* arXiv: 2312.05159 [math.OC]. URL: https://arxiv.org/abs/2312.05159

Theoretical contributions This paper extends Paper I in two ways:

1. It considers strictly causal state estimation.
2. It provides upper and lower bounds on the *achievable* disturbance gains γ_N (from disturbances to Nth time-step estimation error).

The achievable attenuation level indicates whether the set of feasible models is suitable for state estimation, independent of the estimation procedure. The gain-bound can be interpreted as a finite-time performance guarantee.

By computing the upper and lower bounds for γ_N for indistinguishable and distinguishable systems, we argue that distinguishability [Silvestre et al., 2021] alone is neither sufficient nor necessary to guarantee that the estimation error will be small.

In practical multiple-model settings, such as in fault detection scenarios, one is often interested in detecting which model generated the data. The work in Papers I and II is limited to the state estimate and does not consider the ability to identify the underlying model.

Credits The author proposed the problem formulation, derived the theoretical results, implemented the algorithms to verify the theoretical results, validated all algebraic expressions numerically, wrote the first draft of the paper and edited and revised it. The author is thankful to Anders Rantzer for supervising the project and providing valuable feedback, Venkatraman Renganathan for reviewing an early draft, Anders Helmerson for related discussions and reviewing the first submission during the Licentiate seminar, and the anonymous reviewers for their constructive feedback.

Paper III

O. Kjellqvist and A. Rantzer (June 2022b). “Learning-enabled robust control with noisy measurements”. In: R. Firoozi et al. (Eds.). *Proc. 4th Annu. Learning Dyn. Control Conf.* Vol. 168. PMLR, pp. 86–96. URL: https://proceedings.mlr.press/v168/kjellqvist22a.html

This paper was presented at L4DC 2022 in Stanford and was published in an abridged form. The version contained in this thesis is an extended version including proofs that were previously omitted and is published on Arxiv:

O. Kjellqvist and A. Rantzer (2022a). *Learning-enabled robust control with noisy measurements.* DOI: 10.48550/ARXIV.2202.08363.

Theoretical contributions The main contribution of this paper is the equivalence between the following two statements for uncertain scalar linear dynamical systems where the uncertainty belongs to a finite set.

1. There exists a causal *output* feedback controller that achieves a closed-loop ℓ_2-gain bound of at most γ from disturbances to errors.

2. There exists a memoryless function of an $\mathcal{H}_\infty$ *multi observer*, such that certain performance quantities are bounded.

The *multi observer* consists of one $\mathcal{H}_\infty$ observer per feasible model, coupled with a performance quantity related to how well the model explains the observed data. The performance quantities can be evaluated recursively using observed signals and the observer states. We use this result to extend [Vinnicombe, 2004] to the output feedback setting, constructing suboptimal controllers for integrators where the gain's sign is unknown. The controllers are of the certainty-equivalence type, which coincides with a multiple-model adaptive (supervisory) control architecture.

Credits The paper is a joint work with Anders Rantzer, who challenged the author to extend [Vinnicombe, 2004]'s results to the output feedback setting. The author derived the theoretical results (initially with Vinnicombe's variable transformation), implemented the algorithms to verify the theoretical results, and wrote the first draft of the paper. Anders Rantzer pointed out that Vinnicombe's variable transformation was unnecessary, and the author found a more straightforward proof. The author and Anders Rantzer jointly edited the paper, and the author revised it.

Paper IV

This paper was presented in an abridged form, without many of the proofs, at L4DC 2024 in Oxford, July 15-17.

O. Kjellqvist (July 2024b). "Minimax dual control with finite-dimensional information state". In: A. Abate et al. (Eds.). *Proc. 6th Annu. Learning Dyn. Control Conf.* Vol. 242. PMLR, pp. 299–311. URL: https://proceedings.mlr.press/v242/kjellqvist24a.html

This thesis contains the extended version of the paper, including all proofs, which is available on Arxiv:

O. Kjellqvist (2024a). *Minimax dual control with finite-dimensional information state.* DOI: 10.48550/arXiv.2312.05156

Theoretical contributions This article identifies a class of systems where the minimax optimal dual controller admits a finite-dimensional information state. The class of systems is characterized by the bounded number of solutions to the measurement equation $y_t = h(x_t)$, exemplified by the magnitude-measured integrator

$$\begin{aligned} x_{t+1} &= x_t + u_t + w_t, \\ y_t &= |x_t|. \end{aligned}$$

This class of systems is broader than one might think, as uncertain parameters can be lifted to unmeasured states. The system belongs to the class as long as the set of feasible parameter realizations is finite. Practical examples include fault detection with control reconfiguration, where the number of fault models is finite, and real-time optimization of physical systems, where the performance measurement has a set-valued inverse with a bounded number of elements. The class is not restricted to linear systems, and the parameters do not have to be constant.

Additionally, if one has only partial measurements of the state, one can often reconstruct a finite set of feasible state realizations by virtually augmenting the measurement with previous measurements (related to the observability index). We demonstrate how to construct the augmented measurement with an example: magnitude-measured input-output models of arbitrary (finite integer) order.

The information state admits recursive computation, and Theorem 1 shows the equivalence between the minimax dual control problem and an information-state dynamic programming problem. We also provide a dissipativity interpretation in Theorem 2. We apply the results to design an adaptive controller for the magnitude-measured integrator with ℓ_2 gain from w to (x, u) less than 4.

Credits The magnitude-measured integrator example is due to Bo Bernhardsson, who challenged the author to develop a method to systematically derive controllers for such systems. The author formulated the problem as a minimax dual control problem, identified the broader class of systems treated in the paper, derived the theoretical results, implemented the algorithms to verify the theoretical results, validated all algebraic expressions numerically, wrote the first draft of the paper, edited and revised it. The author thanks Bo Bernhardsson for the example, Venkatraman Renganathan for reviewing an early draft, Anders Rantzer and Tore Hägglund for helpful discussions and feedback, and the anonymous reviewers for their constructive feedback. Anders Rantzer supervised the project.

Paper V

O. Kjellqvist and A. Rantzer (2024). "Output feedback minimax adaptive control". *IEEE Transactions on Automatic Control.* Submitted

Preliminary results of this paper were presented at the 25th International Symposium on Mathematical Theory of Networks and Systems (MTNS) in Bayreuth, Germany, in 2022.

Theoretical contributions This paper extends the results in Paper III to higher-order linear MIMO systems and provides a unified treatment of state- and output-feedback minimax adaptive control, thus subsuming [Rantzer, 2021]. We show that if the uncertain parameters belong to a finite set, the minimax optimal controller admits a finite-dimensional information state. The finite-dimensional information state allows for explicit controller synthesis, and we provide heuristics for approximating the minimax optimal dual controller.

Software contributions The paper includes a Julia implementation[2] of the algorithms, which is used to verify the theoretical results.

Credits This paper results from almost five years of research, starting with Anders Rantzer's suggestion to generalize the preliminary results [Rantzer,

[2] https://github.com/kjellqvist/MinimaxAdaptiveControl.jl

2020] to the output feedback setting at the start of the author's PhD studies. The author formalized the problem and the theoretical results based on five years of discussions and investigations with Anders Rantzer. The author implemented the heuristic and the Julia package and wrote the first draft of the paper. The author and Anders Rantzer jointly edited the paper, and the author revised it. The author is thankful to colleagues Venkatraman Renganathan and Richard Pates for reviewing early drafts.

3.2 Excluded papers

The following papers are peer reviewed but excluded from this thesis because the author's contribution was minor, or because they have already been discussed in the author's Licentiate thesis [Kjellqvist, 2022].

Excluded paper I

V. Renganathan et al. (2024). "Distributed adaptive control for uncertain networks". In: *Proc. 2024 Eur. Control Conf.* Pp. 1789–1794. DOI: 10.23919/ECC64448.2024.10591151

Theoretical contributions This paper proposes a distributed adaptive control algorithm for a class of uncertain networked linear systems based on the minimax adaptive control framework proposed in [Rantzer, 2021]. The critical insight is that for the class of studied systems, the controller in [Rantzer, 2021] uses only a subset of the empirical covariance matrix. The subset can be partitioned into a set of local covariance matrices, one for each node in the network, corresponding to the interactions between the nodes and their neighbors. This locality is exploited to reduce the communication between the nodes in the network. The main theoretical contribution is to show that a certainty-equivalence controller is stabilizing with finite ℓ_2 gain from the disturbance to the state.

Credits This paper is a joint work with Venkatraman Renganathan and Anders Rantzer, the main contributors. The author contributed to the problem formulation by pointing out the uncertainty set should be the cartesian product of the individual nodes' uncertainty sets, proved Theorem 1, validated the claims, and assisted in editing the paper.

Expluded paper II

O. Kjellqvist and A. Gattami (2022). "Learning optimal team-decisions". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 1441–1446. DOI: 10.1109/CDC51059.2022.9992786

Theoretical contributions This work concerns static team decision problems where the models are unknown to the players. The goal is to minimize the losses incurred by the team as the team interacts with the environment. We employ online gradient descent to improve the policy over time. The paper's main findings concern upper bounds for the expected regret both when each player has access to the gradient, and when each player only learns the total loss incurred by the team after each action is taken (bandit). In the bandit setting, we use a "zeroth"-order gradient estimate. The gradient estimate is obtained by sampling the corners of the unit cube, as suggested in [Shamir, 2013]. This sampling strategy is a good idea in distributed settings because it does not require coordination between players.

Credits This paper is a joint work with Ather Gattami, who suggested the problem formulation and guided the project. The author derived the theoretical results, and implemented the algorithms to verify the theoretical results. Both authors wrote the first draft of the paper and edited and revised it.

Excluded paper III

O. Kjellqvist and J. Yu (2022). "On infinite-horizon system level synthesis problems". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 5238–5244. DOI: 10.1109/CDC51059.2022.9992443

Theoretical contributions This paper considers the synthesis of spatially localized controllers with delayed communication between controllers. The main contributions are twofold. Firstly, we solve the infinite-horizon state-feedback localized LQR problem with delayed communication. Previous results consider finite-impulse response approximations [Wang et al., 2018] or instantaneous communication [Yu et al., 2021]. Secondly, we combine the state feedback policy with a localized Kalman filter to synthesize localized output feedback controllers. These localized controllers are not LQ optimal but have much smaller memory requirements than those based on the finite-impulse response approximation.

This work assumes that the problem admits feasible solutions and does not discuss how to determine the feasibility.

Credits The project emerged from discussions with Jing Yu during the author's visit to the California Institute of Technology. Both authors contributed equally to the problem formulation, derivation of the theoretical results and writing and revising the paper. Jing Yu implemented the algorithms to verify the theoretical results The author reviewed the implementation.

Excluded paper IV

O. Kjellqvist and J. C. Doyle (2022). "ν-analysis: a new notion of robustness for large systems with structured uncertainties". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 2361–2366. DOI: 10.1109/CDC51059.2022.9992640

Theoretical contributions This paper argues that the current robustness measures for structured uncertainty are inadequate to analyze large systems and proposes an alternative, ν. The argument is based on the observation that structured singular values and ℓ_1-robustness measures certify stability against the largest perturbation and cannot distinguish between dense and sparse perturbations.

The work was motivated by the search for robustness measures compatible with system-level synthesis for control design. In particular, we aimed for a convex and separable quantity so one can synthesize controllers locally.

Credits This paper is part of a joint work with John Doyle, who suggested the problem formulation, and Lisa Li during the author's visit to the California Institute of Technology. The mathematical problem formulation and the theoretical results are based on discussions with John Doyle and Lisa Li, mainly conjectures and counterexamples. These discussions resulted in two papers, where the author wrote and revised [Kjellqvist and Doyle, 2022] and Lisa Li wrote and revised [Li and Doyle, 2022].

Excluded paper V

O. Kjellqvist and O. Troeng (2020). "Numerical pitfalls in Q-design". In: vol. 53. 2, pp. 4404–4408. DOI: 10.1016/j.ifacol.2020.12.368

Theoretical contributions This paper concerns the numerical stability of basis expansions of the Youla-Kucera parametrization. We show that, for quadratic constraints and objectives, using orthogonal basis expansions has numerical advantages and that numerical stability depends heavily on the choice of state-space realization. Further, we provide a realization that empirically demonstrates superior numerical stability.

Credits This work is the result of the author's master's thesis under the supervision of Olof Troeng. The author and Olof Troeng jointly formulated the problem, derived the results, implemented the software, validated the results numerically, wrote the first draft of the paper, and revised it.

4

Relations to Literature

This chapter discusses the relationship between the work in this thesis and other control-related topics. We relate problem formulations, methodologies, difficulties, and results to the literature. The strongest connections, those that have inspired the work, are already discussed in the background. The sections are intended to be self-contained but may refer to the background and the included papers.

4.1 Finite model sets

Papers I–III and V propose a practical approach by assuming a finite uncertainty set. Finite uncertainty sets retain some of the simplicity and tractability of the perfect model case while still accommodating significant parameter variations. The finite model set approach is common in several fields, some discussed below.

Multiple model adaptive estimation

Multiple-model adaptive estimation originated in the '60s with [Magill, 1965; Lainiotis, 1976]. The approach consists of a set of filters, called a filter bank, and a method of combining the estimates to form a composed estimate. There are many ways to design the filter bank, but the most common is to consider a finite set of linear systems driven by white noise. Each system is associated with a Kalman filter and an a priori probability. The posterior probabilities are computed recursively using Bayes' rule. The composed estimate is then the conditional expectation of the state given the measurements, which reduces to a weighted sum of the Kalman filter estimate for each model.

The approach easily extends to switching systems by matching a Kalman filter with each possible trajectory. In that case, the number of filters will grow exponentially with time, which has sparked research into more efficient methods. Notable numerically tractable and suboptimal algorithms for estimation in hybrid systems are the Generalized Pseudo Bayesian [Ackerson and Fu, 1970; Chang and Athans, 1978], and the Interacting Multiple Model [Blom and Bar-Shalom, 1988]. These methods efficiently circumvent the exponential growth by reinitializing the filters using the combined estimates. This way, the number of filters is kept constant. The suggested future work in Paper II of bounding the historical cost of switching between models using a finite set of filters would also alleviate the computational burden, but suboptimality could be quantified.

Multiple-model adaptive control

Multiple-model adaptive control originated in the '70s to handle linear stochastic systems with uncertain parameters belonging to a finite set. The framework was tried on equilibrium flight control of an F-8C aircraft [Athans et al., 1977] and STOL F-15 with sensor and actuator failures [Maybeck and Pogoda, 1989] with mixed results. Each model was associated with a Kalman filter and a control law. The composed control law was a weighted sum of the individual control laws, where the likelihood of the model determined the weights. Unfortunately, the closed loop is not guaranteed to be stable and the posterior probability of the "true" model is not guaranteed to converge even if the individual models are stable, controllable and observable. By a model being "true", we mean that the trajectory of the system is generated by that model. This issue is demonstrated in the following counterexample.

Example 1—Counterexample to convergence in probability
Consider the uncertain first-order linear stochastic discrete-time system

$$\begin{aligned} x(t+1) &= ax(t) \pm u(t) + w(t), \\ y(t) &= x(t) + v(t). \end{aligned} \tag{4.1}$$

In (4.1), $x(t), u(t), y(t) \in \mathbb{R}$ are the state, input and measured output. The process disturbance $w(t)$ and the measurement noise $v(t)$ are two jointly Gaussian uncorrelated white-noise random variables with $\mathbb{E}[v(t)] = \mathbb{E}[w(t)] = 0$ and $\mathbb{E}[v(t)^2] = \mathbb{E}[w(t)^2] = 1$, drawn independently and identically distributed at each time t. If we assume the initial probability of each mode to be $\rho_+(0) = \rho_-(0) = 0.5$, we will choose $u(0) = 0$. However, if we do not inject any control signal—the residual of the two Kalman filters will be the same, so $\rho_+(1) = \rho_-(1) = 0.5$. Since $u_+(y(1), y(0)) = -u_-(y(1), y(0))$ we have $u(1) = 0.5u_+(1) + 0.5u_-(1) = 0$. By similar arguments, we will have $u(t) = 0$ for all t and $\rho_+(t) = \rho_-(t)$ for all t, regardless of which model is "true".

Stability can be guaranteed, however, if certain *distinguishability* conditions are fulfilled [Silvestre et al., 2021]. The above example is treated in Paper III, where we provide a stabilizing control law by switching between controllers instead of interpolating.

The uncertain system (4.1) is stabilizable in the ℓ_p sense by a periodic controller. [Khargonekar et al., 1985] showed that any finite collection of finite-dimensional controllable LTI systems is stabilizable by periodically circulating dead-beat controllers. For instance, taking $u(t) = (-1)^t ay(t)$ stabilizes (4.1) with finite ℓ_2 gain. This result was extended to finite collections of internally *stabilizable* linear time-varying systems in [Khargonekar et al., 1988], and [Mårtensson, 1985] showed similar results for continuous sets of parametric uncertainty using an exhaustive dense search in parameter space. Like [Khargonekar et al., 1985; Khargonekar et al., 1988], [Mårtensson, 1985] did not rely on interpolating among feasible candidates. Instead, the results rely on certainty equivalence—using a prerouted search among controllers that work well for each realization.

Stephen Morse proposed using the predictive performance of each feasible model to decide which model to use for certainty equivalence control in [Morse, 1996] and proved that the closed-loop is ℓ_2 stable in [Morse, 1997]. Morse's contribution concerned linear time-invariant SISO systems with possibly uncountable uncertainty sets. It was also assumed that any realization could be satisfactorily controlled by a linear time-invariant controller based on a model from an a priori specified finite collection of "nominal" models.

Figure 4.1 illustrates a supervisory switched control system. The main difference between prerouted search amongst controllers and Morse's adaptive approach is the supervisor determining the switching sequence σ. In the prerouted case, σ is a predetermined function of time, whereas in the adaptive case, σ is a function of the control input u and the process output y.

The adaptive switching algorithms (supervisors) can roughly be divided into two categories: those based on process estimation and those based on a direct performance evaluation of each candidate controller. Our work mainly relates to the estimation-based supervisors, and we will focus on them.

The tutorial [Hespanha, 2001] contains much of the development up to 2002. [Buchstaller and French, 2016a; Buchstaller and French, 2016b] proposed an axiomatic framework providing robust stability and performance bounds for a broad class of estimation-based supervisory control schemes for MIMO LTI plants and some classes of nonlinear plants.

Our work differs from the above in that rather than starting with a finite set of controllers; we start with a set of models and an objective. The finite set of controllers and the switching law are consequences of the objective and the models. Compared to [Buchstaller and French, 2016a; Buchstaller and French, 2016b], we get less complicated conditions for finite gain stability that can be used directly for synthesis.

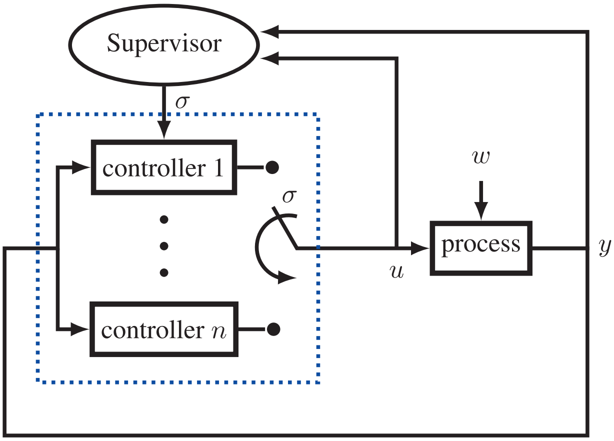


Figure 4.1 Illustration of a supervisory control architecture, recreated from [Hespanha, 2001]

4.2 Data-driven approaches

Learning controllers from data is not new; many practical tuning methods rely on identifying coarse models and applying tried and tested rules related to those parameters. For example, in the classical Ziegler-Nichols methods, the process dynamics are characterized by two parameters. The step-response method identifies a first-order time-delay model. The frequency domain method describes the ultimate point in the Nyquist plot by the ultimate gain and the ultimate frequency. Improved tuning rules exist now [Åström and Hägglund, 1995; Berner, 2017], and many can be automated. The new schemes' performance improvement comes from better-identified models, more expressive models, and improved tuning rules. When the methods that employ low-order approximate models and good tuning rules fail, one can turn to system identification [Ljung, 1999] and robust controller design [Dullerud and Paganini, 2000], or adaptive control [Åström and Wittenmark, 2008].

In the recent "data-driven control" literature, the model used for control design is the "raw" collected trajectories themselves; the controller is synthesized directly from the collected data without first identifying an intermediate model. Related methods include *iterative feedback tuning* [Hjalmarsson et al., 1998], the *intelligent PID* [Fliess and Join, 2009; Tabuada et al., 2017], *virtual reference feedback tuning* [Campi and Savaresi, 2006] and *direct adaptive control* [Åström and Wittenmark, 2008]. Recently, control-design methods based on Willems' fundamental lemma [Willems et al., 2005; Berberich et al., 2023] have gained much attention in the data-driven control community. The result states that for linear systems, future trajectories are linear combinations of past trajectories if the data is rich enough. Linear systems were studied in [De Persis and Tesi, 2020; Berberich et al., 2020], polynomial systems in [Guo et al., 2022], rational systems in [Strässer et al., 2021] and nonlinear stabilization [De Persis et al., 2023]. The recent survey [Martin et al., 2023] provides a comprehensive overview of recent progress in the field, especially on theoretical guarantees.

Model predictive control based on the behavioral framework was proposed in [Yang and Li, 2015]. Coulson et al. [Coulson et al., 2019] argued that the behavioral framework naturally addresses one of the significant shortcomings of most MPC schemes, the assumption of perfect state measurements, and introduced regularization as a means to extend the framework to stochastic nonlinear systems. Berberich et al. proved recursive feasibility, constraint satisfaction, and exponential stability in [Berberich et al., 2021]. They extended their theory of linear MPC to the nonlinear case in [Berberich et al., 2022] by adaptation, in the sense that they update the Hankel matrix online with new observations.

Our problem formulation also leverages data to improve control performance but differs from the recent data-driven frameworks in several ways. We collect data and improve performance online instead of relying on data collected from previous experiments to design a controller. We do, however, allow for different rates of parameter updates and control actions, resembling a periodic controller redesign. We take a grey box approach, allowing users to construct uncertain models and reduce uncertainty online rather than viewing the system as a black box.

4.3 Adaptive control

There is no consensus on the definition of *adaptation* in control theory. We use the definition from [Åström and Wittenmark, 2008], who consider a controller adaptive if a mechanism changes the controller parameters based on the system response.

A thorough review of the field is outside the scope of this thesis. However, the interested reader may consult the textbooks [Åström and Wittenmark, 2008; Goodwin and Sin, 2009; Ioannou and Sun, 1995] for a solid foundation, and the excellent survey [Annaswamy and Fradkov, 2021] for a comprehensive overview of the field and its connections to reinforcement learning. Instead, we will focus on the connection between our work and the adaptive control literature, starting with positioning our work in the taxonomy of adaptive control.

We consider the proposed controllers in Papers III–V as indirect adaptive controllers, as they have internal system models, one for each possible parameter realization. However, one could argue that once the observers and controllers are designed, the system models are not used in updating the control law, and one chooses the control signal based on each controller's performance metric—hence, there is no intermediate model, and the controllers are direct adaptive controllers. This conceptual shift does not alter the controllers, only their interpretation.

The controllers in Papers III and V concern fixed parameters and thus relate to *self-tuning regulators*. Introducing an exponential forgetting factor in the empirical covariance matrix allows self-tuning regulators to adapt to slowly changing parameters. This forgetting factor typically does not work well if the parameters change abruptly, as when the system switches between operating modes, as in the uncertain sign in Paper IV. It is more appropriate to include a mechanism, like a covariance reset, to detect the change and reinitialize the parameter estimation scheme. The controller in Paper IV can detect the change in the state's sign but does not reinitialize the parameter (and uncertainty) estimate, which allows it to reject a false detection more quickly.

Classical adaptive control laws may become unstable in the presence of disturbances and unmodeled dynamics; see [Ioannou and Sun, 1995, Chapter 8.3] and [Anderson, 2005] for thorough discussions. The topic "robust adaptive control" pursues the same ideal as this thesis. Robust adaptive control schemes are modifications of the classical adaptive control laws that include robustifying terms to ensure stability in the presence of disturbances and unmodeled dynamics. The idea of a "dominantly rich" control law [Ioannou and Sun, 1995, Chapter 9] is based on the observation that adaptation deals with low-frequency uncertainty and that the remaining uncertainty is in the high-frequency spectrum. The modification concerns the frequency content of the excitation signal. Care is taken to excite the system in the low-frequency range to avoid exciting the high-frequency range where the unmodeled dynamics reside. Section 2.1 discusses the frequency characteristics of different types of uncertainty. However, we approach the problem starting with robustness in mind rather than attempting to robustify a classical adaptive control law.

[Anderson, 2005] relates some of the (catastrophic) failures that can occur in adaptive control to three fundamental underlying timescales: that of the physical system and the controller, that of the adaptation, and that of the changing parameters. The timescales have a natural ordering in that the physical system and controller are the fastest, the adaptation rate is slower, and the parameter changes are the slowest. Problems tend to occur when the timescales conflict, like when adaptation is too slow to track the changing parameters and when adaptation and control start to interact.

In the self-tuning regulator, Figure 2.2, the two loops are often executed at different rates, with the parameter estimate being the slower. Paper V provides an example where the sufficienct condition of [Rantzer, 2021] does not admit a solution. The example includes a delay between the input and the effect of an uncertain parameter on the output. The resolution is only to update the controller parameters after the effect of the previous update becomes clear, formalized as a periodic Bellman inequality. Effectively, the control and parameter update loops are executed at different rates.

4.4 Reinforcement learning

Reinforcement learning considers problem formulations and methodologies for agents that learn to act in an (unknown) environment to maximize rewards. As such, it is closely related to optimal control theory, and the two fields intersect in many ways. Most of the work in reinforcement learning concerns discrete time, finite states, and action spaces, and the agent interacts with the environment through a sequence of actions. It has been successful in settings such as chess, Go, and video games.

Strong connections exist between reinforcement learning and optimal control, perhaps most apparent in the work concerning continuous states and action spaces. Notable works include the continuous counterpart of temporal-difference methods [Doya, 2000], the adaptation of the "deep Q-learning" framework [Lillicrap et al., 2019]. However, in the terminology of [Matni et al., 2019], this body of work considers *episodic tasks*: iterative experimental design and controller synthesis, assuming access to perfect simulators or the ability to reset the system safely. The work is concerned with a controller design problem, not adaptation. We point the interested reader to the surveys [Recht, 2019; Shin et al., 2019] for a control-theoretic perspective on the larger body of reinforcement learning.

Recently, there has been a surge of interest in applying theory from "single-trajectory", or "continuing task", reinforcement learning to analyze the sample complexity of adaptive control settings. The movement was initiated by [Abbasi-Yadkori and Szepesvári, 2011], who took an "optimism in the face of uncertainty" approach to the adaptive LQR problem. Most of the work in this area concerns linear quadratic control problems with perfect state measurements and stochastic disturbances. The tutorial paper [Matni et al., 2019] provides an excellent introduction and overview of the results up to 2019.

Much of the recent progress in the field has come from the combination of learning theory, much inspired by the work on bandit problems, with high-dimensional statistics and online convex optimization. [Tsiamis et al., 2023] gives an outstanding account of recent progress on the sample complexity of system identification and statistical guarantees for linear-quadratic single-trajectory problems based on high-dimensional statistics. The guarantees are typically given in terms of regret, which is the difference between the learned controller's performance and a suitable baseline and hold in expectation or high probability.

Online control

The work based on online convex optimization [Hazan, 2023] has evolved in a slightly different direction: "online nonstochastic control". The recent book [Hazan and Singh, 2023], currently in draft, provides an introduction to the field. Online nonstochastic control concerns linear systems, possibly unknown, with bounded disturbances and sequentially revealed cost functions. The cost functions are unknown apriori and revealed to the controller after each control action. The strictly causal access to the cost functions prevents closed-form solutions of the optimal controller; optimal controller synthesis occurs online, in feedback with the physical system.

For a comprehensive set of references, we direct the reader to [Hazan and Singh, 2023], but we will list a few seminal papers and highlight some intriguing similarities and differences. The state-feedback case without model uncertainty was introduced in [Agarwal et al., 2019], the uncertain case with an available stabilizing controller in [Hazan et al., 2020], and the "black-box", without a stabilizing controller, in [Chen and Hazan, 2021]. Output feedback was treated in [Simchowitz et al., 2020]. In papers IV and V, the uncertainty is initially in the observations and the system parameters, but after a reformulation, it is moved to the cost functions. Just like this thesis, the theoretical guarantees hold for adversarial disturbance models. However, the adversary selects the disturbance sequences in online control apriori, which may not depend on the control signal realization. This thesis allows for disturbances dependent on both estimates and control signals.

In all the above papers, the control law comprises two parts: a stabilizing state feedback, or dynamical output feedback, controller, and a finite impulse response (FIR) component operating on a finite window of past disturbances Online optimization is used to tune the FIR component, and the window length is a parameter that depends on the control horizon and specific system quantities. The infinite horizon case is treated by a doubling trick, meaning that the memory required to store the FIR components goes to infinity.

Interestingly, none of the above papers discuss the implications of nonminimum phase behavior, like right halfplane zeros and delays, on the achievable performance. Nonminimum phase behavior is a commonly acknowledged problem in adaptive control. It can occur in both state and output-feedback settings. Online optimization has an extra feedback loop that includes the cost function. This cost function may have right-half plane zeros or delay-like behavior even if the controller can access perfect state measurements. In model predictive control, it is well known that nonminimum phase costs may require a long prediction horizon to be stable. The counterexample in Paper

V, which shows the nonexistence of a solution to the LMIs in [Rantzer, 2021], is based on a nonminimum phase cost function. Periodic parameter updates resolve it.

Convex body chasing

For systems with (large) parametric uncertainty, [Ho et al., 2021] introduces a new take on certainty equivalence control. The procedure consists of a robust oracle that provides a controller that robustly satisfies performance measures and a parameter selection mechanism that selects parameters. The procedure is guaranteed to make only a finite number of mistakes, given that each oracle would satisfy the specifications in a small neighborhood around the "true" parameters.

The parameter selection mechanism finds the set of parameters consistent with the observed trajectory and then selects a point in this set for certainty equivalence control. Under the assumption that the feasible sets are convex, the selection procedure is posed as an online convex body-chasing problem. In convex body chasing, the goal is select points $p_0, p_1, \dots$ from a sequence of convex sets $K_0, K_1, \dots$ such that the path length, $|p_1 - p_0| + |p_2 - p_1| + \dots$ is minimized.

The initial work [Ho et al., 2021] concerns fixed parameters, for which the feasible sets are nested (monotonically decreasing in terms of set inclusion). For nested convex body chasing, the path length is bounded by the diameter of the first set. By combining the robustness of the oracle with the finite path length, the controller is guaranteed to make only a finite number of mistakes.

The framework is extended to slowly-varying linear time-varying systems in [Yu et al., 2023], where the feasible sets are no longer nested, using recent results for convex function and body chasing [Sellke, 2023]. The controller guarantees bounded-input bounded-output stability of the closed loop.

Like in Papers III—V, this work uses certainty equivalence control, but the parameter set may be infinite. The focus is on state feedback, but [Ho et al., 2021] remarks that the framework can be applied to the output feedback setting if a common observer state exists so that the robust control oracle is a memoryless function of this shared state. In Paper V's output feedback case, each observer has its own observer state. However, we exploit a reformulation where they are stacked into a vector to reduce output feedback to an instance of the principal problem. This stacked vector then becomes a shared state that the controller uses.

The performance metrics

Regret quantifies a learning method's accumulated suboptimality and gives a clear picture of its learning performance, but the implications for system theoretic properties like stability, sensitivity, bandwidth, et cetera still need to be better understood. [Karapetyan et al., 2023] provides a first step in this direction for linear systems with linear controllers, but the results do not apply to learning systems as they are inherently nonlinear. The performance metric of papers III–V is the soft-constrained reformulation of the induced gain, which has a transparent system theoretic interpretation. Our guarantees are "anytime" and hold for the moment the controller is initialized. However, the worst-case approach may be overly conservative, and optimizing regret instead could potentially lead to more efficient controllers. Further, statistical guarantees make more sense if accurate probabilistic models of the system parameters and noise realization can be provided.

4.5 Dual control

For linear systems with uncertain parameters[1], the relationship between parameter observability (more commonly known as identifiability) and inputs is well understood through the concept of excitation. Perhaps the first work to scrutinize the effects of this interplay in optimal control was Alexander Aronovich Feldbâum [Feldbâum, 1963]. Feldbâum argued that optimal controllers would have two distinct traits for a large class of uncertain systems: they would ensure the system is well-behaved and probe the system for additional information. These traits are called exploitation and exploration in reinforcement learning. Controllers that exhibit these traits are called *dual controllers*. The interested reader is referred to the book [Åström and Wittenmark, 2008, Chapter 7.] [Feldbâum, 1963], and the surveys [Filatov and Unbehauen, 2000; Mesbah, 2018].

Stochastic dual control formulates adaptive control as a stochastic optimal control problem, where uncertain disturbances and parameters are drawn randomly from apriori known distributions. The posterior distribution of the states and uncertain parameters is a sufficient statistic, sometimes called an information state or a hyperstate. Interestingly, the hyperstate does not depend on the objective, only the system dynamics and observations. This is in contrast to the controllers in Papers III–V, and minimax control in general [Witsenhausen, 1966; Bertsekas and Rhodes, 1973; James and Baras, 1995], where the information state characterizes the worst-case realization of the cost functions consistent with data.

The hyperstate sometimes has a finite-dimensional representation that can be updated recursively, like Markov decision processes with finite state and action spaces. If the hyperstate is infinite-dimensional, one can approximate it with a finite-dimensional state: [Alspach, 1972] proposes using a sum of Gaussians.

[1] Such systems are not *jointly* linear in the previous sense

In Papers III–V, the hyperstate describes one quadratic function (of the state) per hypothesis. These functions' offsets and linear terms are updated recursively but depend on the hypothesis' parameter values. These quadratic functions describe precisely the worst-case realization of the costs under each hypothesis. Like Alspach, we see great potential in relaxing the restriction to finite uncertainty sets by approximating the worst-case cost with a finite set of quadratic functions. This would be realized practically by associating a robust observer with each member in a finite covering of the uncertainty set. The robust observer would provide over- and under-approximations of the worst-case cost valid for any parameter realization in its covering member.

[Filatov and Unbehauen, 2000] classifies dual control into *direct* and *indirect* methods. A direct method is a controller with an explicit exploration term, like adding white noise to the control signal to ensure excitation. An indirect method is a controller that is a solution, or an approximation, to an optimal dual control problem. The controllers in Papers III–V are indirect dual controllers, as they are suboptimal solutions to dual control problems.

5

Concluding remarks

This thesis has presented a framework for minimax (sub)optimal control and estimation under additive and parametric uncertainty. We conclude by summarizing the key findings and discussing limitations and future work.

5.1 Summary of key findings

- Paper I: Characterizes the minimax optimal output predictor for a finite set of linear systems, providing a framework for robust output prediction.

- Paper II: Presents a priori computable performance bounds for the estimator developed in Paper I, enhancing the reliability of the estimation process under uncertainty.

- Paper III: Proposes a minimax suboptimal adaptive controller for scalar systems with unknown sign in the b-parameter, achieving a finite ℓ_2-gain and providing an upper bound, which is crucial for ensuring system stability.

- Paper IV: Introduces a class of systems for which the minimax optimal dual controller admits a finite-dimensional information state and derives a suboptimal controller for the magnitude-measured integrator with an ℓ_2-gain less than 4, offering practical control solutions.

- Paper V: Proposes a unified framework for minimax adaptive control that combines output feedback and state feedback cases, along with heuristics for explicit controller synthesis, broadening the applicability of minimax control strategies.

5.2 Limitations and future work

Finite-dimensional information state: The results in this thesis are based on reformulating uncertainty in the system model or measurements as an optimization variable in the objective function. We have limited our attention to settings where we can represent this reformulation by a finite number of variables. For example, Paper IV is limited to systems where the preimage of the measurement map $y_t = h(x_t)$ is a finite set, and Papers I–III and V are limited to parametric uncertainty with finite parameter sets.

A natural extension is to consider infinite uncertainty sets. The challenge, however, is that the information state may be infinite-dimensional, which makes the optimization problem intractable. As a first step, consider a finite cover of the infinite set and associate a robust observer with each element in the cover. Just like we could express the optimization problem for finite uncertainty sets by associating one observer and performance quantity with each element in the set, we could instead *approximate* the optimization problem with one robust observer and performance quantity for each element in the cover.

Conservative bounds The bounds in Papers I–V are conservative; there are three reasons for this:

1. We employ a worst-case analysis, meaning we must consider all possible uncertainty realizations. This is a conservative approach, as the actual realization of the uncertainty tends not to be the worst-case.
2. The bounds in Papers II–V approximate the minimax optimal solution with a suboptimal solution. In Paper II, we investigated the optimality gap empirically, and it was found to be small in certain cases and larger in others. For the other papers, the optimality gap is not investigated.
3. Our suboptimal controllers, based on certainty equivalence, do not contain an active mechanism to explore the uncertainty. This is mitigated to some extent by two assumptions about the parameter set: that the set is finite and the parameters are fixed.

The first point is a limitation of the worst-case analysis and can be mitigated by stronger assumptions on the uncertainty set or different uncertainty models. Concerning structured uncertainty, as our framework provides bounded ℓ_2-gain, it can be combined with robust control techniques like $D-K$ iterations or integral quadratic constraints to provide less conservative bounds. The second point warrants further investigation, and the suboptimal controllers may be closer to the optimal solution than our bounds indicate. The third point is a limitation of the certainty equivalence approach; however, as our controllers and value function approximations are feasible, they can be used as starting points for more complex function approximations and numerical optimization. We are interested in exploring reinforcement learning to approximate the value function, Q-function and policies.

Lower bounds Although the thesis' synthesis methods are based on upper bounds and sufficiency conditions, its framework and tools are well suited for studying lower bounds. We conclude this discussion by deriving a new lower bound on the achievable ℓ_2-gain in the setting of [Megretski and Rantzer, 2003, Theorem 1]. This proof not only underscores the utility of the approach taken in this thesis but also serves as a delightful finale for the mathematically inclined reader.

Consider the system

$$\begin{aligned} x_{t+1} &= ax_t + u_t + w_t, \quad \forall t \geq 0 \\ u_t &= \mu_t(x_t, x_{t-1}, \ldots, x_0), \end{aligned} \tag{5.1}$$

where $x \in \mathbb{R}$, $u_t \in \mathbb{R}$ and $w_t \in \mathbb{R}$ are the state, control input and disturbance at time t, respectively. The control input u_t is a function of the state and past states, and agnostic of the uncertain parameter $a \in [a_0, a_1] \subset \mathbb{R}$, but aware of the end points $a_0 < a_1$.

THEOREM 1
Given real numbers $a_1 > a_0$ and a quantity $0 < \gamma < \sqrt{1 + (a_1 - a_0)^2/4}$. Then, for any causal control law μ_t in (5.1) and initial condition $x_0 \in \mathbb{R}$, there exists a parameter value $a \in [a_0, a_1]$ so that, under the dynamics (5.1),

$$J_\mu^T(a) \triangleq \sup_w \left\{ \sum_{t=0}^T x_t^2 - \gamma^2 \sum_{t=0}^T w_t^2 \right\} = \infty,$$

for all $T \geq 3$.

Proof. Let $0 < \gamma < \sqrt{1 + (a_1 - a_0)^2/4}$. The quantity $J_\mu^T(a)$ is monotonically non-decreasing in T, so it suffices to show that $J_\mu^3(a) = \infty$. Consider the change of variables $v_t = ax_t + u_t + w_t$. Just like in Paper IV, we can rewrite the dynamics into a terminal cost problem:

$$\begin{aligned} x_{t+1} &= v_t \\ r_{t+1}(a) &= r_t(a) + x_t^2 - \gamma^2 |v_t - ax_t - u_t|^2, \quad r_0(a) = 0, \end{aligned} \tag{5.2}$$

where v_t is chosen by the adversary, and $r_t(a) : [a_0, a_1] \to \mathbb{R}$ is a sequence of functions so that $J_\mu^3(a) = \sup_v r_3(a)$.

Even though the realization a is unknown to the controller, the function r_t is constructed using observations of the state $(x_t, x_{t-1}, \ldots)$ and control inputs $(u_{t-1}, u_{t-2}$. Therefore, the controller can choose the control input u_t based on the function r_t and the state x_t. We will now bound $\sup_{a_0 \le a \le a_1} J^3_\mu(a)$ from below using dynamic programming. Consider the value functions

$$\begin{aligned} V_3(x,r) &= \max_{a_0 \le a \le a_1} r(a) \\ V_{k-1}(x,r) &= \min_u \max_v \{V_k(v, a \mapsto r(a) + x^2 - \gamma^2 |v - ax - u|^2)\}, \end{aligned} \tag{5.3}$$

for $k = 3, 2, 1$. The value functions V_k are functions of the state x and the function r and $V_{k-1} \ge V_k$. The notation $a \mapsto r(a) + x^2 - \gamma^2|v - ax - u|^2$ defines a new function of a corresponding to the dynamics (5.2). By standard dynamic programming arguments, $V_0(x_0, 0) = \inf_\mu \sup_{a_0 \le a \le a_1} J^3_\mu(a)$. Explicit computation shows

$$\begin{aligned} V_2(x,r) &= \min_u \max_v \{V_3(v, a \mapsto r(a) + x^2 - \gamma^2|v - ax - u|^2)\} \\ &= \min_u \max_a \{ \underbrace{r(a) + x^2}_{\text{Independent of } u} \underbrace{- \gamma^2 \min_v |v - ax - u|^2}_{=0} \} \\ &= \max_a \left\{ x^2 + r(a) \right\}. \end{aligned}$$

For $V_1(x,r)$, we have

$$\begin{aligned} V_1(x,r) &= \min_u \max_v \{V_2(v, a \mapsto r(a) + x^2 - \gamma^2|v - ax - u|^2)\} \\ &= \min_u \max_{a,v} \{x^2 + v^2 - \gamma^2|v - ax - u|^2 + r(a)\} \end{aligned}$$

The maximization over v is unbounded for $\gamma^2 \le 1$. For $\gamma^2 > 1$, the maximizing v is unique and given by $v^\star = -(1-\gamma^2)^{-1}\gamma^2(ax+u)$, thus

$$\begin{aligned} V_1(x,r) &= \min_u \max_a \left\{ x^2 + \left(\frac{\gamma^4}{\gamma^2 - 1} - \gamma^2 \right) |ax + u|^2 + r(a) \right\} \\ &= \min_u \max_a \left\{ x^2 + \frac{|ax+u|^2}{1 - \gamma^{-2}} + r(a) \right\}. \end{aligned}$$

We will bound the maximum over a using $(1-\gamma^{-2})^{-1} > 1$ and the naive bound

$$\max_a \left\{ |ax+u|^2 + r(a) \right\} \ge \max_a |ax+u|^2 + \min_a r(a).$$

Denote the interval's length by $L = a_1 - a_0$, then $\max_a |ax+u|^2 \ge L^2/4|x^2|$ for any u. and

$$V_1(x,r) \ge \min_a \left\{ \left(1 + \left(\frac{L}{2} \right)^2 \right) x^2 + r(a) \right\}.$$

As the Bellman operator is monotone in V, we have

$$\begin{aligned}
V_0(x,r) &= \min_u \max_v \{V_1(v, a \mapsto r(a) + x^2 - \gamma^2|v - ax - u|^2)\} \\
&\geq \min_u \max_v \min_a \{x^2 + \left(1 + \left(\frac{L}{2}\right)^2\right) v^2 - \gamma^2|v - ax - u|^2 + r(a)\} \\
&\geq \min_u \max_v \min_a \left\{ x^2 + \left(1 + \left(\frac{L}{2}\right)^2 - \gamma^2\right) v^2 \right. \\
&\qquad \left. + \gamma^2 v(ax - u) - \gamma^2|ax + u|^2 + r(a) \right\}
\end{aligned}$$

By assumption $\gamma^2 < 1 + L^2/4$, so the quadratic term in v is strictly convex and the right-hand side is unbounded. Therefore $\sup_{a_0 \leq a \leq a_1} J^3_\mu(a) = \infty$. $\square$

References

Abbasi-Yadkori, Y. and C. Szepesvári (2011). "Regret bounds for the adaptive control of linear quadratic systems". In: Kakade, S. M. et al. (Eds.). *Proc. 24th Conf. Learn. Theory*. Vol. 19. PMLR, Budapest, Hungary, pp. 1–26. URL: https://proceedings.mlr.press/v19/abbasi-yadkori11a.html.

Ackerson, G. and K. Fu (1970). "On state estimation in switching environments". *IEEE Trans. Autom. Control* **15**:1, pp. 10–17. DOI: 10.1109/TAC.1970.1099359.

Agarwal, N., B. Bullins, E. Hazan, S. Kakade, and K. Singh (2019). "Online control with adversarial disturbances". In: Chaudhuri, K. et al. (Eds.). *Proc. 36th Int. Conf. Mach. Learn.* Vol. 97. PMLR, pp. 111–119. URL: https://proceedings.mlr.press/v97/agarwal19c.html.

Alleyne, A. et al. (2023). *Control for Societal-scale Challenges: Road Map 2030*. English. Ed. by A. Annaswamy, K. Johansson, and G. Pappas. IEEE Control Systems Society.

Alspach, D. (1972). "Dual control based on approximate a posteriori density functions". *IEEE Trans. Autom. Control* **17**:5, pp. 689–693. DOI: 10.1109/TAC.1972.1100099.

Anderson, B. and A. Dehghani (2008). "Challenges of adaptive control-past, permanent and future". *Annu. Rev. Control* **32**, pp. 123–135. DOI: 10.1016/j.arcontrol.2008.06.001.

Anderson, B. D. O. (2005). "Failures of Adaptive Control Theory and their Resolution". *Commun. Inf. Syst.* **5**:1, pp. 1–20. DOI: 10.4310/CIS.2005.v5.n1.a1.

Annaswamy, A. M. and A. L. Fradkov (2021). "A historical perspective of adaptive control and learning". *Annu. Rev. Control* **52**, pp. 18–41. ISSN: 1367-5788. DOI: 10.1016/j.arcontrol.2021.10.014.

Åström, K. and T. Hägglund (1995). *PID Controllers: Theory, Design, and Tuning*. English. 2nd ed. ISA - The Instrumentation, Systems and Automation Society, USA. ISBN: 1-55617-516-7.

Åström, K. and B. Wittenmark (2008). *Adaptive Control (2 ed.)* English. Dover, Mineola, NY, USA. ISBN: 0-201-55866-1.

Athans, M., D. Castanon, K. Dunn, C. Greene, W. Lee, N. Sandell, and A. Willsky (1977). "The stochastic control of the f-8c aircraft using a multiple model adaptive control (mmac) method–part i: equilibrium flight". *IEEE Trans. Autom. Control* **22**:5, pp. 768–780. DOI: 10.1109/TAC.1977.1101599.

Basar, T. and P. Bernhard (2008). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach.* 2nd ed. Birkhauser, Boston, MA, USA. DOI: 10.1007/978-0-8176-4757-5.

Bencherki, F. and A. Rantzer (2023). "Robust simultaneous stabilization via minimax adaptive control". In: *Proc. 62nd IEEE Conf. Decis. Control*, pp. 2503–2508. DOI: 10.1109/CDC49753.2023.10384134.

Berberich, J., A. Iannelli, A. Padoan, J. Coulson, F. Dörfler, and F. Allgöwer (2023). "A quantitative and constructive proof of willems' fundamental lemma and its implications". In: *Proc. 2023 IEEE Amer. Control Conf.* Pp. 4155–4160. DOI: 10.23919/ACC55779.2023.10156227.

Berberich, J., A. Koch, C. W. Scherer, and F. Allgöwer (2020). "Robust data-driven state-feedback design". In: *Proc. 2020 IEEE Amer. Control Conf.* Pp. 1532–1538. DOI: 10.23919/ACC45564.2020.9147320.

Berberich, J., J. Köhler, M. A. Müller, and F. Allgöwer (2021). "Data-driven model predictive control with stability and robustness guarantees". *IEEE Trans. Autom. Control* **66**:4, pp. 1702–1717. DOI: 10.1109/TAC.2020.3000182.

Berberich, J., J. Köhler, M. A. Müller, and F. Allgöwer (2022). "Linear tracking mpc for nonlinear systems—part ii: the data-driven case". *IEEE Transactions on Automatic Control* **67**:9, pp. 4406–4421. DOI: 10.1109/TAC.2022.3166851.

Berner, J. (2017). *Automatic Controller Tuning using Relay-based Model Identification.* PhD thesis. Department of Automatic Control, Lund University, Sweden. ISBN: 978-91-7753-446-4.

Bertsekas, D. and I. Rhodes (1973). "Sufficiently informative functions and the minimax feedback control of uncertain dynamic systems". *IEEE Trans. Autom. Control* **18**:2, pp. 117–124. DOI: 10.1109/TAC.1973.1100241.

Blom, H. A. P. and Y. Bar-Shalom (1988). "The interacting multiple model algorithm for systems with markovian switching coefficients". *IEEE Trans. Autom. Control* **33**:8, pp. 780–783. DOI: 10.1109/9.1299.

Buchstaller, D. and M. French (2016a). "Robust stability for multiple model adaptive control: part i—the framework". *IEEE Trans. Autom. Control* **61**:3, pp. 677–692. DOI: 10.1109/TAC.2015.2492518.

Buchstaller, D. and M. French (2016b). "Robust stability for multiple model adaptive control: part ii—gain bounds". *IEEE Trans. Autom. Control* **61**:3, pp. 693–708. DOI: 10.1109/TAC.2015.2492503.

Campi, M. and S. Savaresi (2006). "Direct nonlinear control design: the virtual reference feedback tuning (vrft) approach". *IEEE Trans. Autom. Control* **51**:1, pp. 14–27. DOI: 10.1109/TAC.2005.861689.

Cederberg, D., A. Hansson, and A. Rantzer (2022). "Synthesis of minimax adaptive controller for a finite set of linear systems". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 1380–1384. DOI: 10.1109/CDC51059.2022.9993060.

Chang, C. B. and M. Athans (1978). "State estimation for discrete systems with switching parameters". *IEEE Trans. Aerosp. Electron. Syst.* **AES-14**:3, pp. 418–425. DOI: 10.1109/TAES.1978.308603.

Chen, X. and E. Hazan (2021). "Black-box control for linear dynamical systems". In: Belkin, M. et al. (Eds.). *Proc. 34th Conf. Learn. Theory*. Vol. 134. PMLR, pp. 1114–1143. URL: https://proceedings.mlr.press/v134/chen21c.html.

Coulson, J., J. Lygeros, and F. Dörfler (2019). "Data-enabled predictive control: in the shallows of the deepc". In: *Proc. 2019 Eur. Control Conf.* Pp. 307–312. DOI: 10.23919/ECC.2019.8795639.

De Persis, C., M. Rotulo, and P. Tesi (2023). "Learning controllers from data via approximate nonlinearity cancellation". *IEEE Trans. Autom. Control* **68**:10, pp. 6082–6097. DOI: 10.1109/TAC.2023.3234889.

De Persis, C. and P. Tesi (2020). "Formulas for data-driven control: stabilization, optimality, and robustness". *IEEE Trans. Autom. Control* **65**:3, pp. 909–924. DOI: 10.1109/TAC.2019.2959924.

Didinsky, G. and T. Basar (1994). "Minimax adaptive control of uncertain plants". In: *Proc. 33rd IEEE Conf. Decis. Control*. Vol. 3, pp. 2839–2844. DOI: 10.1109/CDC.1994.411368.

Dorato, P. (1987). "A historical review of robust control". *IEEE Control Syst. Mag.* **7**:2, pp. 44–47. DOI: 10.1109/MCS.1987.1105273.

Doya, K. (2000). "Reinforcement learning in continuous time and space". *Neural Comput.* **12**, pp. 219–45. DOI: 10.1162/089976600300015961.

Doyle, J. (1982). "Analysis of feedback systems with structured uncertainties". *IEE Proc. Control Theory Appl.* **129**:6, pp. 242–250. ISSN: 0143-7054. DOI: 10.1049/ip-d.1982.0053.

Doyle, J. C., J. E. Wall, and G. Stein (1982). "Performance and robustness analysis for structured uncertainty". In: *Proc. 21st IEEE Conf. Decis. Control*, pp. 629–636. DOI: 10.1109/CDC.1982.268218.

Dullerud, G. and F. Paganini (2000). *A Course in Robust Control Theory a convex approach*. ISBN: 978-1-4419-3189-4. DOI: 10.1007/978-1-4757-3290-0.

Fan, M. and A. Tits (1986). "Characterization and efficient computation of the structured singular value". *IEEE Trans. Autom. Control* **31**:8, pp. 734–743. DOI: 10.1109/TAC.1986.1104388.

Fan, M., A. Tits, and J. Doyle (1991). "Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics". *IEEE Trans. Autom. Control* **36**:1, pp. 25–38. DOI: 10.1109/9.62265.

Fan, M. K., A. L. Tits, and J. C. Doyle (1988). "Robustness in the presence of joint parametric uncertainty and unmodeled dynamics". In: *Proc. 1988 IEEE Amer. Control Conf.* Pp. 1195–1200. DOI: 10.23919/ACC.1988.4789902.

Feintuch, A. (1998). *Robust control theory in Hilbert space*. Springer, New York, NY, USA. DOI: 10.1007/978-1-4612-0591-3.

Feldbâum, A. (1963). "Dual control theory problems". In: *Proc. 2nd (IFAC) Congr. Automat. Remote Control Theory*. Vol. 1. 2, pp. 541–550. DOI: 10.1016/S1474-6670(17)69687-3.

Filatov, N. and H. Unbehauen (2000). "Survey of adaptive dual control methods". *IEE Proc. Control Theory Appl.* **147**, pp. 118–128. DOI: 10.1049/ip-cta:20000107.

Fliess, M. and C. Join (2009). "Model-free control and intelligent pid controllers: towards a possible trivialization of nonlinear control?" In: *Proc. 15th IFAC Symp. Syst. Identif.* Vol. 42. 10, pp. 1531–1550. DOI: 10.3182/20090706-3-FR-2004.00256.

Francis, B. A. (1987). *A course in $\mathcal{H}_\infty$ control theory*. Springer Berlin, Heidelberg, Germany.

Goodwin, G. C. and K. S. Sin (2009). *Adaptive Filtering Prediction and Control*. Dover Publications, Inc., USA. ISBN: 0486469328.

Guo, M., C. De Persis, and P. Tesi (2022). "Data-driven stabilization of nonlinear polynomial systems with noisy data". *IEEE Trans. Autom. Control* **67**:8, pp. 4210–4217. DOI: 10.1109/TAC.2021.3115436.

Hazan, E. (2023). *Introduction to online convex optimization*. arXiv: 1909.05207 [cs.LG].

Hazan, E., S. Kakade, and K. Singh (2020). "The nonstochastic control problem". In: Kontorovich, A. et al. (Eds.). *Proc. 31st Annu. Conf. Algorithmic Learn. Theory*. Vol. 117. PMLR, pp. 408–421. URL: https://proceedings.mlr.press/v117/hazan20a.html.

Hazan, E. and K. Singh (2023). *Introduction to online nonstochastic control*. arXiv: 2211.09619 [cs.LG]. URL: https://arxiv.org/abs/2211.09619.

Hespanha, J. P. (2001). "Tutorial on supervisory control". In: *40th IEEE Conf. Decis. Control*. Lecture notes for the workshop *Control using Logic and Switching*. Orlando, FL, USA.

Hjalmarsson, H., M. Gevers, S. Gunnarsson, and O. Lequin (1998). “Iterative feedback tuning: theory and applications”. *IEEE Control Syst. Mag.* **18**:4, pp. 26–41. DOI: 10.1109/37.710876.

Ho, D., H. Le, J. Doyle, and Y. Yue (2021). “Online robust control of nonlinear systems with large uncertainty”. In: Banerjee, A. et al. (Eds.). *Proc. 24th Int. Conf. Artif. Intell. Statist.* Vol. 130. PMLR, pp. 3475–3483. URL: https://proceedings.mlr.press/v130/ho21a.html.

Ioannou, P. A. and J. Sun (1995). *Robust adaptive control.* Prentice-Hall, Inc., USA. ISBN: 0134391004.

James, M. and J. Baras (1995). “Robust $\mathcal{H}_\infty$ output feedback control for nonlinear systems”. *IEEE Trans. Autom. Control* **40**:6, pp. 1007–1017. DOI: 10.1109/9.388678.

James, M. and S. Yuliar (1995). “A nonlinear partially observed differential game with a finite-dimensional information state”. *Syst. Control Lett.* **26**:2, pp. 137–145. ISSN: 0167-6911. DOI: 10.1016/0167-6911(95)00005-T.

James, M. R. (1995). “Recent developments in nonlinear $\mathcal{H}_\infty$ control”. **28**:14, pp. 541–552. ISSN: 1474-6670. DOI: 10.1016/S1474-6670(17)46885-6.

Karapetyan, A., A. Tsiamis, E. C. Balta, A. Iannelli, and J. Lygeros (2023). “Implications of regret on stability of linear dynamical systems”. In: *Proc. 22nd IFAC World Congr.* Vol. 56. 2, pp. 2583–2588. DOI: 10.1016/j.ifacol.2023.10.1342.

Khargonekar, P., K. Poolla, and A. Tannenbaum (1985). “Robust control of linear time-invariant plants using periodic compensation”. *IEEE Trans. Autom. Control* **30**:11, pp. 1088–1096. DOI: 10.1109/TAC.1985.1103841.

Khargonekar, P., A. Pascoal, and R. Ravi (1988). “Strong, simultaneous, and reliable stabilization of finite-dimensional linear time-varying plants”. *IEEE Trans. Autom. Control* **33**:12, pp. 1158–1161. DOI: 10.1109/9.14439.

Kjellqvist, O. (2022). *On Control and Estimation of Large and Uncertain Systems.* English. Licentiate Thesis. Department of Automatic Control.

Kjellqvist, O. (2024a). *Minimax dual control with finite-dimensional information state.* DOI: 10.48550/arXiv.2312.05156.

Kjellqvist, O. (2024b). “Minimax dual control with finite-dimensional information state”. In: Abate, A. et al. (Eds.). *Proc. 6th Annu. Learning Dyn. Control Conf.* Vol. 242. PMLR, pp. 299–311. URL: https://proceedings.mlr.press/v242/kjellqvist24a.html.

Kjellqvist, O. (2024c). “Minimax performance limits for multiple-model estimation”. In: *Proc. 2024 Eur. Control Conf.* Pp. 2540–2546. DOI: 10.23919/ECC64448.2024.10590947.

Kjellqvist, O. (2024d). *Minimax performance limits for multiple-model estimation*. arXiv: 2312.05159 [math.OC]. URL: https://arxiv.org/abs/2312.05159.

Kjellqvist, O. and J. C. Doyle (2022). "ν-analysis: a new notion of robustness for large systems with structured uncertainties". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 2361–2366. DOI: 10.1109/CDC51059.2022.9992640.

Kjellqvist, O. and A. Gattami (2022). "Learning optimal team-decisions". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 1441–1446. DOI: 10.1109/CDC51059.2022.9992786.

Kjellqvist, O. and A. Rantzer (2022a). *Learning-enabled robust control with noisy measurements*. DOI: 10.48550/ARXIV.2202.08363.

Kjellqvist, O. and A. Rantzer (2022b). "Learning-enabled robust control with noisy measurements". In: Firoozi, R. et al. (Eds.). *Proc. 4th Annu. Learning Dyn. Control Conf.* Vol. 168. PMLR, pp. 86–96. URL: https://proceedings.mlr.press/v168/kjellqvist22a.html.

Kjellqvist, O. and A. Rantzer (2022c). "Minimax adaptive estimation for finite sets of linear systems". In: *Proc. 2022 IEEE Amer. Control Conf.* Pp. 260–265. DOI: 10.23919/ACC53348.2022.9867474.

Kjellqvist, O. and A. Rantzer (2024). "Output feedback minimax adaptive control". *IEEE Transactions on Automatic Control*. Submitted.

Kjellqvist, O. and O. Troeng (2020). "Numerical pitfalls in Q-design". In: vol. 53. 2, pp. 4404–4408. DOI: 10.1016/j.ifacol.2020.12.368.

Kjellqvist, O. and J. Yu (2022). "On infinite-horizon system level synthesis problems". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 5238–5244. DOI: 10.1109/CDC51059.2022.9992443.

Lainiotis, D. G. (1976). "Partitioning: a unifying framework for adaptive systems, i: estimation". *Proc. IEEE* **64**:8, pp. 1126–1143. DOI: 10.1109/PROC.1976.10284.

Li, J. S. L. and J. C. Doyle (2022). "Distributed robust control for systems with structured uncertainties". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 1702–1707. DOI: 10.1109/CDC51059.2022.9992622.

Liberzon, M. R. (2001). "Lur'e problem of absolute stability - a historical essay". In: *Proc. 5th IFAC Symp. Nonlinear Control Syst.* Vol. 34. 6. St Petersburg, Russia, pp. 25–28. DOI: 10.1016/S1474-6670(17)35141-8.

Lillicrap, T. P., J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra (2019). *Continuous control with deep reinforcement learning*. DOI: 10.48550/arXiv.1509.02971. arXiv: 1509.02971 [cs.LG].

Ljung, L. (1999). *System identification : theory for the user*. 2nd ed. Prentice-Hall, Upper Saddle River, NJ, USA, p. 609. ISBN: 0-13-656695-2.

Magill, D. (1965). “Optimal adaptive estimation of sampled stochastic processes”. *IEEE Trans. Autom. Control* **10**:4, pp. 434–439. DOI: 10.1109/TAC.1965.1098191.

Mårtensson, B. (1985). “The order of any stabilizing regulator is sufficient a priori information for adaptive stabilization”. *Syst. Control Lett.* **6**:2, pp. 87–91. ISSN: 0167-6911. DOI: 10.1016/0167-6911(85)90002-7.

Martin, T., T. B. Schön, and F. Allgöwer (2023). “Guarantees for data-driven control of nonlinear systems using semidefinite programming: a survey”. *Annu. Rev. Control* **56**, p. 100911. ISSN: 1367-5788. DOI: 10.1016/j.arcontrol.2023.100911.

Matni, N., A. Proutiere, A. Rantzer, and S. Tu (2019). “From self-tuning regulators to reinforcement learning and back again”. In: *Proc. 58th IEEE Conf. Decis. Control*, pp. 3724–3740. DOI: 10.1109/CDC40024.2019.9029916.

Maybeck, P. and D. Pogoda (1989). “Multiple model adaptive controller for the stol f-15 with sensor/actuator failures”. In: *Proc. 28th IEEE Conf. Decis. Control*, 1566–1572 vol.2. DOI: 10.1109/CDC.1989.70412.

Megretski, A. and A. Rantzer (1997). “System analysis via integral quadratic constraints”. *IEEE Trans. Autom. Control* **42**:6, pp. 819–830. DOI: 10.1109/9.587335.

Megretski, A. and A. Rantzer (2003). *Bounds on the optimal ℓ_2-gain in adaptive control of a first order linear system*. Tech. rep. Mittag-Leffler, Stockholm, Sweden. ISRN: IML-R-41-02/03-SE+spring.

Mesbah, A. (2018). “Stochastic model predictive control with active uncertainty learning: a survey on dual control”. *Annu. Rev. Control* **45**, pp. 107–117. ISSN: 1367-5788. DOI: 10.1016/j.arcontrol.2017.11.001.

Morse, A. (1996). “Supervisory control of families of linear set-point controllers - part i. exact matching”. *IEEE Trans. Autom. Control* **41**:10, pp. 1413–1431. DOI: 10.1109/9.539424.

Morse, A. (1997). “Supervisory control of families of linear set-point controllers. 2. robustness”. *IEEE Trans. Autom. Control* **42**:11, pp. 1500–1515. DOI: 10.1109/9.649687.

Packard, A., M. Fan, and J. Doyle (1988). “A power method for the structured singular value”. In: *Proc. 27th IEEE Conf. Decis. Control*, 2132–2137 vol.3. DOI: 10.1109/CDC.1988.194710.

Pan, Z. and T. Basar (1998). “Adaptive controller design for tracking and disturbance attenuation in parametric strict-feedback nonlinear systems”. *IEEE Trans. Autom. Control* **43**:8, pp. 1066–1083. DOI: 10.1109/9.704978.

Poola, K. and A. Tikku (1995). “Robust performance against time-varying structured perturbations”. *IEEE Trans. Autom. Control* **40**:9, pp. 1589–1602. DOI: 10.1109/9.412628.

Popov, V. M. and R. Georgescu (1973). *Hyperstability of Control Systems*. Springer-Verlag, Berlin, Heidelberg. ISBN: 0387063730.

Rantzer, A. (2020). “Minimax adaptive control for state matrix with unknown sign”. In: *Proc. 21st IFAC World Congr.* Vol. 53. 2, pp. 58–62. DOI: 10.1016/j.ifacol.2020.12.049.

Rantzer, A. (2021). “Minimax adaptive control for a finite set of linear systems”. In: *Proc. 3th Annu. Learning Dyn. Control Conf.* Vol. 144. Corrected version: https://arxiv.org/abs/2011.10814, pp. 893–904. URL: https://proceedings.mlr.press/v144/rantzer21a.html.

Recht, B. (2019). “A tour of reinforcement learning: the view from continuous control”. *Annu. Rev. Control, Robot. Auton. Syst.* **2**, pp. 253–279. ISSN: 2573-5144. DOI: 10.1146/annurev-control-053018-023825.

Renganathan, V., A. Rantzer, and O. Kjellqvist (2024). “Distributed adaptive control for uncertain networks”. In: *Proc. 2024 Eur. Control Conf.* Pp. 1789–1794. DOI: 10.23919/ECC64448.2024.10591151.

Safonov, M. G. (1978). In: *Proc. 16th Allerton Conf. Commun. Control Comput.* Pp. 451–460.

Safonov, M. G. (1981). “Stability margins of diagonally perturbed multivariable feedback systems”. In: *Proc. 20th IEEE Conf. Decis. Control*, pp. 1472–1478. DOI: 10.1109/CDC.1981.269503.

Safonov, M. G. (2012). “Origins of robust control: early history and future speculations”. *Annu. Rev. Control* **36**:2, pp. 173–181. ISSN: 1367-5788. DOI: 10.1016/j.arcontrol.2012.09.001.

Schaft, A. van der (1992). “$\mathcal{L}_2$-gain analysis of nonlinear systems and nonlinear state-feedback $\mathcal{H}_\infty$ control”. *IEEE Trans. Autom. Control* **37**:6, pp. 770–784. DOI: 10.1109/9.256331.

Sellke, M. (2023). “Chasing convex bodies optimally”. In: Eldan, R. et al. (Eds.). *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2020–2022*. Springer International Publishing, Cham, pp. 313–335. ISBN: 978-3-031-26300-2. DOI: 10.1007/978-3-031-26300-2_12.

Shamir, O. (2013). “On the complexity of bandit and derivative-free stochastic convex optimization”. In: Shalev-Shwartz, S. et al. (Eds.). *Proc. 24th Conf. Learn. Theory.* Vol. 30. Proceedings of Machine Learning Research. PMLR, Princeton, NJ, USA, pp. 3–24. URL: https://proceedings.mlr.press/v30/Shamir13.html.

Shamma, J. (1994). “Robust stability with time-varying structured uncertainty”. *IEEE Trans. Autom. Control* **39**:4, pp. 714–724. DOI: 10.1109/9.286248.

Shin, J., T. A. Badgwell, K.-H. Liu, and J. H. Lee (2019). "Reinforcement learning—overview of recent progress and implications for process control". *Comput. Chem. Eng.* **127**, pp. 282–294. ISSN: 0098-1354. DOI: 10.1016/j.compchemeng.2019.05.029.

Silvestre, D., P. Rosa, and C. Silvestre (2021). "Distinguishability of discrete-time linear systems". *Int. J. Robust Nonlinear Control* **31**:5, pp. 1452–1478. DOI: 10.1002/rnc.5367.

Simchowitz, M., K. Singh, and E. Hazan (2020). "Improper learning for non-stochastic control". In: Abernethy, J. et al. (Eds.). *Proc. 33th Conf. Learn. Theory*. Vol. 125. PMLR, pp. 3320–3436. URL: https://proceedings.mlr.press/v125/simchowitz20a.html.

Strässer, R., J. Berberich, and F. Allgöwer (2021). "Data-driven control of nonlinear systems: beyond polynomial dynamics". In: *Proc. 60th IEEE Conf. Decis. Control*, pp. 4344–4351. DOI: 10.1109/CDC45484.2021.9683211.

Striebel, C. (1965). "Sufficient statistics in the optimum control of stochastic systems". *J. Math. Anal. Appl.* **12**:3, pp. 576–592. ISSN: 0022-247X. DOI: 10.1016/0022-247X(65)90027-2.

Tabuada, P., W.-L. Ma, J. Grizzle, and A. D. Ames (2017). "Proc. 56th IEEE conf. decis. control". In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pp. 6265–6270. DOI: 10.1109/CDC.2017.8264603.

Tadmor, G. (1993). "The standard $\mathcal{H}_\infty$ problem and the maximum principle: the general linear case". *SIAM J. Control Optim.* **31**, pp. 813–846.

Tsiamis, A., I. Ziemann, N. Matni, and G. J. Pappas (2023). "Statistical learning theory for control: a finite-sample perspective". *IEEE Control Syst. Mag.* **43**:6, pp. 67–97. DOI: 10.1109/MCS.2023.3310345.

Vinnicombe, G. (2004). "Examples and counterexamples in finite ℓ_2-gain adaptive control". In: *16th Int. Symp. Mathematical Theory Networks Systems*. Leuven, Belgium.

Wang, Y.-S., N. Matni, and J. C. Doyle (2018). "Separable and localized system-level synthesis for large-scale systems". *IEEE Trans. Autom. Control* **63**:12, pp. 4234–4249.

Willems, J. C., P. Rapisarda, I. Markovsky, and B. L. De Moor (2005). "A note on persistency of excitation". *Syst. Control Lett.* **54**:4, pp. 325–329. ISSN: 0167-6911. DOI: 10.1016/j.sysconle.2004.09.003.

Witsenhausen, H. (1966). *Minimax Control of Uncertain Systems*. Massachusetts Institute of Technology. URL: https://books.google.se/books?id=PCxMnQEACAAJ.

Yang, H. and S. Li (2015). "A data-driven predictive controller design based on reduced hankel matrix". In: *Proc. 10th Asian Control Conf.* Pp. 1–7. DOI: 10.1109/ASCC.2015.7244723.

Young, P. and J. Doyle (1990). "Computation of mu with real and complex uncertainties". In: *Proc. 29th IEEE Conf. Decis. Control*, 1230–1235 vol.3. DOI: 10.1109/CDC.1990.203804.

Young, P. M., M. P. Newlin, and J. C. Doyle (1992). "Practical computation of the mixed μ problem". In: *Proc. 1992 IEEE Amer. Control Conf.* Pp. 2190–2194. DOI: 10.23919/ACC.1992.4792521.

Yu, J., V. Gupta, and A. Wierman (2023). "Online adversarial stabilization of unknown linear time-varying systems". In: *Proc. 62nd IEEE Conf. Decis. Control*, pp. 8320–8327. DOI: 10.1109/CDC49753.2023.10383849.

Yu, J., Y.-S. Wang, and J. Anderson (2021). "Localized and distributed $\mathcal{H}_2$ state feedback control". In: *Proc. 2021 IEEE Amer. Control Conf.* Pp. 2732–2738. DOI: 10.23919/ACC50511.2021.9483301.

Zames, G. (1966). "On the input-output stability of time-varying nonlinear feedback systems–part ii: conditions involving circles in the frequency plane and sector nonlinearities". *IEEE Trans. Autom. Control* **11**:3, pp. 465–476. DOI: 10.1109/TAC.1966.1098356.

Zames, G. (1981). "Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses". *IEEE Trans. Autom. Control* **26**:2, pp. 301–320. DOI: 10.1109/TAC.1981.1102603.

Zhou, K. and J. C. Doyle (1998). *Essentials of Robust Control.* Prentice-Hall, Upper Saddle River, New Jersey, USA.

Paper I

Minimax Adaptive Estimation for Finite Sets of Linear Systems

Olle Kjellqvist **Anders Rantzer**

Abstract

For linear time-invariant systems with uncertain parameters belonging to a finite set, we present a purely deterministic approach to multiple-model estimation and propose an algorithm based on the minimax criterion using constrained quadratic programming. The estimator tends to learn the dynamics of the system, and once the uncertain parameters have been sufficiently estimated, the estimator behaves like a standard Kalman filter.

1. Introduction

1.1 Problem statement

In this article, we consider output prediction for linear systems of the form

$$\begin{aligned} x_{t+1} &= F x_t + G u_t + w_t \\ y_t &= H x_t + v_t, \qquad 0 \le t \le N-1, \end{aligned} \tag{1}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^p$ and $y_t \in \mathbb{R}^m$ are the states and the measured input and output at time-step t, respectively. $w_t \in \mathbb{R}^n$ and $v_t \in \mathbb{R}^m$ are unmeasured process disturbance and measurement noise. The model, (F, H, G) is fixed but unknown, belonging to some finite set

$$\{(F_1, H_1, G_1), \cdots, (F_K, H_K, G_K)\}.$$

consiting of of triplets of real-valued matrices. In particular, we are interested in strictly causal estimation of y_N, such that the gain from disturbance trajectories $(w_t, v_t)_{t=0}^{N-1}$ to pointwise estimation error $(y_N - Hx_N)$ in some weigthed ℓ_2-norm is bounded by a constant $\gamma_N > 0$. This means that given positive definite matrices $P_0 \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times m}$ and $Q \in \mathbb{R}^{n\times n}$ and a nominal value of the initial state, $\hat{x}_0$,

$$\frac{|\hat{y}_N - Hx_N|^2}{|x_0 - \hat{x}_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1}\left(|w_t|^2_{Q^{-1}} + |v_t|^2_{R^{-1}}\right)} \le \gamma_N^2, \tag{2}$$

should hold for all disturbances and models compatible with the measurement history $(y_t, u_t)_{t=0}^{N-1}$. This approach is different from the Bayesian approach to filtering where one takes the conditional expectation as the estimate $\hat{y}_N$. The interest in worst-case gain is motivated by robust feedback-control from estimates. In such settings instability or lack of performance due to model errors is a larger concern than robustness to outliers.

1.2 Background

Simultaneous estimation of states and parameters in linear systems is a bilinear estimation problem. The Maximum-likelihood approach leads to estimates which cannot be put in recursive form and must be obtained by iteration [Bar-Shalom, 1972]. A recursive method can be obtained by parametrizing the dynamical equations and the observer and learning the parameters using the *sequential prediction error* approach. Alternatively, one can augment the state vector with the uncertain parameters and apply nonlinear filtering methods such as the Extended Kalman filter [Goodwin and Sin, 1984]. Unfortunately, optimality guarantees for such methods are difficult to obtain. One exception is when the system can be modeled as a finite set of linear systems and the noise is Gaussian, then the Maximum-likelihood estimates can be put on a recursive form [Crassidis and Junkins, 2011].

Solutions based on the multiple-model approach have been tremendously successful in modeling and estimating complex engineering systems. In essence, it consists of two parts: 1) design simpler models for a finite set of possible operating regimes. 2) Run a filter for each model and cleverly combine the estimates. Multiple-model adaptive estimation has been around since the '60s [Magill, 1965; Lainiotis, 1976] and has been an active research field since. The estimation approach easily extends to systems where the active model can switch (hybrid systems) by matching a Kalman filter with each possible trajectory. In that case, the number of filters will grow exponentially, which has sparked research into more efficient methods. Notable numerically tractable and suboptimal algorithms for estimation in hybrid systems are the Generalized Pseudo Bayesian [Ackerson and Fu, 1970; Chang and Athans, 1978], and the Interacting Multiple Model [Blom and Bar-Shalom, 1988]. The algorithms have been coupled with extended and unscented Kalman filters to deal with non-linear systems [Akca and Efe, 2019], and [Xiong et al., 2015] studied robustness to identification error. In [Ronghua et al., 2008], the authors pointed out that methods based on Kalman filters are sensitive to noise distributions and proposed an Interactive Multiple Model algorithm based on particle filters to handle non-Gaussian noise at the expense of a 100 fold increase in computation. Recently, machine-learning approaches to classification have been combined with the Interacting Multiple Model estimator [Li et al., 2021; Deng et al., 2020] and showed improved accuracy in simulations.

The Bayesian approach to the *Multiple-model* estimation problem involves assigning probability distributions to disturbances (w_t, v_t) and models (F, G, H). The estimate is taken as the expected value of y_N conditioned on past measurements. If the disturbances are zero-mean and Gaussian, then the conditional expectation can be computed as the weighted average of Kalman filter estimates (one for each model), weighted by the conditional probability that its model is active.

It is evident in practice that the estimator's performance depends on the quality of the model set. The models must be distinguishable using measured signals, and the models should accurately describe the operating regimes. Since the estimates can be susceptible to non-Gaussian noise, it is surprising that deterministic approaches similar to those studied by the control community in the '80s and '90s have gathered little attention. Recent progress to minimax adaptive control of linear systems with uncertain parameters belonging to a finite set [Rantzer, 2021] under the assumption of *perfect measurements* has inspired this research into compatible estimation techniques.

1.3 Contribution

In this paper, we formulate the multiple-model estimation problem as a deterministic, two-player dynamic game. In particular, this formulation allows for online computation of the worst-case gain from disturbances to estimation error and tractable synthesis of suboptimal estimators that minimize the worst-case gain. Deterministic dynamic games have played a key role in solving and understanding $\mathcal{H}_\infty$ filtering [Shen and Deng, 1997; Basar and Bernhard, 2008]; our goal in this work has been to take a first step towards extending the advantages of that framework to the multiple model setting.

1.4 Outline

The outline is as follows: First, we introduce notation in Section 2, then we introduce minimax multiple-model filtering and the main results in Section 3. In Section 4, we present a simplified form for time-invariant systems. We illustrate the theory through a numerical example in Section 5. Section 6 contains concluding remarks, and supporting lemmata are given in the Appendix.

2. Notation

The set of $n \times m$-dimensional matrices with real coefficients is denoted $\mathbb{R}^{n\times m}$. The transpose of a matrix A is denoted $A^\top$. For a symmetric matrix $A \in \mathbb{R}^{n\times n}$, we write $A \succ (\succeq) 0$ to say that A is positive (semi)definite. Given $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$, $|x|_A^2 := x^\top A x$. For a vector $x_t \in \mathbb{R}^n$ we denote the sequence of such vectors up to time t by $x_{[0:t]} := (x_k)_{k=0}^t$.

3. Minimax multiple model filtering

In contrast to the Bayesian approach, our approach is fully deterministic; similarly to [Shen and Deng, 1997; Basar and Bernhard, 2008], we do not make explicit assumptions on the distribution of the noise trajectories $w_{[0:t]}$ and $v_{[0:t]}$. We will instead construct a two-player dynamic game between a minimizing player that chooses the estimate, and a maximizing player that chooses dynamics and disturbances. Recall that we are interested in characterizing an estimator $\hat{y}_N$ such that the gain from disturbances to the pointwise estimation error is bounded by γ_N. I.e., (2) holds for all disturbances consistent with (1) and the data $(y_{[0:N-1]}, u_{[0:N-1]})$. Since the disturbances are unknown, we cannot evaluate (2) directly. However, define

$$J_N(y_{[0:N-1]}, u_{[0:N-1]}, \hat{y}_N) := \sup_{x_0, w_{[0:N-1]}, v_{[0:N-1]}, (F,G,H)} \Bigg\{ |\hat{y}_N - Hx_N|^2 - \gamma_N^2 \left(|x_0 - \hat{x}_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1} \left(|w_t|^2_{Q^{-1}} + |v_t|^2_{R^{-1}} \right) \right) \Bigg\}, \quad (3)$$

where the maximization is performed subject to the constraints (1). Then (2) holds if and only if

$$J_N(y_{[0:N-1]}, u_{[0:N-1]}, \hat{y}_N) \leq 0.$$

In this setting, $w_t = x_{t+1} - Fx_t - Gu_t$ and $v_t = y_t - Hx_t$ are uniquely determined by the states, the measurements and the active model. Inserting into (3), we get

$$J_N(y_{[0:N-1]}, u_{[0:N-1]}, \hat{y}_N) = \sup_{x_{[0:N]}, (F,G,H)} \Bigg\{ |\hat{y}_N - Hx_N|^2 - \gamma_N^2 |x_0 - \hat{x}_0|^2_{P_0^{-1}} - \gamma_N^2 \sum_{t=0}^{N-1} \left(|x_{t+1} - Fx_t - Gu_t|^2_{Q^{-1}} + |y_t - Hx_t|^2_{R^{-1}} \right) \Bigg\}. \quad (4)$$

We will call an estimator $\hat{y}_N^\star$ a minimax estimator if

$$\inf_{\hat{y}_N} J_N(y_{[0:N-1]}, u_{[0:N-1]}, \hat{y}_N) = J_N(y_{[0:N-1]}, u_{[0:N-1]}, \hat{y}_N^\star) =: J_N^\star(y_{[0:N-1]}, u_{[0:N-1]}), \quad (5)$$

holds, where $\hat{y}_N$ are functions of past data $y_{[0:N-1]}$ and $u_{[0:N-1]}$. This constitutes a two-player dynamic game and would be linear quadratic if not for the model being chosen by the maximizing player. The intuition behind (5) makes sense in the following way. The minimizing player is penalized for deviating from the true (noiseless) output, and the maximizing player is penalized for selecting a model which requires large disturbances w and v to be compatible with the data. As N increases, the penalty for selecting a model different from the truth grows too large, resulting in a learning mechanism. It turns out that the cost associated with the disturbance trajectories required to explain each model corresponds to the accumulated prediction errors from a corresponding Kalman filter and that the minimax estimate is a weighted interpolation between the Kalman filter estimates.

Theorem 1
Consider matrices $F_1, \dots, F_K \in \mathbb{R}^{n\times n}$, $H_1, \dots, H_K \in \mathbb{R}^{m\times n}$, $G_1, \dots, G_K \in \mathbb{R}^{n\times p}$ *and positive definite* $Q, P_0 \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times m}$. *Define* $P_{t,i}$ *according to*

$$\begin{aligned} P_{0,i} &= P_0 \\ P_{t+1,i} &= Q + F_i(P_{t,i} - P_{t,i}H_i^\top(R + H_iP_{t,i}H_i^\top)^{-1}H_iP_{t,i})F_i^\top, \end{aligned}$$

and assume that $H_iP_{N,i}H_i^\top \prec \gamma_N^2 I$. *Then the cost* (4) *is equivalent to*

$$J_N(y_{[0:N-1]}, u_{[0:N-1]}, \hat{y}_N) = \max_i \left\{ |\hat{y}_N - H_i\breve{x}_{N,i}|^2_{(I-\gamma_N^{-2}H_iP_{N,i}H_i^\top)^{-1}} - \gamma_N^2 c_{N,i} \right\}. \quad (6)$$

$\breve{x}_{N,i}$ *is the Kalman filter estimate of* x_N *using the ith model, and* $c_{N,i}$ *are generated according to*

$$\begin{aligned} \breve{x}_{0,i} &= x_0 \\ \breve{x}_{t+1,i} &= F_i\breve{x}_{t,i} + K_{t,i}(y_t - H_i\breve{x}_{t,i}) + G_iu_t \\ K_{t,i} &= F_iP_{t,i}H_i^\top(R + H_iP_{t,i}H_i^\top)^{-1} \\ c_{0,i} &= 0 \\ c_{t+1,i} &= |H_i\breve{x}_{t,i} - y_t|^2_{(R+H_iP_{t,i}H_i^\top)^{-1}} + c_{t,i}. \end{aligned}$$

Proof. We will perform the maximization over state-trajectories in (4) in two steps. First over past trajectories ($x_{[0:N-1]}$) and then over the future state x_N[1]. The right-hand side of (4) becomes

$$\sup_{x_N,i}\Bigg\{|\hat{y}_N - H_i x_N|^2 - \gamma_N^2 \inf_{x_{[0:N-1]}}\Bigg\{|x_0-\hat{x}_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1}\left(|x_{t+1}-F_i x_t - G_i u_t|^2_{Q^{-1}} + |y_t - H_i x_t|^2_{R^{-1}}\right)\Bigg\}\Bigg\},$$

where $i = 1,\dots K$ is an index for the active model (F_i, H_i, G_i). Apply Lemma 1 to get

$$\begin{aligned} J_N(y_{[0:N-1]}, u_{[0:N-1]}, \hat{y}_N) &= \sup_{x_N,i}\left\{|\hat{y}_N - H_i x_N|^2 - \gamma_N^2 V_{N,i}((x_N, y_{[0:N-1]})\right\} \\ &= \sup_{i,x_N}\left\{|\hat{y}_N - H_i x_N|^2 - \gamma_N^2\left(|x_N - \breve{x}_N|^2_{P_{N,i}^{-1}} + c_{N,i}\right)\right\}. \end{aligned}$$

For fix $\hat{y}_N$ and i, the assumption $H_i P_{N,i} H_i^\top \prec \gamma_N^2 I$ guarantees that we maximize a concave function of x_N and we apply Lemma 2 with $A = H_i,\ X = I,\ Y = P_{N,i}$ to conclude[2],

$$J_N(y_{[0:N-1]}, u_{[0:N-1]}, \hat{y}_N) = \max_i |\hat{y}_N - H_i \breve{x}_{N,i}|^2_{(I-\gamma_N^{-2}H_i P_{N,i} H_i^\top)^{-1}} - \gamma_N^2 c_{N,i}.$$

□

Remark 1
Theorem 1 holds also for time-varying systems, if F_i and H_i are replaced by $F_{t,i}$ and $H_{t,i}$. Further, P_0, Q and R can be time-varying and differ between models.

Remark 2
Equation (6) is monotonically increasing in γ_N and the smallest $\gamma_N^\star$ such that $J_N(y_{[0:N-1]}, u_{[0:N-1]}, \hat{y}_N) \le 0$ can be found efficiently through bisection.

The below Corollary follows from Theorem 1 and describes how to compute the minimax estimator as a convex quadratic program.

Corollary 1
With assumptions as in Theorem 1, consider the convex program

$$\begin{aligned} &\underset{\hat{y}_N,t}{\textit{minimize}} && t \\ &\textit{subject to:} && |\hat{y}_N - H_i\breve{x}_{N,i}|^2_{(I-\gamma_N^{-2}H_i P_{N,i} H_i^\top)^{-1}} - \gamma_N^2 c_{N,i} \le t \\ &&& \forall i = 1\dots K. \end{aligned}$$

[1] $\max_{x_{[0:N]}}\{\dots\} = \max_{x_N}\left\{\max_{x_{[0:N-1]}}\{\dots\}\right\}$.

[2] The maximizing argument is given by $x_N^\star(\hat{y}_N, i) = (H_i^\top H_i - \gamma_N^2 P_{N,i}^{-1})^{-1}(H_i^\top \hat{y}_N - P_{N,i}^{-1}\gamma_N^2 \breve{x}_{N,i})$

The minimizing argument $\hat{y}_N^\star$ satisfies (5).

REMARK 3
If the model set is a singleton, then $\hat{y}_N^\star = Hx_N^\star = H\breve{x}_N$ is the estimate generated by the Kalman filter, which is a well known result [Basar and Bernhard, 2008].

3.1 On $c_{N,i}$ and the relation to conditional probability.

It is known (see for instance [Crassidis and Junkins, 2011]) that if w_t and v_t are uncorrelated Gaussian white noise with covariances Q and R, the conditional probability that the measured output $y_{[0:N]}$ has been generated by the model (F_i, G_i, H_i) and the input $u_{[0:N]}$ can be expressed as

$$p(i|y_{[0:N]}, u_{[0:N]}) = \frac{\alpha_N e^{-|y_N - H_i\breve{x}_{N,i}|^2_{\tilde{R}_{N,i}}}}{\det(2\pi\tilde{R}_{N,i})^{1/2}} p(i|y_{[0:N-1]}, u_{[0:N-1]}).$$

α_N is some normalization constant independent of i, and

$$\tilde{R}_{N,i} = R + H_i P_{N,i} H_i^\top,$$

with $P_{N,i}$ as in Theorem 1. Taking $c_{N,i}$ as in Theorem 1 we see that the conditional probability is proportional to $e^{-c_{N+1,i}}$,

$$p(i|y_{[0:N-1]}, u_{[0:N-1]}) \propto e^{-c_{N+1,i}} \prod_{t=1}^{N} \det(2\pi\tilde{R}_{t,i})^{-1/2}.$$

4. Stationary solution

For a set of time-invariant systems, we summarize a simple version of the filter in the below theorem.

THEOREM 2
Consider matrices $F_1, \ldots, F_K \in \mathbb{R}^{n\times n}$, $H_1, \ldots, H_K \in \mathbb{R}^{m\times n}$ and positive definite $Q, P_0 \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{m\times m}$. Assume that the algebraic Riccati equations

$$P_i = Q + F_i(P_i - P_i H_i^\top (R + H_i P_i H_i^\top)^{-1} H_i P_i) F_i^\top,$$

have solutions $H_i P_i H_i^\top \prec \gamma_N^2 I$. Then a minimax strategy $\hat{y}_N^\star$ for the game defined by

$$\min_{\hat{y}_N} \max_{x_{[0:N]}, i} \Bigg\{ |\hat{y}_N - H_i x_N|^2 - \gamma_N^2 |x_0 - \hat{x}_0|^2_{P_i^{-1}} - \gamma_N^2 \sum_{t=0}^{N-1} \Big(|x_{t+1} - F_i x_t - G_i u_t|^2_{Q^{-1}} + |y_t - H_i x_t|^2_{R^{-1}} \Big) \Bigg\},$$

and (1)*, is the minimizing argument of*

$$\min_{\hat{y}_N} \max_i \left\{ |\hat{y}_N - H_i \breve{x}_{N,i}|^2_{(I-\gamma_N^{-2} H_i P_i H_i^\top)^{-1}} - \gamma_N^2 c_{N,i} \right\}.$$

$\breve{x}_{N,i}$ *is the Kalman filter estimate of* x_N *using the* i*th model, and* $c_{N,i}$ *are generated according to*

$$\begin{aligned}
\breve{x}_{0,i} &= x_0 \\
\breve{x}_{t+1,i} &= F_i \breve{x}_{t,i} + K_i(y_t - H_i \breve{x}_{t,i}) + G_i u_t \\
K_i &= F_i P_i H_i^\top (R + H_i P_i H_i^\top)^{-1} \\
c_{0,i} &= 0 \\
c_{t+1,i} &= |H_i \breve{x}_{t,i} - y_t|^2_{(R+H_i P_i H_i^\top)^{-1}} + c_{t,i}.
\end{aligned}$$

Proof. This is a special case of Theorem 1, by replacing P_0 with P_i. □

5. Example

In this example, we compare a minimax estimator synthesized using Corollary 1, bisecting over γ_N, to find the estimator $\hat{y}_N^\star$ such that (2) is satisfied for the smallest possible γ_N. We compare this to a Bayesian multiple-model estimator [Crassidis and Junkins, 2011] and calculate the corresponding bound γ_N using Theorem 1 and bisection. Consider the uncertain linear system

$$\begin{aligned} x_{t+1} &= F x_t + w_t \\ y_t &= x_t + v_t \end{aligned}, \quad F \in \{-1, 1\}.$$

The weights in (2) are chosen to be $Q = R = P_0 = 1$. We generate data $y_{[0:N-1]}$ by simulating the system with $F = 1$ and w_t, v_t as independent Gaussian white noise with intensity 1. For $N = 5$ we find

$$P_{5,1} = P_{5,-1} = 1{,}62,$$

$$\breve{x}_{5,1} = -2{,}34, \quad \breve{x}_{5,-1} = 1{,}50,$$

$$c_{5,1} = 3{,}56, \quad c_{5,-1} = 8{,}11.$$

In Fig. 1, we illustrate (6) for $N = 5$ and the estimates. Note that $\gamma = 1{,}51$ can be guaranteed for the minimax estimator, but not the Bayesian. Fig. 2 contains a comparison between the smallest γ_N so that (2) can be guaranteed for the minimax estimator and the Bayesian estimator when $N = 1 \dots 20$.

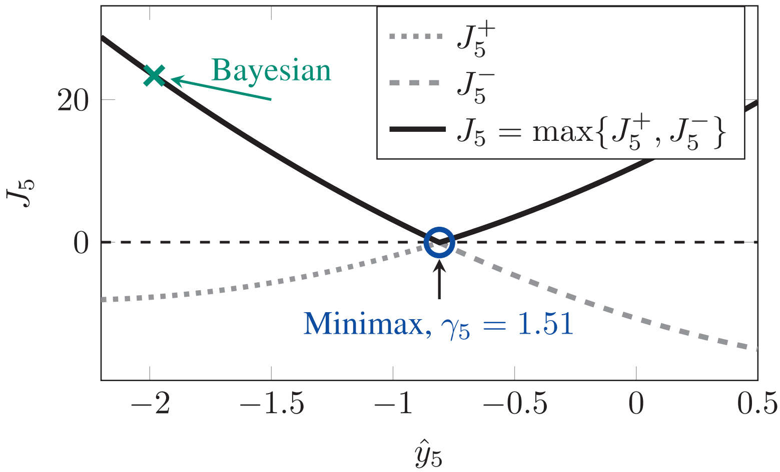


Figure 1. Illustration of the optimization problem (6) for $N = 5$, together with the minimax solution and the one given by a Bayesian multiple model estimator for $\gamma_N = 1.51$. The minimax estimate has a guaranteed worst-case gain bound from disturbances to observer error lower than 1.51, whereas the Bayesian estimator does not. Here $J_5^+ = |\hat{y}_5 - \breve{x}_{5,1}|^2_{(I-\gamma_5^{-2}P_{5,1})^{-1}} - c_{5,1}$ corresponds to $F = 1$, whereas J_5^- (defined similarly) corresponds to $F = -1$. $J_5 = J_5(y_{[0:5]}, 0, \hat{y}_5)$ is then equivalent to (6).

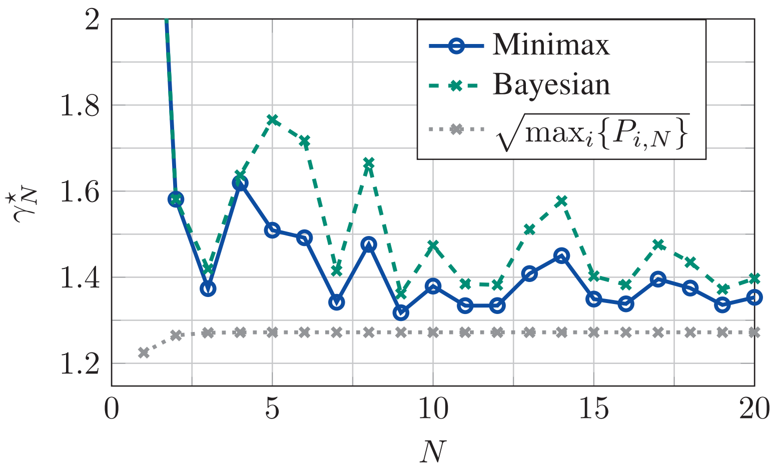


Figure 2. The smallest γ_N such that $J_N(y_{[0:N-1]}, 0, \hat{y}_N) \leq 0$ for the minimax estimator (blue) compared to the Bayesian multiple-model adaptive estimator (green) for one realization.

6. Conclusions

We stated the minimax criterion for output prediction, where the dynamics belong to a finite set of linear systems and proposed a minimax estimation strategy. The strategy can be implemented as a convex program, and the resulting estimate is a weighted interpolation of Kalman filter estimates. We showed in a numerical example how to apply the theoretical results to compute the worst-case gain from disturbances to error for any multi-model estimation algorithm online and how to generate estimates that minimize the said gain.

By running a minimax estimator in parallel to another estimator, we can measure the worst-case performance level of the other estimator. A large difference in performance levels indicates that the nominal estimator may be highly sensitive to errors in the noise model.

Predetermining the smallest achievable gain from disturbances to estimation errors is still an open research problem, that is, finding necessary and sufficient conditions such that

$$\sup_{y_{[0:N-1]}} J_N^\star(y_{[0:N-1]}, u_{[0:N-1]}) \leq 0.$$

In future work, we plan to develop a Multiple-model adaptive estimator with a prescribed ℓ_2-gain bound from disturbance to error and methods for infinite sets of linear systems.

Acknowledgment

We are grateful to our colleague Dr. Carolina Bergeling for comments that greatly improved the manuscript. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 834142 (ScalableControl).

Appendix — supporting lemmata

LEMMA 1
The cost function

$$V_{N,i}(x_N, y_{[0:N-1]}) = \min_{x_{[0:N-1]}} \Bigg\{ |x_0 - \hat{x}_0|^2_{P_0^{-1}} + \sum_{k=1}^{N-1} \left(|x_{t+1} - F_i x_t - G_i u_t|^2_{Q^{-1}} + |y_t - H_i x_t|^2_{R^{-1}} \right) \Bigg\} \quad (7)$$

under the dynamics (1), *is of the form*

$$V_{t,i}(x, y_{[0:t-1]}) = |x - \breve{x}_{t,i}|^2_{P_{t,i}} + c_{t,i},$$

where $P_{t,i}$ *and* $c_{t,i}$ *are generated as*

$$\begin{aligned} P_{0,i} &= P_0 \\ P_{t+1,i} &= Q + F_i P_{t,i} F_i^\top - F_i P_{t,i} H_i^\top (R + H_i P_{t,i} H_i^\top)^{-1} H_i P_{t,i} F_i^\top \\ \breve{x}_{0,i} &= x_0 \\ \breve{x}_{t+1,i} &= F_i \breve{x}_{t,i} + K_{t,i}(y_t - H_i \breve{x}_{t,i}) + G_i u_t \\ K_{t,i} &= F_i P_{t,i} H_i^\top (R + H_i P_{t,i} H_i^\top)^{-1} \\ c_{0,i} &= 0 \\ c_{t+1,i} &= |H_i \breve{x}_{t,i} - y_t|^2_{(R + H_i P_{t,i} H_i^\top)^{-1}} + c_{t,i}. \end{aligned}$$

Proof. The proof builds on forward dynamic programming [Cox, 1964], and is similar to one given in [Goodwin et al., 2005] but differ in the assumption that F_i is not invertible. Further, the constant terms $c_{t,i}$ are explicitly computed. The cost function V_N[3] can be computed recursively

$$V_1(x, y_{[0:0]}) = |x - x_0|^2_{P_0^{-1}} \tag{8}$$

$$V_{t+1}(x, y_{[0:t]}) = \min_\xi |x - F\xi - Gu_t|^2_{Q^{-1}} + |y_t - H\xi|^2_{R^{-1}} + V_t(\xi, y_{[0:t-1]}). \tag{9}$$

With a slight abuse of notation, we assume a solution of the form $V_t(x) = |x - \breve{x}_t|_{P_t^{-1}} + c_t$ and solve for the minimum

$$\begin{aligned} V_{t+1}(x) = \min_\xi\, & |x - Gu_t|^2_{Q^{-1}} + |\xi|^2_{F^\top Q^{-1} F + H^\top R^{-1} H + P_t^{-1}} \\ & - 2(F^\top Q^{-1}(x - Gu_t) + H^\top R^{-1} y_t + P_t^{-1} \breve{x}_t)^\top \xi + |y_t|^2_{R^{-1}} + |\breve{x}|_{P_t^{-1}}. \end{aligned}$$

Assume at this stage $S_t := F^\top Q^{-1} F + H^\top R^{-1} H + P_t^{-1} \succ 0$, then the minimizing $\xi^\star$ is a stationary point

$$\xi^\star = S_t^{-1}(F^\top Q^{-1}(x - Gu_t) + H^\top R^{-1} y_t + P_t^{-1} \breve{x}_t)$$

and the resulting partial cost

$$\begin{aligned} |x - \breve{x}_{t+1}|^2_{P_{t+1}^{-1}} + c_{t+1} = {} & |x - Gu_t|^2_{Q^{-1}} + |y_t|^2_{R^{-1}} + |\breve{x}_t|^2_{P_t^{-1}} \\ & - |F^\top Q^{-1}(x - Gu_t) + H^\top R^{-1} y_t + P_t^{-1} \breve{x}_t|^2_{S_t^{-1}} + c_t. \end{aligned} \tag{10}$$

[3] We relax the index i in this proof

Since this should hold for arbitrary x and

$$x - \breve{x}_{t+1} = (x - Gu_t) - (\breve{x}_{t+1} - Gu_t),$$

we get

$$\begin{aligned} P_{t+1}^{-1} &= Q^{-1} - Q^{-1} F S_t^{-1} F^\top Q^{-1} \\ \breve{x}_{t+1} - Gu_t &= P_{t+1} Q^{-1} F S_t^{-1} (H^\top R^{-1} y_t + P_t^{-1} \breve{x}_t) \end{aligned}$$

The expression for calculating P_{t+1} can be further simplified using the Woodbury identity,

$$\begin{aligned} P_{t+1}^{-1} &= (Q + F(H^\top R^{-1} H + P_t^{-1})^{-1} F^\top)^{-1} \\ P_{t+1} &= Q + F P_t F^\top - F P_t H^\top (R + H P_t H^\top)^{-1} H P_t F^\top, \end{aligned}$$

where we used the Woodbury matrix identity twice. Inserting these expressions into (10), applying the Woodbury matrix identity to $S_t^{-1} F^\top (Q - F S_t^{-1} F^\top)^{-1} S_t^{-1} + S_t^{-1} = (S_t - F^\top Q^{-1} F)^{-1} = (H^\top R^{-1} H + P_t^{-1})^{-1}$ gives

$$\begin{aligned} c_{t+1} &= -|H^\top R^{-1} y_t + P_t^{-1} \breve{x}_t|^2_{(H^\top R^{-1} H + P_t^{-1})^{-1}} + |y_t|^2_{R^{-1}} + |\breve{x}_t|^2_{P_t^{-1}} + c_t \\ &= |H\hat{x}_t - y_t|^2_{(R + H P_t H^\top)^{-1}} + c_t \end{aligned}$$

Next we show that $\breve{x}$ can be formulated as a state-observer

$$\begin{aligned} \breve{x}_{t+1} - Gu_t &= P_{t+1} Q^{-1} F S_t^{-1} (H^\top R^{-1} y_t + P_t^{-1} \breve{x}) \\ &= P_{t+1} Q^{-1} F S_t^{-1} H^\top R^{-1} (y_t - H\breve{x}_t) + \\ &\quad\quad P_{t+1} Q^{-1} F S_t^{-1} (H^\top R^{-1} H + P_t^{-1}) \breve{x}_t \end{aligned}$$

Use the matrix inversion lemma $(A + BCD)^{-1} BC = A^{-1} B (C + D A^{-1} B)^{-1}$.

$$\begin{aligned} P_{t+1} Q^{-1} F S_t^{-1} &= -(-Q^{-1} + Q^{-1} F S_t^{-1} F^\top Q^{-1})^{-1} Q^{-1} F S_t^{-1} \\ &= -(-Q^{-1})^{-1} (Q^{-1} F)(S_t - F^\top Q^{-1} F)^{-1} \\ &= F(H^\top R^{-1} H + P_t^{-1})^{-1}. \end{aligned}$$

Insert in to the previous expression and conclude

$$\breve{x}_{t+1} = F\breve{x}_t + K_t (y_t - H\breve{x}) + Gu_t,$$

where

$$K_t = F P_t H^\top (R + H P_t H^\top)^{-1} \qquad \square$$

Lemma 2
For $x \in \mathbb{R}^n$, v, $y \in \mathbb{R}^m$, a non-zero matrix $A \in \mathbb{R}^{n\times m}$, positive-definite matrices $X \in R^{n\times n}$ and $Y \in \mathbb{R}^{m\times m}$, and a positive real number $\gamma_N > 0$ such that

$$A^\top X^{-1}A - \gamma_N^2 Y^{-1} \prec 0,$$

it holds that

$$\max_v \left\{ |x - Av|^2_{X^{-1}} - \gamma_N^2 |y - v|^2_{Y^{-1}} \right\} = |x - Ay|^2_{(X-\gamma_N^{-2}AYA^\top)^{-1}}. \tag{11}$$

Proof. Expanding the left-hand side of (11) and equating the gradient with 0 we get

$$\begin{aligned}
&\max_v \left\{ |x - Av|^2_{X^{-1}} - \gamma_N^2 |y - v|^2_{Y^{-1}} \right\} \\
&= \max_v \left\{ |v|^2_{A^\top X^{-1}A-\gamma_N^2 Y} + |x|^2_{X^{-1}} - \gamma_N^2 |y|^2_{Y^{-1}} - 2v^\top (A^\top X^{-1}x - \gamma_N^2 Y^{-1})y \right\} \\
&= |x|^2_{X^{-1}} - \gamma_N^2 |y|^2_{Y^{-1}} - |A^\top X^{-1}x - \gamma_N^2 Y^{-1}y|_{(A^\top X^{-1}A-\gamma_N^2 Y^{-1})^{-1}} \\
&= |x|^2_{X^{-1}-X^{-1}A^\top (A^\top X^{-1}A-\gamma_N^2 Y^{-1})^{-1}A^\top X^{-1}} \\
&\qquad + |y|^2_{-\gamma_N^2 Y^{-1}-\gamma_N^2 Y^{-1}(A^\top X^{-1}A-\gamma_N^2 Y^{-1})^{-1}Y^{-1}\gamma_N^2} \\
&\qquad - 2x^\top X^{-1}A(A^\top X^{-1}A - \gamma_N^2 Y^{-1})^{-1}(-\gamma_N^2 Y^{-1})y \\
&= |x|^2_{(X-\gamma_N^{-2}AYA^\top)^{-1}} + |Ay|^2_{(X-\gamma_N^{-2}AYA^\top)^{-1}} - 2x^\top (X - \gamma_N^{-2}AYA^\top)^{-1}Ay \\
&= |x - Ay|^2_{(X-\gamma_N^{-2}AYA^\top)^{-1}}. \qquad \square
\end{aligned}$$

References

Ackerson, G. and K. Fu (1970). “On state estimation in switching environments”. *IEEE Trans. Autom. Control* **15**:1, pp. 10–17. DOI: 10.1109/TAC.1970.1099359.

Akca, A. and M. Ö. Efe (2019). “Multiple model kalman and particle filters and applications: a survey”. **52**:3, pp. 73–78. ISSN: 2405-8963. DOI: 10.1016/j.ifacol.2019.06.013.

Bar-Shalom, Y. (1972). “Optimal simultaneous state estimation and parameter identification in linear discrete-time systems”. *IEEE Trans. Autom. Control* **17**:3, pp. 308–319. DOI: 10.1109/TAC.1972.1100005.

Basar, T. and P. Bernhard (2008). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach*. 2nd ed. Birkhauser, Boston, MA, USA. DOI: 10.1007/978-0-8176-4757-5.

Blom, H. A. P. and Y. Bar-Shalom (1988). “The interacting multiple model algorithm for systems with markovian switching coefficients”. *IEEE Trans. Autom. Control* **33**:8, pp. 780–783. DOI: 10.1109/9.1299.

Chang, C. B. and M. Athans (1978). “State estimation for discrete systems with switching parameters”. *IEEE Trans. Aerosp. Electron. Syst.* **AES-14**:3, pp. 418–425. DOI: 10.1109/TAES.1978.308603.

Cox, H. (1964). “On the estimation of state variables and parameters for noisy dynamic systems”. *IEEE Trans. Autom. Control* **9**:1, pp. 5–12. DOI: 10.1109/TAC.1964.1105635.

Crassidis, J. L. and J. L. Junkins (2011). *Optimal Estimation of Dynamic Systems, Second Edition*. 2nd. Chapman & Hall/CRC, Boca Raton, FL, USA. ISBN: 1439839859.

Deng, L., D. Li, and R. Li (2020). “Improved IMM algorithm based on RNNs”. *J. Phys. Conf. Ser.* **1518**, p. 012055. DOI: 10.1088/1742-6596/1518/1/012055.

Goodwin, G. C., J. A. De Dona, and M. M. Seron (2005). *Constrained Control and Estimation — An Optimization Approach*. Springer-Verlag, London, UK. DOI: 10.1007/b138145.

Goodwin, G. C. and K. S. Sin (1984). *Adaptive filtering prediction and control*. Prentice-Hall, Englewood Cliffs, N.J. ISBN: 013004069X.

Lainiotis, D. G. (1976). “Partitioning: a unifying framework for adaptive systems, i: estimation”. *Proc. IEEE* **64**:8, pp. 1126–1143. DOI: 10.1109/PROC.1976.10284.

Li, D., P. Zhang, and R. Li (2021). “Improved IMM algorithm based on XGBoost”. In: *J. Phys. Conf. Ser.* Vol. 1748, p. 032017. DOI: 10.1088/1742-6596/1748/3/032017.

Magill, D. (1965). "Optimal adaptive estimation of sampled stochastic processes". *IEEE Trans. Autom. Control* **10**:4, pp. 434–439. DOI: 10.1109/TAC.1965.1098191.

Rantzer, A. (2021). "Minimax adaptive control for a finite set of linear systems". In: *Proc. 3th Annu. Learning Dyn. Control Conf.* Vol. 144. Corrected version: https://arxiv.org/abs/2011.10814, pp. 893–904. URL: https://proceedings.mlr.press/v144/rantzer21a.html.

Ronghua, G., Q. Zheng, L. Xiangnan, and C. Junliang (2008). "Interacting multiple model particle-type filtering approaches to ground target tracking". *J. Comput.* **3**, pp. 23–30. DOI: 10.4304/jcp.3.7.23-30.

Shen, X. and L. Deng (1997). "Game theory approach to discrete H_∞ filter design". *IEEE Trans. Signal Process.* **45**:4, pp. 1092–1095.

Xiong, K., C. Wei, and L. Liu (2015). "Robust multiple model adaptive estimation for spacecraft autonomous navigation". *Aerosp. Sci. Technol.* **42**, pp. 249–258. ISSN: 1270-9638. DOI: 10.1016/j.ast.2015.01.021.

Paper II

Minimax Performance Limits for Multiple-Model Estimation

Olle Kjellqvist

Abstract

This article concerns the performance limits of strictly causal state estimation for linear systems with fixed, but uncertain, parameters belonging to a finite set. In particular, we provide upper and lower bounds on the smallest achievable gain from disturbances to the pointwise estimation error. The bounds rely on forward and backward Riccati recursions—one forward recursion for each feasible model and one backward recursion for each pair of feasible models. We give simple examples where the lower and upper bounds are tight.

1. Introduction

Multiple-model estimation is a valuable tool for state estimation of systems that operate in different modes, for problems involving unknown parameters, for dealing with systems subject to faults, and for target tracking. If the mode is known, one selects the filter corresponding to the current mode. Otherwise, one can use a bank of filters, one for each mode, and cleverly combine the estimates. The latter approach is precisely what is called multiple-model estimation.

Almost all of the literature assumes that the system is affected by stochastic noise and that good noise statistics are available. Unfortunately, many popular methods are sensitive to a mismatch between the assumed and actual noise statistics. This assumption limits the applicability of in control systems, where we often use simplified models and disguise the model mismatch as additive disturbances. These disturbances are sometimes poorly modeled by Gaussian noise, and the noise statistics are often unknown.

In this article, we consider the problem of predicting the state of a linear system with unknown but fixed parameters belonging to a finite set. We assume that the system is affected by disturbances but make no assumptions about the noise statistics. We study the *minimax performance level*, defined as the gain from disturbances to point-wise estimation error, and are concerned with bounding the optimal (smallest achievable) performance level. See Fig. 1 for an illustration of our problem.

1.1 Contributions

This author, and Rantzer, recently proposed an estimator that achieves the optimal performance level but the performance level itself was not characterized [Kjellqvist and Rantzer, 2022b]. The main contribution of this article is to extend the framework in [Kjellqvist and Rantzer, 2022b] with a method to compute upper and lower bounds of the optimal performance level. These bounds are computed offline, a priori, and depend on the pairwise interaction between candidate models.

1.2 Background

The idea of using multiple models to reduce uncertainty is prevalent in many fields. It has been used in adaptive estimation since the '60s [Magill, 1965], where it is called multiple-model estimation and in feedback control since the '70s [Athans et al., 1977], where it is called multiple-model adaptive control [Buchstaller and French, 2016], or supervisory control [Hespanha, 2001]. The concept has been known in machine learning at least since Dasarty and Sheela introduced the "Composite classifier system" in 1979 [Dasarathy and Sheela, 1979], and is commonly referred to as ensemble learning [Dietterich, 2000]. In the field of economics, the idea of multiple models is known as model averaging [Steel, 2020], and was popularized by the work of Bates and Granger [Bates and Granger, 1969].

The task usually falls into one of two categories: *model selection*, where the goal is to find the best performing model, or *model averaging*, where the goal is to use all the models to generate an estimate of some common quantity. In this article, the focus is on predicting the state in dynamical systems, which falls into the latter category.

When the model is known, the Kalman filter is a realization of many reasonable estimation strategies. The minimum variance estimate, the maximum-likelihood estimate, and the conditional expectation under white-noise assumptions [Anderson and Moore, 1979] all coincide with the estimate generated by the Kalman filter. The filter also has appealing deterministic interpretations as the minimum energy estimate [Willems, 2004; Buchstaller et al., 2020], and as Krener showed [Krener, 1980], it constitutes a minimax optimal estimate.

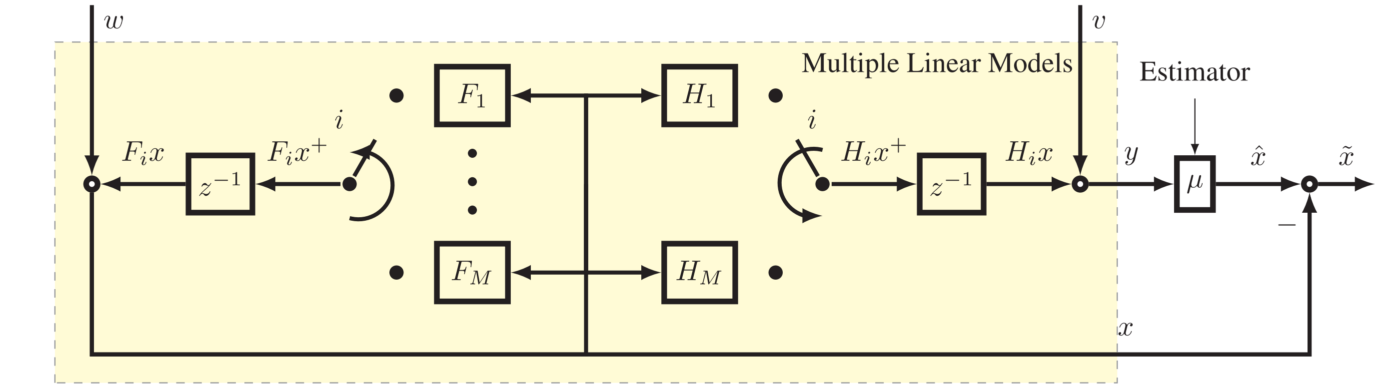


Figure 1. An illustration of the multiple-model estimation problem. A linear system with uncertain state-space matrices (F, H), driven by process noise w and measurement noise v, generates state and measurement sequences $x_{[0:N+1]}$ and $y_{[0:N]}$. The state transition matrix F and the measurement matrix H are unknown but belong to a finite set $(F_1, H_1), \ldots, (F_M, H_M)$. The goal of minimax estimation is to predict the next state x_{N+1} given the past measurements $y_{[0:N]}$ such that the gain from disturbance trajectories to the point-wise error $|\tilde{x}_{N+1}|^2$ is minimized. This paper concerns the optimal performance level (smallest achievable gain).

Interestingly, a minimax optimal estimate can be derived and computed without explicit knowledge of its *minimax performance level*, a property not shared with the $\mathcal{H}_\infty$-optimal estimate [Shen and Deng, 1997] and controller [Basar and Bernhard, 2008], which require knowledge of their performance levels. Tamer Başar showed that the optimal performance level can be obtained from the finite escape times of some related Riccati recursions [Başar, 1991].

In the case of multiple fixed models, the different estimation strategies give rise to different estimates[1]. The stochastic multiple-model approach to adaptive estimation was introduced in the '60s [Magill, 1965; Lainiotis, 1976] for linear systems with fixed, but unknown parameters, and has numerous applications in fault detection, state estimation and target tracking [Rong Li and Jilkov, 2005]. This estimation algorithm applies the Bayes rule recursively under white-noise assumptions on (w, v) and is well described in many textbooks like [Gustafsson, 2000; Crassidis and Junkins, 2011; Anderson and Moore, 1979]. The book [Anderson and Moore, 1979] also contains a convergence result, stating that given a certain distinguishability condition[2], the conditional probability for the active model generating the data converges to 1 as time goes to infinity. Vahid et al., [Hassani et al., 2009], proposed a minimum-energy condition for multiple-model estimation and proved a convergence result given a persistency-of-excitation-like criterion.

Multiple-model estimation has also been extended to the case with changing parameters, the case when i in Fig. 1 evolves on a Markov chain. One can, in principle, solve exactly for the Baysian average, but this is computationally intractable as the number of feasible trajectories grows exponentially with time. Instead, there exist sub-optimal algorithms that cleverly combine estimates at each time-step, compressing the feasible trajectories, like Blom's Interacting-Multiple-Model algorithm, [Blom and Bar-Shalom, 1988]. This idea was further generalized by Li and Bar-Shalom to the case when the model set varies with time, [Li and Bar-Shalom, 1996].

The work in this article is inspired by recent progress in *minimax adaptive control* [Rantzer, 2020; Rantzer, 2021; Kjellqvist and Rantzer, 2022a], and in a broader sense, the search for performance guarantees in learning-based control and identification [Matni et al., 2019; Mania et al., 2022].

1.3 Outline

The rest of this paper is organized as follows. We establish notation in Section 2. Section 3 contains the problem formulation and solution. Illustrative examples are in Section 4. We give conclusions and final remarks in Section 5.

[1] Except the maximum likelihood estimate under white-noise assumptions and a uniform prior over $\mathcal{M}$ coinciding with the minimum-energy estimate.

[2] Silvestre et al., [Silvestre et al., 2021], recently reexamined the distinguishability requirements from a multiple-model adaptive control perspective.

The proofs of the main results and supporting Lemmata are contained in the appendix.

2. Notation

The set of $n \times m$-dimensional matrices with real coefficients is denoted $\mathbb{R}^{n\times m}$. The transpose of a matrix A is denoted A^T. For a symmetric matrix $A \in \mathbb{R}^{n\times n}$, we write $A \succ (\succeq) 0$ to say that A is positive (semi)definite. The $n \times n$-dimensional identity matrix is denoted I_n, and the $n \times m$-dimensional zero matrix is denoted $0_{n\times m}$. Given $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times n}$, $|x|_A^2 := x^\mathsf{T} A x$. For a vector $x_t \in \mathbb{R}^n$ we denote the sequence of such vectors up to time t by $x_{[0:t]} := (x_k)_{k=0}^t$. For a sequence of square matrices $(A_i)_{i=1}^M$, we denote the corresponding block-diagonal matrix as $\operatorname{BlockDiag}(A_i)_{i=1}^M$.

3. Minimax performance limits

3.1 Problem statement

In this article, we consider strictly causal[3] state estimation for uncertain linear systems of the form

$$\begin{aligned} x_{t+1} &= F x_t + w_t, \qquad (F, H) \in \mathcal{M} \\ y_t &= H x_t + v_t, \quad 0 \le t \le N-1, \end{aligned} \tag{1}$$

where $x_t \in \mathbb{R}^n$, and $y_t \in \mathbb{R}^m$ are the states and the measured output at time t. $w_t \in \mathbb{R}^n$ and $v_t \in \mathbb{R}^m$ are unmeasured process disturbance and measurement noise. We employ a deterministic framework and make no assumptions on the distributions of w_t and v_t. Instead, they are adversarially chosen to maximize the objective of a related minimax problem that we will define shortly. The model, $(F, H) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{m\times n}$ is unknown but fixed, belonging to a (known) finite set

$$\mathcal{M} = \{(F_1, H_1), \ldots, (F_M, H_M)\}.$$

The state estimate at time N, $\hat{x}_N$, is generated by a causal estimator, μ, that depends on previous measurements but is unaware of the model, (F, H), and noise, (w, v), realizations,

$$\hat{x}_N = \mu(y_{N-1}, \ldots, y_0).$$

[3] The ideas in this paper extend to other information structures like filtering, k-step prediction, and smoothing, but they require some extra steps.

We are interested in describing the smallest γ_N, denoted $\gamma_N^\star$, such that the below expression has finite value.

$$J_N^\star(\hat{x}_0) := \inf_{\mu} \sup_{x_0, w_{[0:N-1]}, v_{[0:N-1]}, i} \Bigg\{ |x_N - \hat{x}_N|^2 - \gamma_N^2 \Bigg(|x_0 - \hat{x}_{0,i}|^2_{P_{0,i}} + \sum_{t=0}^{N-1} \Big[|w_t|^2_{Q_i^{-1}} + |v_t|^2_{R_i^{-1}} \Big] \Bigg) \Bigg\}, \quad (2)$$

where the trajectory $x_{[0:N]}$ in (2) is generated according to (1) with $(F_i, H_i) \in \mathcal{M}$. The problem set-up is a two-player game where the adversary picks the disturbance sequences $w_{[0:N-1]}$ and $v_{[0:N-1]}$, the initial state x_0, and the active model $i = 1, \dots, M$. The minimizing player picks the estimation policy μ. The matrices $Q_i \in \mathbb{R}^{n\times n}$ and $R_i \in \mathbb{R}^{m\times m}$ are positive definite matrices that weights the norms on w and v. The matrices $P_{0,i} \in \mathbb{R}^{n\times n}$ are positive definite and quantify the uncertainty in the estimates of the initial states $\hat{x}_{0,i}$.

3.2 Forward recursions

The forward recursions describe the worst-case disturbances consistent with the dynamics and an observed trajectory. They are also fundamental in constructing a minimax-optimal estimator $\mu_\star$. The recursions are equivalent to those of a Kalman filter of a system driven by zero-mean independent white noise sequences w_t and v_t with covariance matrices Q_i and R_i respectively,

$$\begin{aligned} K_{t,i} &= F_i P_{t,i} H_i^\top (R_i + H_i P_{t,i} H_i^\top)^{-1}, \\ P_{t+1,i} &= Q_i + F_i P_{t,i} F_i - K_{t,i}(R_i + H_i P_{t,i} H_i^\top) K_{t,i}^\top. \end{aligned} \quad (3)$$

The relation between the stochastic interpretation and our deterministic framework lies in that the least-squares estimate coincedes with the maximum-likelihood estimate under white-noise assumptions.

Remark 1
$P_{0,i}$ is a regularization term that penalizes deviations from an initial state estimate $\hat{x}_{0,i}$ and can be interpreted as the covariance of the initial estimate $\hat{x}_{0,i}$. It is practical to choose $P_{0,i}$ as the stationary solution to (3), and we will do so in the sequel to simplify the notation by removing the time index. The results in this section are valid for any positive semi-definite choice of $P_{0,i}$. However, the resulting observer dynamics will be time-varying. We leave it to the reader to reintroduce the dependence on t.

The solution, P_i, to the Riccati equation (3) quantifies the uncertainty of the state estimate given the observations $y_{0:t}$ and the model i and bounds the smallest achievable gain from below if the model is known. This is formalized in the following proposition, whose proof is in the Appendix.

Proposition 1
$\gamma_N \geq \gamma_N^\star$ only if $P_i \preceq \gamma_N^2 I$ for all $i = 1, \dots, M$.

In our previous work, [Kjellqvist and Rantzer, 2022b], we show how to construct the minimizing argument $\mu_\star$ of (2) in the case of output-prediction. The estimator uses the forward recursions (3) and requires a γ_N that fulfills Proposition 1. The following proposition shows how to construct a state predictor that is optimal for (2).

Proposition 2—Minimax multiple-model estimator
Given matrices $F_i \in \mathbb{R}^{n\times n}$ and $H_i \in \mathbb{R}^{m\times n}$, positive definite Q_i, $P_{0,i} \in \mathbb{R}^{n\times n}$ and $R_i \in \mathbb{R}^{m\times m}$ for $i = 1, \dots, M$. With $P_{0,i}$, P_i and K_i as the stationary solutions to (3),

$$\tilde{R}_i = R_i + H_i P_i H_i^\mathsf{T}, \tag{4}$$

a quantity γ_N such that $\gamma_N^2 I \succ P_i$, the below estimate achieves the infimum in (2):

$$\hat{x}_N^\star = \min_{\hat{x}_N} \max_i \left\{ |\hat{x}_N - \breve{x}_{N,i}|^2_{(I-\gamma^{-2}P_i)^{-1}} - \gamma_N^2 c_{N,i} \right\},$$

where $\breve{x}_{N,i} \in \mathbb{R}^n$ and $c_{N,i} \in \mathbb{R}$ are generated according to

$$\breve{x}_{0,i} = x_0, \quad c_{0,i} = 0, \tag{5a}$$

$$\breve{x}_{t+1,i} = F_i \breve{x}_{t,i} + K_i(y_t - H_i \breve{x}_{t,i}), \tag{5b}$$

$$c_{t+1,i} = |H_i \breve{x}_{t,i} - y_t|^2_{\tilde{R}_i^{-1}} + c_{t,i}. \tag{5c}$$

Proof. The proof is identical to that of Theorem 1 in [Kjellqvist and Rantzer, 2022b] but with the following modifications: $P_{0,i}$ is replaced by the stationary solution to (3) leading to $K_{t,i}$ and $P_{t,i}$ being replaced by K_i and P_i, the term $\hat{y}_N - H_i x_N$ is replaced by $\hat{x}_N - x_N$. □

3.3 Backward recursions

The backward recursions are similar to those of the linear-quadratic regulator and relate to the worst-case trajectories, in contrast to the forward recursions, which relate to the worst-case disturbances consistent with any given trajectory. They play no role in constructing the optimal estimator, $\mu_\star$, once a performance level γ has been found, but form the basis for *a priori* analysis of the optimal performance level $\gamma_N^\star$ that holds for any realization. Let

$$F^{ij} = \begin{bmatrix} F_i - K_i H_i & 0_{n\times n} \\ 0_{n\times n} & F_j - K_j H_j \end{bmatrix}, \quad K^{ij} = \begin{bmatrix} K_i \\ K_j \end{bmatrix}.$$

F_t^{ij} corresponds to the closed-loop of a pair (i, j) of Kalman filters with filter gains K_i and K_j as in (3). We will express the necessary and sufficient conditions using the following Riccati recursions. Given some symmetric matrix

$T_N^{ij} \in \mathbb{R}^{2n\times 2n}$ and $t = N-1, \ldots, 0$,

$$\begin{aligned} X_t^{ij} &= (K^{ij})^\mathsf{T} T_{t+1}^{ij} K^{ij} + (\tilde{R}_i^{-1} + \tilde{R}_j^{-1}), \\ L_t^{ij} &= (X_t^{ij})^{-1} \left((K^{ij})^\mathsf{T} T_{t+1}^{ij} F^{ij} - \begin{bmatrix} \tilde{R}_i^{-1} H_i & \tilde{R}_j^{-1} H_j \end{bmatrix} \right), \\ T_t^{ij} &= (F^{ij})^\mathsf{T} T_{t+1}^{ij} F^{ij} - (L_t^{ij})^\mathsf{T} X_t^{ij} L_t^{ij} + \begin{bmatrix} H_i^\mathsf{T} \tilde{R}_i^{-1} H_i & \\ & H_j^\mathsf{T} \tilde{R}_j^{-1} H_j \end{bmatrix}. \end{aligned} \tag{6}$$

For these recursions to be well-defined, the matrix X_t^{ij} must be invertible. The conditions for bounding $\gamma_N^\star$ are related to the positive definiteness of X_t^{ij} and are summarized in Theorems 1 and 2 below. The first concerns sufficient conditions and can be used to obtain upper bounds.

Theorem 1—Sufficient Condition
Given matrices $F_i \in \mathbb{R}^{n\times n}$ and $H_i \in \mathbb{R}^{m\times n}$, positive definite $Q_i \in \mathbb{R}^{n\times n}$ and $R_i \in \mathbb{R}^{m\times m}$ for $i = 1, \ldots, M$. Further, let $P_{0,i} = P_i$ and K_i be the stationary solutions to (3), and consider a quantity γ_N such that $\gamma_N^2 I \succ P_i$. Let $\underline{Q} \in \mathbb{R}^{n\times n}$ be a positive definite matrix such that $\underline{Q} \preceq I - \gamma_N^{-2} P_i$ for all $i = 1, \ldots, M$ and initialize the backward recursions (6) with the terminal state

$$T_N^{ij} = -\begin{bmatrix} \underline{Q}^{-1} & -\underline{Q}^{-1} \\ -\underline{Q}^{-1} & \underline{Q}^{-1} \end{bmatrix} / \gamma_N^2.$$

Assume that X_t^{ij} in (6) is negative definite for all i, j. Then $\gamma_N^\star \leq \gamma_N$ and

$$J_N^\star(\hat{x}_0) \leq \frac{1}{2} \max_{i,j} \left\{ -\gamma_N^2 \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix}^\mathsf{T} T_0^{ij} \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix} \right\}.$$

The second theorem concerns necessary conditions and helps obtaining lower bounds.

Theorem 2—Necessary Condition
Given matrices $F_i \in \mathbb{R}^{n\times n}$ and $H_i \in \mathbb{R}^{m\times n}$, positive definite $Q_i \in \mathbb{R}^{n\times n}$ and $R_i \in \mathbb{R}^{m\times m}$ for $i = 1, \ldots, M$. Further, let $P_{0,i} = P_i$ and K_i be the stationary solutions to (3), and consider a quantity γ_N such that $\gamma_N^2 I \succ P_i$. Initialize the backward recursion (6) with the terminal state

$$\begin{aligned} T_N^{ij} &= -\begin{bmatrix} Q^{ij} & -Q^{ij} \\ -Q^{ij} & Q^{ij} \end{bmatrix} / \gamma_N^2, \\ Q^{ij} &= (2I - \gamma_N^{-2}(P_{N,i} + P_{N,j}))^{-1}. \end{aligned}$$

If $X_t^{ij} \not\succ 0$ for some pair i, j and $0 \leq t \leq N-1$, then $\gamma_N^\star > \gamma_N$. If $X_t^{ij} \succ 0$, for all $t = 0, \ldots, N-1$ then

$$J_N^\star(\hat{x}_0) \geq \frac{1}{2} \max_{ij} \left\{ -\gamma_N^2 \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix}^\mathsf{T} T_0^{ij} \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix} \right\}.$$

Table 1. Parameters for the systems in Fig. 2a–2d. In all cases $Q_1 = Q_2 = R_1 = R_2 = 1$ and $P_{0,i}$ is the stationary solution to (3).

System	F_1	F_2	H_1	H_2	P_1	P_2
2a	1.1	1.1	1	-1	1.77	1.77
2b	0.9	0.9	1	-1	1.48	1.48
2c	0.7	0.9	1.5	1	1.16	1.48
2d	2	1	1	16	4.23	1.00

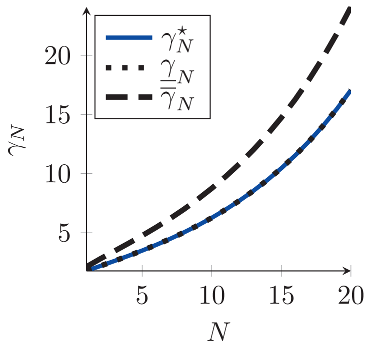
Remark 2
Theorems 1 and 2 give upper and lower bounds on $J_n^\star$ that can be translated upper and lower bounds on $\gamma_N^\star$ by bisecting over γ_N.

4. Examples

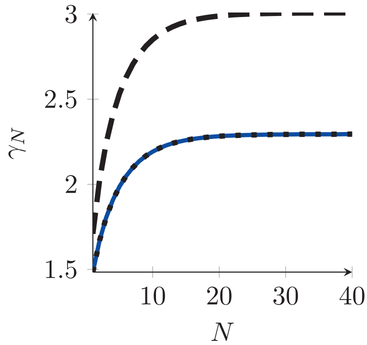
Figures 2a–2d show $\gamma_N^\star$ along with upper bounds, $\overline{\gamma}_N$, and lower bounds, $\underline{\gamma}_N$ for four different pairs of scalar systems, defined in Table 1. The optimal performance level, $\gamma_N^\star$, was computed using the construction in Appendix A.4, gridding the probability simplex $\{(\theta, 1-\theta) : \theta = 0, 10^{-3}, \dots, 1-10^{-3}, 1\}$ and the bounds were computed using Theorem 1 and 2, bisecting over γ to an accuracy of $\pm 10^{-3}$. The systems in Fig. 2a are unstable and *indistinguishable*, and the resulting optimal performance level $\gamma_N^\star$ grows exponentially in N. Fig. 2b is also *indistinguishable*, but here both systems are stable. The optimal performance level $\gamma_N^\star$ is bounded and is equal to the lower bound $\underline{\gamma}_N$. This is because the systems are BIBO stable, so picking $\hat{x}_N = 0$ results in an estimation error bounded by the disturbance's norm. Fig. 2c contains two stable systems that are *distinguishable*. The performance level $\gamma^\star$ is similar to the case where the system is known, and the bounds are close. $\gamma_N^\star$ is smaller than the other examples. Fig. 2d contains two unstable *distinguishable* systems. Here $\gamma_N^\star$ is bounded and approaches the upper bound $\overline{\gamma}_N$.

5. Conclusions

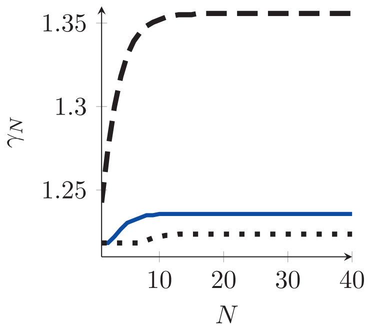
This article proposed a method to compute upper and lower bounds for the optimal minimax performance level for uncertain linear systems, where the uncertainty belongs to a finite set. The bounds are computed by evaluating the positive-definiteness of matrices appearing in coupled Riccati recursions. The performance level refines the notion of *distinguishability* in a priori analysis of the problem set-up for multiple-model estimation, and answers the question "To what extent can I guarantee the performance multiple-model



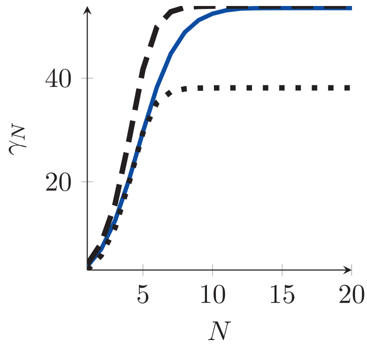
(a) Two *unstable* indistinguishable systems.



(b) Two *stable* indistinguishable systems.



(c) Two *stable* distinguishable systems.



(d) Two *unstable* distinguishable systems.

Figure 2. Numerically evaluated optimal performance levels, upper and lower bounds for the four system pairs considered in Section 4. Only stable *and or* distinguishable systems have bounded performance levels. In two pairs $\gamma_N^\star$ achieves the lower bound, and in Fig 2d it approaches the upper bound.

estimation applied to my problem?". Our experiments indicate that if similar output trajectories come from similar state trajectories, the gain is small. This agrees with the intution that such systems generate similar estimates, and that in order for these estimates to be poor, the disturbances must be large. However, if similar output trajectories come from different state trajectories, the state estimates will be different even for small disturbances, and as the optimal estimate is an interpolation of the estimates from the different models, the term $x - \hat{x}_N$ will be large even for small distrubances. The provided examples show that there are systems where the optimal performance level is equal to its lower bound, approaches its upper bound, and where neither bound is ever tight.

As with $\mathcal{H}_\infty$-control and estimation, the results are valid for any disturbance realization but are conservative if good disturbance statistics are available.

5.1 Future work

The numerical examples show that the bounds are tight for some systems, but not for others. The difference between the upper and lower bounds trivially bounds the conservativeness, but obtaining general conditions, and classifying systems where the bounds are tight, would enhance the practical utility of the results.

In this work, the system parameters F_i and H_i are assumed to be fixed. The extension to time-varying parameters is straightforward, but the extension to jump-linear systems is not. The reason is that the number of feasible parameter trajectories grows exponentially with time. There are heuristic ways of combining the Kalman filter estimates from different models, such as Blom's interacting-multiple-model estimator, [P. Blom, 1984].

The worst-case history can be losslessly compressed to quadratic functions, but the number of functions will grow exponentially in time. However, it is possible to upper bound the time-evolution of the worst-case data-consistent parameter realization by updating a constant number of quadratic functions, similar to how we combine many Kalman-filter estimates into one estimate in this paper. It would be interesting to exploit this bound to extend the results to jump-linear systems.

6. Acknowledgements

The author expresses sincere gratitude to colleagues Anders Rantzer and Venkatraman Renganathan for their stimulating and insightful discussions on the content and organization of this paper. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No

834142 (ScalableControl). The author is a member of the ELLIIT Strategic Research Area at Lund University.

Appendix

A. Proofs

This section proves Theorems 1 and 2. In doing so we obtain an expression that can be used to evaluate the value (2), but is computationally intractable for problems with uncertainties belonging to moderately-sized sets.

A.1 Proof strategy

We reparameterize the disturbance trajectory $(w_{[0:N-1]}, v_{[0:N-1]})$ in the state-output trajectory and the active model $(x_{[0:N-1]}, y_{[0:N-1]}, i)$. This reparameterization allows us to partially switch the order of the minimization and the maximization, as μ is a function of $y_{[0:N-1]}$, yielding a problem of the form $\max_{y_{[0:N-1]}} \min_\mu \max_{i,x_{[0:N]}}$. Previous work, [Kjellqvist and Rantzer, 2022b], shows how to maximize over $x_{[0:N]}$ using forward dynamic programming, resulting in the forward Riccati recursions (3).

We then reformulate the maximization over the feasible set to maximizing over its convex hull. This reformulation allows us to switch the order of minimizing with respect to μ and maximizing with respect to the model. The catch is that while the value is unchanged, the maximizing θ is not necessarily the same. As we are interested in the value, we can ignore this issue.

The inner minimization problem is unconstrained and convex-quadratic in the estimate $\hat{x}_N$, which has a closed-form solution. The maximization over the convex hull of the model set is then bounded from above and from below by a maximum over a finite number of functions that linear-quadratic regulator costs in $y_{[0:N-1]}$, which has a solution expressed by the backward Riccati recursion, (6).

A.2 Reparameterization

The disturbance w_t is uniquely determined by $F = F_i$ and (x_{t+1}, x_t), and v_t is uniquelly determined by $H = H_i$, y_t and x_t. As the maximizing player is aware of the dynamics, i, we can substitute $w_t = x_{t+1} - F_i x_t$ and $v_t =$

$y_t - H_i x_t$ into (2),

$$J_N^\star(\hat{x}_0) = \inf_{\mu} \sup_{x_{[0:N]}, y_{[0:N-1]}, i} \Bigg\{ |x_N - \hat{x}_N|^2 - \gamma^2 |x_0 - \hat{x}_{0,i}|^2_{P_{0,i}} - \gamma^2 \sum_{t=0}^{N-1} \left[|x_{t+1} - F_i x_t|^2_{Q_i^{-1}} + |y_t - H_i x_t|^2_{R_i^{-1}} \right] \Bigg\}. \quad (7)$$

Furthermore, as μ is a function of $y_{[0:N-1]}$, we can move the maximization over output trajectories outside of the minimization and minimize directly over the estimate $\hat{x}_N \in \mathbb{R}^n$. Consider the inner maximization over state trajectories, which is a function of the observations and estimates,

$$J_N^{\text{inner}}(y_{[0:N-1]}, \hat{x}_N, \hat{x}_0) = \sup_{x_{[0:N]}, i} \Bigg\{ |x_N - \hat{x}_N|^2 - \gamma_N^2 |x_0 - \hat{x}_{0,i}|^2_{P_{0,i}} - \gamma^2 \sum_{t=0}^{N-1} \left[|x_{t+1} - F_i x_t|^2_{Q_i^{-1}} + |y_t - H_i x_t|^2_{R_i^{-1}} \right] \Bigg\}. \quad (8)$$

Then (7) can be written as

$$J_N^\star(\hat{x}_0) = \sup_{y_{[0:N-1]}} \inf_{\hat{x}_N} J_N^{\text{inner}}(y_{[0:N-1]}, \hat{x}_N, \hat{x}_0). \quad (9)$$

A.3 Forward recursion

Following the proof of Theorem 1 in [Kjellqvist and Rantzer, 2022b], with $P_{0,i}$ as the stationary solution to (3), we see that the value inner optimization problem (8) is equal to

$$\sup_{i, x_N} \left\{ |\hat{x}_N - x_N|^2 - \gamma^2 \left(|x_N - \breve{x}_{N,i}|^2_{P_i^{-1}} + c_{N,i} \right) \right\} = \max_i \left\{ |\hat{x}_N - \breve{x}_{N,i}|^2_{(I - \gamma^{-2} P_i)^{-1}} - \gamma^2 c_{N,i} \right\}, \quad (10)$$

if $I \succ \gamma^{-2} P_i^{-1}$ for all i. The value is unbounded if $I \not\succeq \gamma^{-2} P_i^{-1}$ for some i, which proves Proposition 1. Proposition 2 shows how to compute $\breve{x}_{N,i}$, P_i and $c_{N,i}$ in (10)

A.4 Exact computations of $J_N^\star$

By substituting (10), we see that the value of (9) is equal to

$$\sup_{y_{[0:N-1]}} \min_{\hat{x}_N} \max_i \left\{ |\hat{x}_N - \breve{x}_{N,i}|^2_{(I - \gamma^{-2} P_i)^{-1}} - \gamma^2 c_{N,i} \right\}. \quad (11)$$

Maximizing over the finite set $\mathcal{M}$ in (11) is equivalent to optimizing for convex combinations over the probability simplex $\Theta = \{\theta \in \mathbb{R}^n : 0 \leq \theta_i \leq 1, \sum_{i=1}^{M} \theta_i = 1\}$. The equivalence is because the optimal value of a linear program over a simplex is located on a vertix. As (11) is convex in $\hat{x}$, the minimizing $\hat{x}$ can be bounded in terms of $\breve{x}_{N,i}$. The convex combination is affine in θ, so Von Neumann's minmax theorem applies and the value (11) is equal to

$$\sup_{\theta \in \Theta} \min_{\hat{x}} \left\{ \sum_{i=1}^{M} \theta_i \left(|\hat{x} - \breve{x}_{N,i}|^2_{Q_{N,i}} - \gamma^2 c_{N,i} \right) \right\},$$

where $Q_{N,i} = (I - \gamma^{-2} P_i)^{-1}$. Applying Lemma 3 to the inner minimization problem means that the value (11) is equal to

$$\sup_{y_{[0:N-1]},\theta} \left\{ \sum_{i=1}^{M} \theta_i \left(|\breve{x}_{i,i}|^2_{Q_{N,i}} - \gamma^2 c_{N,i} \right) - \left| \sum_{i=1}^{M} \theta_i Q_{N,i} \breve{x}_{N,i} \right|^2_{(\sum \theta_i Q_{N,i})^{-1}} \right\}. \quad (12)$$

For a fixed θ, this is a sequential quadratic optimization problem in y that can be solved using dynamic programming. In fact this can be reformulated into a standard linear-quadratic regulator problem, except that the terminal penalty is indefinite. This indefinite term will, for small values of γ_N, lead to a loss of concavity in $y_{[0:N-1]}$. This means that the value is unbounded, and $\gamma_N < \gamma_N^\star$. Larger values of γ_N will compensate for the indefinite term and ensure concavity in $y_{[0:N-1]}$. Testing for concavity amounts to evaluating whether $\mathbf{X}_t$ in (16) is positive definite for all t. If concavity in $y_{[0:N-1]}$ holds for all $\theta \in \Theta$, then the value is finite and $\gamma_N \geq \gamma_N^\star$. Define

$$\mathbf{F} \triangleq \text{BlockDiag}\left(\{F_i - K_i H_i\}_{i=1}^{M}\right)$$
$$\breve{\mathbf{x}}_\mathbf{t} \triangleq \begin{bmatrix} \breve{x}_{t,1}^\mathsf{T} & \cdots & \breve{x}_{t,M}^\mathsf{T} \end{bmatrix}^\mathsf{T}, \quad \mathbf{K} \triangleq \begin{bmatrix} K_1^\mathsf{T} & \cdots & K_M^\mathsf{T} \end{bmatrix}^\mathsf{T}.$$

Then, the multi-observer update (5b) becomes,

$$\breve{\mathbf{x}}_{t+1} = \mathbf{F}\breve{\mathbf{x}}_\mathbf{t} + \mathbf{K} y_t.$$

Further, let

$$\mathbf{Q}_N \triangleq \text{BlockDiag}\{\theta_i Q_{N,i}\}_{i=1}^{M} - \begin{bmatrix} \theta_1 Q_{N,1} \\ \vdots \\ \theta_M Q_{N,M} \end{bmatrix} \left((\sum \theta_i Q_{N,i})^{-1} \right)^{-1} \begin{bmatrix} \theta_1 Q_{N,1} \\ \vdots \\ \theta_M Q_{N,M} \end{bmatrix}^\mathsf{T}, \quad (13)$$

$$\begin{bmatrix} \mathbf{Q} & \mathbf{N}^\mathsf{T} \\ \mathbf{N} & \mathbf{R} \end{bmatrix} \triangleq \begin{bmatrix} \text{BlockDiag}\left(\{-H_i^\mathsf{T}\}_{i=1}^{M}\right) \\ I \quad \cdots \quad I \end{bmatrix}$$
$$\times \text{BlockDiag}\left(\{\theta_i \tilde{R}_i^{-1}\}_{i=1}^{M}\right) \begin{bmatrix} \text{BlockDiag}\left(\{-H_i^\mathsf{T}\}_{i=1}^{M}\right) \\ I \quad \cdots \quad I \end{bmatrix}^\mathsf{T}, \quad (14)$$

where $\times$ denotes standard matrix product. With

$$l(\theta, \breve{\mathbf{x}}_t, y_t) = \gamma_N^2 \left(|\breve{\mathbf{x}}_t|^2_{\mathbf{Q}} - 2y_t^\mathsf{T} \mathbf{N} \breve{\mathbf{x}}_\mathbf{t} + |y_t|^2_{\mathbf{R}} \right),$$

(9) becomes

$$J_N^\star(\hat{x}_0) = -\inf_\theta \underbrace{\inf_{y_{[0:N-1]}} \left\{ |\breve{\mathbf{x}}_\mathbf{t}|^2_{\mathbf{Q}_N} + \sum_{t=0}^{N-1} l(\theta, \breve{\mathbf{x}}_t, y_t) \right\}}_{\triangleq J_N(\theta, \hat{x}_0)}. \tag{15}$$

It is apparent that l is strictly convex in y_t. However, the terminal penalty matrix, $\mathbf{Q}_N$, is indefinite, which may cause (15) to lose convexity and become unbounded.

Remark 3

The stage cost is a convex combination of the Kalman filter residuals $l(\theta, \breve{\mathbf{x}}_t, y_t) = \gamma_N^2 \sum_{i=1}^M (\theta_i c_{t,i})$.

The Riccati recursions corresponding to the linear-quadratic regulator are well described in many textbooks, for instance in [Åström and Wittenmark, 1997, Chapter 11.2], and can be used to compute the value provided that $\theta \in \Theta$ fixed:

$$\begin{aligned} \mathbf{X}_t &= \mathbf{K}^\mathsf{T} \mathbf{T}_t \mathbf{K} + \mathbf{R}, \quad \mathbf{L}_t = \mathbf{X}_t^{-1} (\mathbf{K}^\mathsf{T} \mathbf{T}_t \mathbf{F} - \mathbf{N}) \\ \mathbf{T}_{t-1} &= \mathbf{F}^\mathsf{T} \mathbf{T}_t \mathbf{F} + \mathbf{Q} - \mathbf{L}_t^\mathsf{T} \mathbf{X}_t \mathbf{L}_t. \end{aligned} \tag{16}$$

The relationship between the solution to the above Riccati equations and the value of the game are summarized in the below lemma.

Lemma 1

Consider the backward Riccati equations above with terminal condition $\mathbf{T}_N = -\mathbf{Q}_N / \gamma_N^2$. *Let* $J_N^\star(x_0)$ *be the value of the game* (2) *and* $J_N(\theta, x_0)$ *be value of the inner, sequantial, optimization problem in* (15). *If* $\mathbf{X}_t \not\succ 0$ *for some* $\theta \in \Theta$, *then* $J_N^\star(x_0)$ *is unbounded. If* $\mathbf{X}_t \succ 0$ *for all* $\theta \in \Theta$ *then* $J_N(\theta, \hat{x}_0) = -\gamma_N^2 |\breve{\mathbf{x}}_0|^2_{\mathbf{T}_0}$, *and*

$$J_N^\star(\hat{x}_0) = \max_\theta \left(J_N(\theta, \hat{x}_0) \right).$$

B. Upper- and lower bounds of $J_N^\star$

This section develops upper and lower bounds on the objective, (2). As the maximum is greater than the average of any two points, we have that

$$\begin{aligned}
J_N^\star(\hat{x}_0) &\geq \sup_{i,j,y_{[0:N-1]}} \min_{\hat{x}_N} \frac{1}{2}\Big\{|\hat{x}_N - \breve{x}_{N,i}|^2_{(I-\gamma^{-2}P_i)^{-1}} \\
&\quad - \gamma^2 c_{N,i} + |\hat{x}_N - \breve{x}_{N,j}|^2_{(I-\gamma^{-2}P_j)^{-1}} - \gamma^2 c_{N,j}\Big\} \\
&= \sup_{i,j,y_{[0:N-1]}} \frac{1}{2}\Big\{|\breve{x}_{N,i} - \breve{x}_{N,j}|^2_{(2I-\gamma^{-2}P_i-\gamma^{-2}P_j)^{-1}} - \gamma^2 c_{N,i} - \gamma^2 c_{N,j}\Big\} \\
&\triangleq \max_{i,j} \underline{J}_N^{ij}(\hat{x}_0). \qquad (17)
\end{aligned}$$

Thus $\gamma_N < \gamma_N^\star$ only if $\underline{J}_N^{ij}(\hat{x}_0)$ is bounded for all pairs (i,j). Towards finding a sufficient condition, let $S \in \mathbb{R}^{n\times n}$ be a positive definite matrix such that $S \preceq I - \gamma^{-2}P_i$ for all $i = 1,\dots,M$. Then, applying Lemma 2 to (12), we have

$$\begin{aligned}
J_N^\star(\hat{x}_0) &\leq \max_{y,\theta}\Big\{\sum_{i,j}^{M} \theta_i\theta_j |\breve{x}_{N,i} - \breve{x}_{N,j}|^2_{S^{-1}}/2 - \gamma_N^2 \sum_i \theta_i c_{N,i}\Big\} \\
&\leq \frac{1}{2}\max_y \max_\theta \sum_i^M \theta_i \Big[-\gamma^2 c_{N,i} + \max_\sigma \Big\{\sum_j^M \sigma_j(|\breve{x}_{N,i} - \breve{x}_{N,j}|^2_{S^{-1}} - \gamma^2 c_{N,j})\Big\}\Big] \\
&= \max_{i,j} \underbrace{\max_y \frac{1}{2}\left\{|\breve{x}_{N,i} - \breve{x}_{N,j}|^2_{S^{-1}} - \gamma^2(c_{N,i} + c_{N,j})\right\}}_{\overline{J}_N^{ij}(\hat{x}_0)}. \qquad (18)
\end{aligned}$$

Thus, if $\overline{J}_N^{ij}(\hat{x}_0)$ is bounded for all pairs (i,j), then $\gamma_N^\star \leq \gamma_N$. The only difference between the expressions of $\overline{J}^{ij}(\hat{x}_0)$ and $\underline{J}^{ij}(\hat{x}_0)$ is the penalty of the term $|\breve{x}_{N,i} - \breve{x}_{N,j}|^2_*$.

Theorems 1 and 2 follow from applying Lemma 1 to the upper bound $\overline{J}_N^{ij}(\hat{x}_0)$ in (18) and the lower bound $\underline{J}_N^{ij}(\hat{x}_0)$ in (17) with $\theta_i = \theta_j = \frac{1}{2}$.

C. Lemmata

Lemma 2

Let $X_i \succ 0$ and $\theta_i \in (0,1)$ for $i = 1, \dots, M$ and that $\sum_{i=1}^{M} \theta_i = 1$. Let $S = \sum \theta_{i=1}^{M} X_i^{-1}$, then

$$\min_v \left\{ \sum \theta_i |v - x_i|^2_{X_i^{-1}} \right\} = \sum_{i=1}^{M} \theta_i \Big(|X_i^{-1} x_i|^2_{X_i - S^{-1}} + \frac{1}{2} \sum_{j=1}^{M} \theta_j \left(|X_i^{-1} x_i - X_j^{-1} x_j|^2_{S^{-1}} \right) \Big).$$

Proof. As $X_i \succ 0$, we have that $\sum \theta_i X_i^{-1} \succ 0$ and the (unique) minimum is a stationary point. We have

$$\min_v \left\{ \sum \theta_i |v - x_i|^2_{X_i^{-1}} \right\} = \sum_{i=1}^{M} \theta_i |x_i|^2_{X_i^{-1}} - |\sum_{i=1}^{M} \theta_i X_i^{-1} x_i|^2_{(\sum_1^M \theta_i X_i^{-1})^{-1}}$$

With $S := (\sum_1^M \theta_i X_i^{-1})$, we have that

$$\begin{aligned} -|\sum_{i=1}^{M} \theta_i X_i^{-1} x_i|^2_S &= -\sum_{i=1}^{M} \sum_{j=1}^{M} \theta_i \theta_j x_i^\top X_i^{-\top} S X_j^{-1} x_j \\ &= \frac{1}{2} \sum_{i=1}^{M} \sum_{j=1}^{M} \theta_i \theta_j \left(|X_i^{-1} x_i - X_j^{-1} x_j|^2_S \right) - \sum_{i=1}^{M} \theta_i |X_i^{-1} x_i|^2_S. \end{aligned}$$

□

Lemma 3—Interpolation

Let $z_k \in \mathbb{R}^n$ and $Z_k \in \mathbb{R}^{n \times n}$ be matrices such that $\sum_{k=1}^{K} Z_k \succ 0$ for $k = 1, \dots, K$. Then,

$$\min_x \left\{ \sum_{k=1}^{K} |x - z_k|^2_{Z_k} \right\} = \sum_{k=0}^{K} |z_k|^2_{Z_k} - \left| \sum_{k=1}^{K} Z_k z_k \right|^2_{(\sum_{k=1}^{K} Z_k)^{-1}}.$$

Proof. The problem is unconstrained and strictly convex—the minimizing solution is given by the stationary point. □

References

Anderson, B. and J. Moore (1979). *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ.

Åström, K. J. and B. Wittenmark (1997). *Computer-Controlled Systems (3rd Ed.)* Prentice-Hall, Inc., USA. ISBN: 0133148998.

Athans, M., D. Castanon, K. Dunn, C. Greene, W. Lee, N. Sandell, and A. Willsky (1977). "The stochastic control of the f-8c aircraft using a multiple model adaptive control (mmac) method–part i: equilibrium flight". *IEEE Trans. Autom. Control* **22**:5, pp. 768–780. DOI: 10.1109/TAC.1977.1101599.

Basar, T. and P. Bernhard (2008). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach*. 2nd ed. Birkhauser, Boston, MA, USA. DOI: 10.1007/978-0-8176-4757-5.

Başar, T. (1991). "Optimum performance levels for minimax filters, predictors and smoothers". *Syst. Control Lett.* **16**:5, pp. 309–317. ISSN: 0167-6911. DOI: 10.1016/0167-6911(91)90052-G.

Bates, J. M. and C. W. J. Granger (1969). "The combination of forecasts". *J. Oper. Res. Soc.* **20**, pp. 451–468. URL: https://api.semanticscholar.org/CorpusID:14178367.

Blom, H. A. P. and Y. Bar-Shalom (1988). "The interacting multiple model algorithm for systems with markovian switching coefficients". *IEEE Trans. Autom. Control* **33**:8, pp. 780–783. DOI: 10.1109/9.1299.

Buchstaller, D. and M. French (2016). "Robust stability for multiple model adaptive control: part i—the framework". *IEEE Trans. Autom. Control* **61**:3, pp. 677–692. DOI: 10.1109/TAC.2015.2492518.

Buchstaller, D., J. Liu, and M. French (2020). "The deterministic interpretation of the kalman filter". *Int. J. Control* **94**, pp. 1–19. DOI: 10.1080/00207179.2020.1755895.

Crassidis, J. L. and J. L. Junkins (2011). *Optimal Estimation of Dynamic Systems, Second Edition*. 2nd. Chapman & Hall/CRC, Boca Raton, FL, USA. ISBN: 1439839859.

Dasarathy, B. and B. Sheela (1979). "A composite classifier system design: concepts and methodology". *Proc. IEEE* **67**:5, pp. 708–713. DOI: 10.1109/PROC.1979.11321.

Dietterich, T. G. (2000). "Ensemble methods in machine learning". In: *Proc. 2000 Multiple Classifier Syst.* Cagliari, Italy. DOI: 10.1007/3-540-45014-9_1.

Gustafsson, F. (2000). *Adaptive Filtering and Change Detection*. Wiley, West Sussex, England. DOI: 10.1002/0470841613.

Hassani, V., A. P. Aguiar, M. Athans, and A. M. Pascoal (2009). "Multiple model adaptive estimation and model identification usign a minimum energy criterion". In: *Proc. 2009 IEEE Amer. Control Conf.* Pp. 518–523. DOI: 10.1109/ACC.2009.5160446.

Hespanha, J. P. (2001). "Tutorial on supervisory control". In: *40th IEEE Conf. Decis. Control.* Lecture notes for the workshop *Control using Logic and Switching.* Orlando, FL, USA.

Kjellqvist, O. and A. Rantzer (2022a). "Learning-enabled robust control with noisy measurements". In: Firoozi, R. et al. (Eds.). *Proc. 4th Annu. Learning Dyn. Control Conf.* Vol. 168. PMLR, pp. 86–96. URL: https://proceedings.mlr.press/v168/kjellqvist22a.html.

Kjellqvist, O. and A. Rantzer (2022b). "Minimax adaptive estimation for finite sets of linear systems". In: *Proc. 2022 IEEE Amer. Control Conf.* Pp. 260–265. DOI: 10.23919/ACC53348.2022.9867474.

Krener, A. (1980). "Kalman-bucy and minimax filtering". *IEEE Trans. Autom. Control* **25**:2, pp. 291–292. DOI: 10.1109/TAC.1980.1102301.

Lainiotis, D. G. (1976). "Partitioning: a unifying framework for adaptive systems, i: estimation". *Proc. IEEE* **64**:8, pp. 1126–1143. DOI: 10.1109/PROC.1976.10284.

Li, X.-R. and Y. Bar-Shalom (1996). "Multiple-model estimation with variable structure". *IEEE Trans. Autom. Control* **41**:4, pp. 478–493. DOI: 10.1109/9.489270.

Magill, D. (1965). "Optimal adaptive estimation of sampled stochastic processes". *IEEE Trans. Autom. Control* **10**:4, pp. 434–439. DOI: 10.1109/TAC.1965.1098191.

Mania, H., A. Jadbabaie, D. Shah, and S. Sra (2022). "Time varying regression with hidden linear dynamics". In: Firoozi, R. et al. (Eds.). *Proc. 4th Annu. Learning Dyn. Control Conf.* Vol. 168. PMLR, pp. 858–869. URL: https://proceedings.mlr.press/v168/mania22a.html.

Matni, N., A. Proutiere, A. Rantzer, and S. Tu (2019). "From self-tuning regulators to reinforcement learning and back again". In: *Proc. 58th IEEE Conf. Decis. Control*, pp. 3724–3740. DOI: 10.1109/CDC40024.2019.9029916.

P. Blom, H. A. (1984). "An efficient filter for abruptly changing systems". In: *Proc. 23rd IEEE Conf. Decis. Control*, pp. 656–658. DOI: 10.1109/CDC.1984.272089.

Rantzer, A. (2020). "Minimax adaptive control for state matrix with unknown sign". In: *Proc. 21st IFAC World Congr.* Vol. 53. 2, pp. 58–62. DOI: 10.1016/j.ifacol.2020.12.049.

Rantzer, A. (2021). "Minimax adaptive control for a finite set of linear systems". In: *Proc. 3th Annu. Learning Dyn. Control Conf.* Vol. 144. Corrected version: https://arxiv.org/abs/2011.10814, pp. 893–904. URL: https://proceedings.mlr.press/v144/rantzer21a.html.

Rong Li, X. and V. Jilkov (2005). "Survey of maneuvering target tracking. part v. multiple-model methods". *IEEE Trans. Aerosp. Navig. Electron.* **41**:4, pp. 1255–1321. DOI: 10.1109/TAES.2005.1561886.

Shen, X. and L. Deng (1997). "Game theory approach to discrete H_∞ filter design". *IEEE Trans. Signal Process.* **45**:4, pp. 1092–1095.

Silvestre, D., P. Rosa, and C. Silvestre (2021). "Distinguishability of discrete-time linear systems". *Int. J. Robust Nonlinear Control* **31**:5, pp. 1452–1478. DOI: 10.1002/rnc.5367.

Steel, M. F. J. (2020). "Model averaging and its use in economics". *J. Econ. Lit.* **58**:3, pp. 644–719. DOI: 10.1257/jel.20191385.

Willems, J. (2004). "Deterministic least squares filtering". *J. Econom.* **118**:1, pp. 341–373. DOI: 10.1016/S0304-4076(03)00146-5.

Paper III

Learning-Enabled Robust Control with Noisy Measurements

Olle Kjellqvist **Anders Rantzer**

Abstract

We present a constructive approach to bounded ℓ_2-gain adaptive control with noisy measurements for linear time-invariant scalar systems with uncertain parameters belonging to a finite set. The gain bound refers to the closed-loop system, including the learning procedure. The approach is based on forward dynamic programming to construct a finite-dimensional information state consisting of $\mathcal{H}_\infty$-observers paired with a recursively computed performance metric. We do not assume prior knowledge of a stabilizing controller.

1. Introduction

The great control engineer is lazy; her models are simplified and imperfect, the operating environment may be poorly controlled — yet her solutions perform well. Robust control provides excellent tools to guarantee performance if the uncertainty is small [Zhou and Doyle, 1998]. If the uncertainty is large, one can perform laborious system identification offline to reduce model uncertainty and synthesize a robust controller. An appealing alternative is to trade the engineering effort for a more sophisticated controller, particularly a learning-based component that improves controller performance as more data is collected. However, for such a controller to be implemented, it had better be robust to any prevalent unmodelled dynamics. Currently, there is considerable research interest in the boundary between machine learning, system identification, and adaptive control. For a review, see for example [Matni et al., 2019]. Most of the studies concern stochastic uncertainty

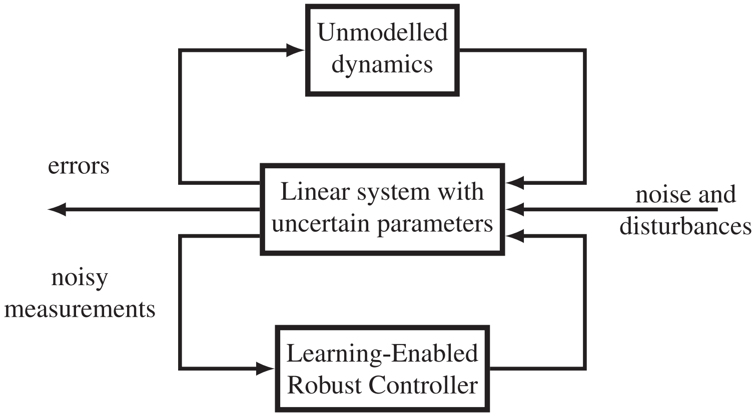


Figure 1. For a finite set of linear time-invariant models, the Learning-Enabled Robust Controller minimizes the ℓ_2-gain from noise and disturbances to errors for any realization of the unknown model parameters. This gain bound guarantees robustness to unmodelled dynamics.

and disturbances and assume perfect state measurements. Recently, works connecting to worst-case disturbances have started to appear. For example, non-stochastic control was introduced for known systems with unknown cost functions in [Agarwal et al., 2019] and extended to unknown dynamics and output feedback, under the assumption of bounded disturbances and prior knowledge of a stabilizing proportional feedback controller in [Simchowitz, 2020]. In [Dean et al., 2019] the authors leverage novel robustness results to ensure constraint satisfaction while actively exploring the system dynamics. In this contribution, the focus is on worst-case models for disturbances and uncertain parameters as discussed in [Didinsky and Basar, 1994], [Vinnicombe, 2004] and more recently in [Rantzer, 2021], but differ in that we consider output-feedback. See Figure 1 for an illustration of the considered problem. Unlike most recent contributions, the approach taken in this paper:

1. does not assume prior knowledge of a stabilizing controller. In particular, we allow for uncertain systems that a linear controller cannot stabilize,
2. assumes that the measurements are corrupted by additive noise,
3. provides guarantees on the ℓ_2-gain from disturbance and noise to state for the entire control duration.

1.1 Contributions and outline

We formalize the problem of finding a causal output-feedback controller with guaranteed finite ℓ_2-gain stability that is agnostic to the realization of the system parameters in Section 3. Section 4 is devoted to characterizing the Learning-Enabled Robust Controller in known or computable quantities. In Theorem 1 we show that ensuring finite ℓ_2-gain is equivalent to running one $\mathcal{H}_\infty$-observer for each feasible model, checking the sign of the associated cumulative cost and that each cumulative cost can be computed recursively. We show that it is necessary and sufficient to consider observer-based feedback in Theorem 2. In other words, the history can be compressed to a finite number of recursively computable quantities, growing linearly in the number of feasible models. In Section 5, we apply these results to synthesize a controller for an integrator with unknown input sign with a guaranteed bound on the ℓ_2-gain from noise and disturbances to error. All results in this paper are in discrete-time and for scalar systems, but sections 3 and 4 are readily extended to multivariable time-invariant systems.

2. Notation

The set of $n \times m$ matrices with real coefficients is denoted $\mathbb{R}^{n\times m}$. The transpose of a matrix A is denoted $A^\top$. For a symmetric matrix $A \in \mathbb{R}^{n\times n}$ and a vector $x \in \mathbb{R}^n$ we use the expression $|x|_A^2$ as shorthand for $x^\top Ax$. We write $A \prec (\preceq)\ 0$ to say that A is positive (semi)definite. We refer to the value of a signal w at time t as $w(t)$. The space of square-summable sequences from $\{T_0, T_0 + 1, \dots, T_f\}$ taking values in $\mathbb{R}$ is denoted $\ell_2[T_0, T_f]$. For a set $\mathcal{S}$, we let $\#(\mathcal{S})$ be the cardinality.

3. Learning-enabled control with guaranteed finite ℓ_2 gain

Given a positive quantity $\gamma > 0$ and a finite set of feasible models $\mathcal{M} \subset \mathbb{R}^3$, we concern ourselves with the uncertain linear system

$$\begin{aligned} x(t+1) &= ax(t) + bu(t) + w(t), \quad x(0) = x_0 \\ y(t) &= cx(t) + v(t), \quad t \geq 0 \end{aligned} \tag{1}$$

where the control signal $u(t) \in \mathbb{R}$ is generated by a causal output-feedback control policy

$$u(t) = \mu_t\left(y(0), y(1), \dots, y(t)\right). \tag{2}$$

In (1), $x(t) \in \mathbb{R}$ is the state, $y(t) \in \mathbb{R}$ is the measurement, the model $M := (a, b, c)$ is unknown but belongs to $\mathcal{M}$. The noise v and disturbances w satisfy $w, v \in \ell_2([0, T])$ for all $T \geq 0$. We are interested in control that makes the closed-loop system finite gain, with gain from (w, v) to x bounded above by γ. That is,

$$\alpha(T) := \sum_{\tau \leq T+1} x(\tau)^2 - \gamma^2 \sum_{\tau \leq T} w(\tau)^2 - \gamma^2 \sum_{\tau \leq T+1} v(\tau)^2 - P_M x(0)^2 \leq 0 \quad (3)$$

must hold for all $T \geq 0$, any admissible disturbances, initial state and the possible realizations M of (1). P_M quantifies prior information on the initial state and is taken as a positive solution to the Riccati equation

$$P_M = \left(a^2 \left(P_M + \gamma^2 c^2 - 1\right)^{-1} + \gamma^{-2}\right)^{-1}. \quad (4)$$

In this article, we explicitly construct controllers satisfying the finite-gain property and give conditions under which such controllers exist for the case when $c = 1$ and $b = \pm 1$.

Remark 1
The cases $b = -1$ and $b = 1$ cannot be simultaneously stabilized by a static feedback controller when $a \geq 1$

Remark 2
P_M could be any positive quantity. Our choice leads to stationary observer dynamics, simplifying the coming sections.

4. An information-state condition

In this section we will apply a slight modification to the $\mathcal{H}_\infty$-observer from [Basar and Bernhard, 2008] to bound (3) in a way which leads itself to recursive computation. We need the following lemma:

Lemma 1—Past cost
Given a known model $M = (a, b, c)$, a positive quantity γ, assume that the Riccati equation (4) has a positive solution P_M. For fixed $u \in \ell_2([0, t])$, $y \in \ell_2[0, t])$ and $x(t+1) \in \mathbb{R}$, we have that

$$\sup_{w,v \in \ell_2[0,t], x_0 \in \mathbb{R}} \left\{ \sum_{\tau \leq t} x(\tau)^2 - \gamma^2 \sum_{\tau \leq t} \left(w(t)^2 + v(t)^2\right) - Px(0)^2 : \ s.t.\ (1) \right\}$$
$$= -P_M(x(t+1) - \hat{x}_M(t+1))^2 + l_M(t+1). \quad (5)$$

The state observer $\hat{x}_M(t)$, and the past cost $l_M(t)$ are defined by the recursion

$$K_M = \frac{\gamma^2 c_M^2}{P_M + \gamma^2 c^2 - 1}, \quad \hat{w}_M(t) = \frac{\hat{x}_M(t)}{P_M + \gamma^2 c^2 - 1},$$
$$\hat{x}_M(t+1) = a\hat{x}(t) + bu(t) + K_M\left(y(t) - c\hat{x}(t)\right) + \hat{w}_M(t), \quad \hat{x}_M(0) = 0, \quad (6)$$
$$l_M(t+1) = l_M(t) - P_M\hat{x}_M(t)^2 - \gamma^2(y(t))^2 + \frac{\left(P_M\hat{x}_M(t) + \gamma^2 cy(t)\right)^2}{P_M + \gamma^2 c^2 - 1}, \quad (7)$$
$$l_M(0) = 0.$$

Remark 3
The observer form (6) makes sense for linear systems where we can design a state-feedback controller and observer separately and then join them together using the separation principle in [Basar and Bernhard, 2008]. The assumptions for the separation principle are not satisfied in our case, so we find it simpler to use the equivalent form

$$\hat{x}_M(t+1) = \hat{a}_M x(t) + bu(t) + \hat{g}_M y(t),$$

where $\hat{a}_M = aP_M/(P_M + \gamma^2 c^2 - 1)$ and $\hat{g}_M = \gamma^2 ac/(P_M + \gamma^2 c^2 - 1)$.

Proof Lemma 1. The system is equivalent to (6.1) and (6.2) in [Basar and Bernhard, 2008, p. 243] but with $D_k = \begin{bmatrix} I & 0 \end{bmatrix}$ and $E_k = \begin{bmatrix} 0 & I \end{bmatrix}$. Note that the term $-P_M x(0)^2$ in (5) ensures that $P_{k+1} = P_k = \ldots = P_M$, i.e. stationarity. Explicitly computing $l_M(t)$ requires some extra bookkeeping; in 6.35 the *terms independent of ξ and w* is equivalent to $\gamma^2|y(t)|^2_{(HH^\top)^{-1}} + |\hat{x}(t)|^2_{P(t)} - |u(t)|^2_R - l(t)$, the notational differences are $(HH^\top) \to N$, $P(t) \to K(t)$ and $l(t) \to c(t)$. After application of Lemma 6.2 on *p.* 259 we identify

$$m_k = -|P(t)\hat{x}(t) + \gamma^2 C^\top (HH^\top)^{-1} y(t)|^2_{(P(t)+\gamma^2 C^\top (HH^\top)^{-1}C - Q)^{-1}} + \gamma^2 |y(t)|^2_{(HH^\top)^{-1}} + |\hat{x}(t)|^2_{P(t)} - |u(t)|^2_R - l(t)$$

and conclude $l_M(t+1) = -m_k$. □

Lemma 1 lets us express the worst-case accumulated cost compatible with the dynamics as a function of the past trajectory (u, y) and the next state $x(t+1)$, if the dynamics M of the system (1) are known. As $x(t+1)$ changes, so does the set of trajectories w, v that are compatible with $x(t+1)$. In particular, the entire sequence of a maximizing trajectory will change as $x(t+1)$ is varied. With that in mind, it is remarkable that the effect to the accumulated cost is captured completely by the term $-P\left(x(t+1) - \hat{x}(t+1)^2\right)$. The second term $l(t+1)$ contains the terms of the cost that depend only on past inputs and outputs and is independent of $x(t+1)$.

We will study the value of the left-hand side of (3) for each model separately. Define for $M = (a, b, c) \in \mathcal{M}$, $y \in \ell_2[0, t]$ and an arbitrary output-feedback control policy μ the quantities

$$\alpha_M(t) := \sup_{w,v \in \ell_2[0,t], x_0 \in \mathbb{R}} \{\alpha(t) : (a, b, c) = M, \text{subject to (1) and (2)}\} \quad (8)$$

Then $\max_M \alpha_M(t)$ is the largest possible value of (3) at time t. In the following theorem, we use Lemma 1 to express α_M recursively and construct equivalent conditions using computable quantities.

Theorem 1—Information-state condition

Given a causal output-feedback control policy μ, a positive quantity γ, and an uncertainty set $\mathcal{M}$. Assume that for all $(a, b, c) = M \in \mathcal{M}$ the Riccati equation

$$P_M = \left(\frac{a^2}{P_M + \gamma^2 c^2 - 1} + \gamma^{-2} \right)^{-1} \quad (9)$$

a positive solution P_M and let

$$\hat{a}_M = \frac{a P_M}{P_M + \gamma^2 c^2 - 1}, \qquad \hat{g}_M = \gamma^2 \frac{ac}{P_M + \gamma^2 c^2 - 1}.$$

Further let

$$\hat{x}_M(t+1) = \hat{a}_M \hat{x}_M(t) + bu(t) + \hat{g}_M y(t), \;\; \hat{x}_M(0) = 0, \quad (10)$$

$$l_M(t+1) = l_M(t) - P_M \hat{x}_M(t)^2 - \gamma^2 y(t)^2 + \frac{(P_M \hat{x}_M(t) + \gamma^2 c y(t))^2}{P_M + \gamma^2 c^2 - 1}, \quad (11)$$

$$l_M(0) = 0. \quad (12)$$

Then the closed-loop system (1), (2) with control μ is finite gain for any realization $M \in \mathcal{M}$ if and only if $l_M(t+1) \leq 0$ holds for all $M \in \mathcal{M}$, $t \geq 0$ and $y \in \ell_2([0, t])$. If $P_M < 1$ for some M, γ is not an upper bound of the ℓ_2-gain from disturbance to error.

Proof. Let $\alpha_M(t)$ be defined as in (8). Then (3) holds for all (w, v, x_0), $M \in \mathcal{M}$ and T if and only if $\alpha_M(T) \leq 0$ for all $M \in \mathcal{M}$ and $y \in \ell_2[0, T]$. We now apply Lemma 1 to express $\alpha_M(t)$ in the known quantities $\hat{x}_M(t)$, P_M

and $l_M(t)$[1]:

$$\alpha_M(t) = \sup_{x(t),v(t)\in\mathbb{R}} \sup_{w,v\in\ell_2[0,t-1],x_0\in\mathbb{R}} \Bigg\{ x(t)^2 - \gamma^2 v(t)^2 + \sum_{\tau\le t-1} x(\tau)^2 - \gamma^2 \sum_{\tau\le t-1} \left(w(t)^2 + v(t)^2\right) : x(t+1) = ax(t) + bu(t) + w(t),\ y(t) = cx(t) + v(t),\ (a,b,c) = M \Bigg\}$$
$$= \sup_{x\in\mathbb{R},v\in\mathbb{R}} \left\{ x^2 - \gamma^2 v^2 - P_M\left(x - \hat{x}_M(t)\right)^2 + l_M(t) \right\}$$
$$= \left(P_M \hat{x}_M(t) + \gamma^2 c y(t)\right)^2 / (P_M + \gamma^2 c^2 - 1) - P_M \hat{x}_M^2(t) - \gamma^2 y(t)^2 + l_M(t) = l_M(t+1).$$

Finally, note that if for some M, $P_M < 1$, then $l_M(t+1)$ is strictly convex in $y(t)$ and thus unbounded from above. □

From Theorem 1 we see that the observer states $\hat{x}_M(t)$ and cumulative objectives $l_M(t+1)$ contain the information necessary and sufficient to evaluate the finite-gain condition (3). In other words, we can tell everything we need about the current state of affairs by running one $\mathcal{H}_\infty$ observer and computing l_M(t+1) for each model M in parallel; *but is it sufficient to consider observer-based feedback for control? If so, is it also necessary?*. the next theorem, we show that the observer states and cumulative objectives contain precisely the information required to synthesize a finite-gain control policy.

Theorem 2—Observer-based feedback

Given a positive quantity $\gamma > 0$ and an uncertainty set $\mathcal{M} \in \mathbb{R}^3$. The following are logically equivalent.

(i) There exists a causal output-feedback control policy $\mu^\star$ such that the closed-loop system (1) and (2) is finite-gain.

(ii) There exist observers $(\hat{x}_M, l_M)$ for each model $m \in \mathcal{M}$ generated by (10), (12) and an observer-based control policy $\eta^\star$

$$u(t) = \eta^\star \left\{ (\hat{x}_M(t), l_M(t+1), y(t)) : m \in \mathcal{M} \right\},$$

such that $l_M(t+1) \le 0$ for all $m \in \mathcal{M}$, $y \in \ell_2[0,t]$ and $t \ge 0$.

If $\eta^\star$ satisfies (ii), the following control policy satisfies (i):

$$\mu_t^\star\left(y(0), y(1) \ldots, y(t)\right) = \eta^\star \left\{ (\hat{x}_M(t), l_M(t+1), y(t)) : m \in \mathcal{M} \right\} \tag{13}$$

[1] We let subscript M denote quantities using $(a,b,c) = M$.

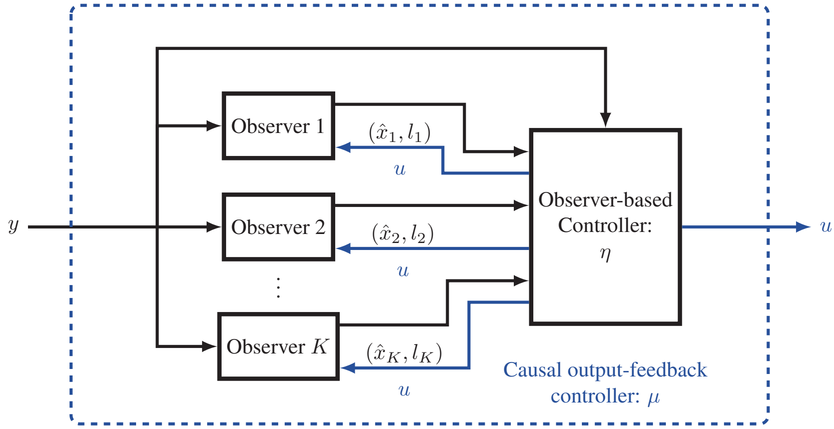


Figure 2. Illustration of the controller architecture in Theorem 2 for uncertainty sets consisting of K linear models. The controller η only considers the current state of the observers.

REMARK 4
By compressing the past trajectory to a finite set of cumulative performance quantities l_M, policies of this type learns the actual dynamics of the system as time goes on. This leads to a kind of multi-observer controller. The architecture is illustrated in 2.

Proof. Theorem 2 *(ii) implies (i)* follows from that $\hat{x}_M(t), l_M(t+1)$ depend causally on y, thus the observer-based control policy is a special case of causal feedback control policies. By assumption, $l_M(T) \leq 0$ for all T, M and $y \in \ell_2[0,T]$ for the controller (13), which we know implies that the system is finite gain by Theorem 1.

(i) implies (ii): Assume that the controller $\mu^\star$ fulfills (i). By the construction of (3) the Riccati equations have positive solutions P_M, therefore the assumptions of Theorem 1 are fulfilled and there exist observers $\hat{x}_M$ and l_M generated by (10) and (12). Define the set of feasible generating trajectories given observer states $\hat{x}_M(t)$, $l(t)$ and current measurement $y(t)$:

$$\mathcal{T}\left\{(\hat{x}_M(t), \hat{l}_M(t+1), y(t)) : M \in \mathcal{M}\right\}$$
$$:= \left\{(\breve{y}(\tau))_{\tau=0}^{T} : \breve{x}_M(T) = \hat{x}_M(t), \breve{y}(T) = y(t),\right.$$
$$\left.\breve{l}_M(T+1) = l_M(t+1), (\breve{x}_M, \breve{l}_M) \text{ generated by } \breve{y} \text{ and } u(\tau) = \mu^\star(\breve{y}(0), \ldots \breve{y}(\tau))\right\}.$$

Then $\mathcal{T}\{(\hat{x}(0), l_M(1), y(0)) : M \in \mathcal{M}\}$ is nonempty since it is compatible with any trajectory of length 1 such that $\breve{y}(0) = y(0)$. Fix $t \geq 0$

and observer states $\hat{x}_M(t), l_M(t+1)$ and measurement $y(t)$. Assume that $\mathcal{T}\{\hat{x}_M(t), l_M(t), y(t)) : M \in \mathcal{M}\}$ is non empty. Then there exists a sequence $\breve{y}$, and final time T so that $l_M(t+1) = \alpha_M(T)$ with $\alpha_M(t)$ as in (8) generated by $\breve{y}$ and the controller $u(\tau) = \mu^\star(\breve{y}(0), \ldots, \breve{y}(\tau))$. By assumption, $l_M(t+1) = \alpha_M(T) \leq 0$. Taking

$$\eta^\star \{(\hat{x}_M(t), l_M(t+1), y(t)) : M \in \mathcal{M}\} = \mu^\star(\breve{y}),$$

for some $\breve{y}, \in \mathcal{T}\{(\hat{x}_M(t) l_M(t+1), y(t)) : M \in \mathcal{M}\}$ ensures that $\mathcal{T}$ will be nonempty the next time step. By induction $\mathcal{T}$ will be nonempty for all $T \geq 0$ and thus u is well defined and $l_M(T) \leq 0$ for all T. □

5. Certainty equivalence control

We will now leverage these results to synthesize a control policy for the case when the pole $a \in \mathbb{R}$ is known, $b = \pm 1$ and $c = 1$. Emboldened by Theorem 2 we will construct a simple observer-based supervisory controller in the following way: We will run two observers in parallel corresponding to the cases $b = \pm 1$. The supervisor will monitor the cumulative objectives $l_{-1}(t)$ and $l_1(t)$ and determine which observer and model to use for computing the control signal. The policy computes the control signal as if the selected model were true. Let $i \in \{-1, 1\}$ index the observers. The Riccati equations (9) reduce to

$$P_i = P = \frac{1}{2}(1 - \gamma^2 a^2) + \sqrt{\gamma^2(-1+\gamma^2) + (\gamma^2 a^2 - 1)^2/4}. \tag{14}$$

Construct the observers $\hat{x}_i$ and cumulative objectives l_i using (10) and (12) with $b_i = i$ and

$$\hat{a}_i = \hat{a} = \frac{aP}{P + \gamma^2 - 1}, \quad \hat{g}_i = \hat{g} = \frac{\gamma^2 a}{P + \gamma^2 - 1}.$$

Define the *certainty-equivalence dead-beat controller* as the function

$$u(t) = \begin{cases} -(\hat{a}\hat{x}_1(t) + \hat{g}y(t)) & \text{if } l_1(t+1) \geq l_{-1}(t+1) \\ \hat{a}\hat{x}_{-1}(t) + \hat{g}y(t) & \text{if } l_1(t+1) < l_{-1}(t+1). \end{cases} \tag{15}$$

The dead-beat controller[2] ensures that for every t, either $\hat{x}_1(t)$ or $\hat{x}_{-1}(t)$ will be zero. This simplifies the observer dynamics $\hat{x}$ and the cost associated with the history l. We summarize the properties in the following proposition.

[2] The controller is dead-beat for the observer state corresponding to the model with the hightest cumulative cost. The observers themselves are not dead-beat.

PROPOSITION 1

With $\hat{a}$, $\hat{g}$, P as above, $\hat{x}_i$ and l_i as in (10) and (12), and the control signal given by (15), let

$$\hat{x}(t+1) = \hat{a}\hat{x}(t) + 2\hat{g}y(t), \quad \hat{x}(0) = 0.$$

Then the following is true:

$$\hat{x}_1(t) = \begin{cases} 0, & \text{if } l_1(t) \geq l_{-1}(t) \\ \hat{x}(t), & \text{if } l_1(t) < l_{-1}(t) \end{cases}, \quad \hat{x}_{-1}(t) = \begin{cases} \hat{x}(t), & \text{if } l_1(t) \geq l_{-1}(t) \\ 0, & \text{if } l_1(t) < l_{-1}(t), \end{cases}$$

and

$$l_1(t+1) = \begin{cases} l_1(t) - \gamma^2 y(t)^2 + \frac{(\gamma^2 y(t))^2}{P+\gamma^2-1} & \text{if } l_1(t) \geq l_{-1}(t) \\ l_1(t) - P\hat{x}(t)^2 - \gamma^2 y(t)^2 + \frac{(P\hat{x}(t)+\gamma^2 y(t))^2}{P+\gamma^2-1}, & \text{if } l_1(t) < l_{-1}(t) \end{cases}$$
$$l_{-1}(t+1) = \begin{cases} l_{-1}(t) - P\hat{x}(t)^2 - \gamma^2 y(t)^2 + \frac{(P\hat{x}(t)+\gamma^2 y(t))^2}{P+\gamma^2-1}, & \text{if } l_1(t) \geq l_{-1}(t) \\ l_{-1}(t) - \gamma^2 y(t)^2 + \frac{(\gamma^2 y(t))^2}{P+\gamma^2-1}, & \text{if } l_1(t) < l_{-1}(t) \end{cases} \tag{16}$$

Proof. We start by proving the first claim. Consider the case when $l_1(t+1) \geq l_{-1}(t+1)$. Then $\hat{x}_1(t+1) = 0$ and $\hat{x}_{-1}(t+1) = \hat{a}(\hat{x}_1(t) + \hat{x}_{-1}(t)) + 2\hat{g}y(t)$. The case when $l_1(t+1) < l_{-1}(t+1)$ is similar. Taking $\hat{x}(t) = \hat{x}_1(t) + \hat{x}_{-1}(t)$ completes the proof. To see that the second claim is true, note that if $l_1(t) \geq l_{-1}(t)$ then $\hat{x}_1(t) = 0$ and $\hat{x}_{-1}(t) = \hat{x}(t)$. The claim follows by substitution into (12). □

5.1 Conditions for finite-gain stability

This section determines sufficient conditions for the certainty-equivalence controller to guarantee a gain-bound of at most γ. We first give conditions on $l_1(t)$ and $l_{-1}(t)$ such that both quantities are negative for the next time step. We will then give conditions on γ so that the negativity conditions hold for all t. We summarize the non-negativity conditions in the following Lemma.

LEMMA 2

Given $P > 1$, $\gamma > 0$, $\hat{x}(t) \in \mathbb{R}$, $l_1(t)$ and $l_{-1}(t)$. Assume that $\max_{i\in\{-1,1\}} l_i(t) \leq 0$ and that

$$\min_i l_i(t) \leq -\frac{P}{P-1}\hat{x}(t)^2.$$

Then with $l_i(t+1)$ as in (16), it holds that $l_i(t+1) \leq 0$ for $i \in \{1, -1\}$.

Proof Lemma 2, full. We will give the proof for the case $0 \geq l_1(t) \geq l_{-1}(t)$. The case $0 \geq l_{-1}(t) \geq l_1(t)$ is similar. Note that $l_1(t+1)$ and $l_{-1}(t+1)$ are concave in $y(t)$ if and only if

$$\frac{1}{\gamma^2} \geq \frac{1}{P+\gamma^2-1} \iff P+\gamma^2-1 \geq \gamma^2,$$

and we conclude that $l_1(t+1)$ and $l_{-1}(t+1)$ are bounded from above if and only if $P \geq 1$. Secondly, we see that $l_1(t+1) = l_1(t) - cy^2 \leq 0$ for some positive constant c. Finally, let $X = P + \gamma^2 - 1$ and consider

$$\begin{aligned}
\max_{y(t)} l_{-1}(t+1) &= \max_{y(t)} \Big\{ l_{-1}(t) - P\hat{x}(t)^2 - \gamma^{-2}\left(\gamma^2 y(t)\right)^2 \\
&\quad + (P\hat{x}(t) + \gamma^2 y(t))^2/X \Big\} \\
&= \max_{y(t)} \Big\{ l_{-1}(t) + \left(-\gamma^{-2} + X^{-1}\right)\left(\gamma^2 y(t)\right)^2 \\
&\quad + 2X^{-1}P\hat{x}(t)\gamma^2 y(t) - (P - P^2/X)\hat{x}(t) \Big\} \\
&= l_{-1}(t) - \left(\frac{X^{-2}P^2}{-\gamma^{-2} + X^{-1}} + P - P^2/X \right) \hat{x}(t)^2 \\
&= l_{-1}(t) - \frac{\gamma^2 P^2/X + P(\gamma^2 - X) - P^2/X(\gamma^2 - X)}{\gamma^2 - X} \hat{x}(t)^2 \\
&= l_{-1}(t) - \frac{P(\gamma^2 - X) + P^2}{\gamma^2 - X} \hat{x}(t)^2 \\
&= l_{-1}(t) - \frac{P(1-P) + P^2}{1-P} \hat{x}(t)^2 \\
&= l_{-1}(t) + \frac{P}{P-1} \hat{x}(t)^2
\end{aligned}$$

Which is negative if and only if $l_{-1}(t) \leq -\frac{P}{P-1}\hat{x}(t)^2$. □

Next we give conditions on γ so that the assumptions in Lemma 2 are fulfilled for all t. This is illustrated in Figure 3, where subfigure (a) illustrates a case where $l_1(t+1)$ and $l_{-1}(t+1)$ cannot simultaneously be greater than $-\frac{P}{P-1}\hat{x}(t+1)^2$ and subfigure (b) illustrates the case when the condition is not guaranteed to hold for the next time step. For values of γ so that the system behaves as in Figure 3 (a), if the assumptions are fulfilled for some t, then (by induction) they will be fulfilled for all $T \geq t$. This is formalized in the next theorem.

THEOREM 3—CERTAINTY EQUIVALENCE, UPPER BOUND

Given a real number a and a quantity $\gamma > 0$. Assume that

$$P = \frac{1}{2}(1 - \gamma^2 a^2) + \sqrt{\gamma^2(-1+\gamma^2) + (\gamma^2 a^2 - 1)^2/4} > 1.$$

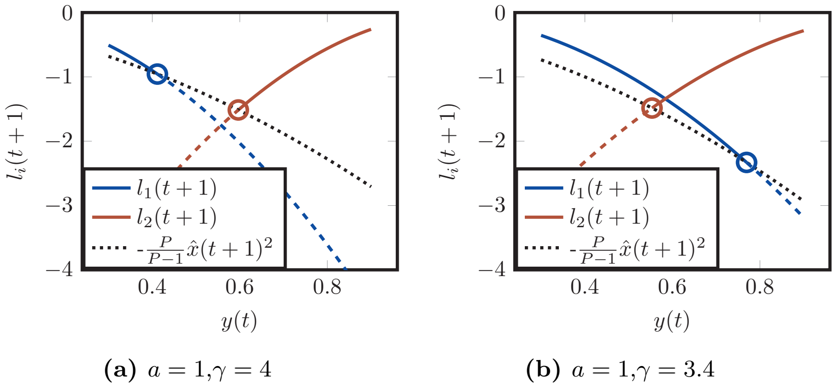


Figure 3. Illustrations of $l_1(t+1)$, $l_{-1}(t+1)$ and $-\frac{P}{P-1}\hat{x}(t+1)$ when $l_1(t) = 0$, $l_{-1}(t) = -\frac{P}{P-1}\hat{x}(t)^2$. The solid lines highlight the values of $y(t)$ where $l_i(t+1) \geq -\frac{P}{P-1}\hat{x}(t+1)^2$. We see that in (a) the solid lines do not overlap, i.e. given that the assumptions of Lemma 2 are fulfilled for some t, they will be fulfilled the next time step as well. In (b) the solid lines overlap, i.e. there are values for $y(t)$ so that the assumptions are violated the next time step.

If P and γ fulfill the curvature condition (17) *and* strong negativity condition (18) *below, then the closed-loop system* (1) *controlled with the certainty-equivalence deadbeat controller* (15) *has gain from $(w, v) \to x$ bounded above by γ.*

$$P > 2\gamma - 1 \tag{17}$$

$$(P + 2\gamma^2 - 1)\left(P - 1 - 2\sqrt{\gamma^2 - P})^2\right) \geq (P - 1)\left((P + 1)^2 - 4\gamma^2\right) \tag{18}$$

Remark 5
We can solve (18) *with equality restricted to the domain $P > 2\gamma - 1$. The resulting γ satisfies $(|a| + \sqrt{a^2+1})\sqrt{a^2+1} \leq \gamma \leq 2.1a^2 + 2$, and is shown in Figure 4.*

Remark 6
In [Vinnicombe, 2004], Vinnicombe studied the state-feedback version of the problem and found that the bound $\gamma = |a| + \sqrt{a^2+1}$ is achieved by the control policy

$$u(t) = \begin{cases} ax(t), & \text{if } \alpha_1(t) \leq \alpha_{-1}(t) \\ -ax(t), & \text{else,} \end{cases}$$

where $\alpha_b(t) = \sum_{\tau \leq t-1} (x(\tau+1) - ax(\tau) - bu(\tau))^2$. If we apply this control policy to the noisy measurements $y(t) = x(t) + v(t)$ we have that $x(t+1) = ax(t) + bu(t) + w(t) \pm av(t)$, and we get $\|x\|_2 \leq \gamma \| \begin{bmatrix} 1 & a \end{bmatrix} (w, v)\|_2 \leq (|a| + \sqrt{1+a^2})\sqrt{1+a^2}\|(w,v)\|_2$ which is the lower bound in Figure 4.

Proof Theorem 3, full. By assumption $P > 1$ is positive so Theorem 1 applies. We will show that if the *curvature condition* and the *strong negativity condition* are fulfilled, then the assumptions in Lemma 2 will hold for all t. Then, by Theorem 2 the observer-based controller is finite-gain for the original system. For $t = 0$, we have that $l_i(0) = 0$, $\hat{x}(0) = 0$ and that $l_i(t) \leq -\frac{P}{P-1}\hat{x}(0)^2$ holds trivially. Fix $t \geq 0$, assume without loss of generality that $0 \geq l_1(t) \geq l_{-1}(t)$ and that $l_{-1}(t) \leq -\frac{P}{P-1}\hat{x}(t)^2$. By Lemma 2 $\max_i\{l_i(t+1)\} \leq 0$. It remains to show that

$$\min_i\{l_i(t+1)\} \leq -\frac{P}{P-1}\hat{x}(t+1)^2. \tag{19}$$

Let $z(t) := y(t) - \frac{P}{2\gamma^2}\hat{x}(t)$. Then $\hat{x}(t+1) = 2\hat{g}z(t)$ and using Proposition 1, letting $X = P + \gamma^2 - 1$ we have

$$l_1(t+1) = l_1(t) + \left(-\frac{P\hat{x}(t)}{2} + \gamma^2 z(t)\right)^2 (1/X - 1/\gamma^2)$$

$$l_{-1}(t+1) = l_{-1}(t) + \left(\frac{P\hat{x}(t)}{2} + \gamma^2 z(t)\right)^2 /X - \left(-\frac{P\hat{x}(t)}{2} + \gamma^2 z(t)\right)^2 /\gamma^2 - P\hat{x}(t)^2$$

Curvature: For (19) to be true for all $z(t) \in \mathbb{R}$ it is necessary that $l_i(t+1) + 4\frac{P}{P-1}\hat{g}^2 z(t)^2$ is concave in $z(t)$. This is the case if and only if

$$\gamma^4(1/X - 1/\gamma^2) \leq -4\frac{P}{P-1}\hat{g}^2 \tag{20}$$

$$\iff \quad \gamma^4 \geq -4\frac{P}{P-1}\frac{1}{1/X - 1/\gamma^2}\hat{g}^2$$

Insert $\hat{g} = \gamma^2 a^2/X$ to get

$$-4\frac{P}{P-1}\frac{1}{1/X - 1/\gamma^2}\hat{g}^2 = 4\frac{P}{P-1}\frac{\gamma^2 X}{X - \gamma^2}\hat{g}^2 = \frac{4P}{(P-1)^2}\gamma^2 a^2/X\gamma^4.$$

Further, insert

$$P = \frac{1}{a^2/X + \gamma^{-2}} \iff \frac{a^2}{X} = \frac{1}{P} - \gamma^{-2}$$

to get

$$-4\frac{P}{P-1}\frac{1}{1/X - 1/\gamma^2}\hat{g}^2 = 4\frac{\gamma^2 - P}{(P-1)^2}\gamma^4. \tag{21}$$

The concavity condition (20) simplifies to the *curvature condition* (17),

$$1 \geq 4\frac{\gamma^2 - P}{(P-1)^2} \iff (P+1)^2 \geq 4\gamma^2 \iff \quad P \geq 2\gamma - 1.$$

Strong negativity: Define the upper bounds

$$\bar{l}_1(t+1) := \left(-\frac{P\hat{x}(t)}{2} + \gamma^2 z(t)\right)^2 (1/X - 1/\gamma^2)$$
$$\bar{l}_{-1}(t+1) := -\frac{P}{P-1}\hat{x}(t)^2 + \left(\frac{P\hat{x}(t)}{2} + \gamma^2 z(t)\right)^2 /X$$
$$- \left(-\frac{P\hat{x}(t)}{2} + \gamma^2 z(t)\right)^2 /\gamma^2 - P\hat{x}(t)^2.$$

Also define the sets

$$\mathcal{I}_i := \left\{ z \in \mathbb{R} : l_i(t+1) \geq -4\frac{P}{P-1}\hat{g}^2 z(t)^2 \right\}.$$

and $\bar{\mathcal{I}}_i$ anagolously. Then the inequality (19) is satisfied if and only if $\#\left(\mathcal{I}_1 \cap \mathcal{I}_{-1}\right) \leq 1$. Since $\bar{l}_i \geq l_i$ we have that $\mathcal{I}_i \subseteq \bar{\mathcal{I}}_i$, and a sufficient condition is that they intersection contains at most one point, i.e. $\#\left(\bar{\mathcal{I}}_1 \cap \bar{\mathcal{I}}_{-1}\right) \leq 1$. The reason we allow for the intersection to contain one point, is that at such a point both $l_1(t+1)$ and $l_{-1}(t+1)$ fulfills (19) with equality. We will start with characterizing $\bar{\mathcal{I}}_1$ by looking for the solutions to $\bar{l}_1(t+1) = -4\frac{P}{P-1}\hat{g}^2 z(t)^2$:

$$\left(-\frac{P\hat{x}(t)}{2} + \gamma^2 z(t)\right)^2 (1/X - 1/\gamma^2) = -4\frac{P}{P-1}\hat{g}^2 z(t)^2$$
$$\iff \left(-\frac{P\hat{x}(t)}{2} + \gamma^2 z(t)\right)^2 = 4\frac{\gamma^2 - P}{(P-1)^2}(\gamma^2 z(t))^2$$
$$\iff \left(-\frac{P\hat{x}(t)}{2} + \gamma^2 \left(1 + 2\frac{\sqrt{\gamma^2 - P}}{P-1}\right) z(t)\right)$$
$$\times \left(-\frac{P\hat{x}(t)}{2} + \gamma^2 \left(1 - 2\frac{\sqrt{\gamma^2 - P}}{P-1}\right) z(t)\right) = 0$$

We conclude that for positive $\hat{x}(t)$

$$\bar{\mathcal{I}}_1 = \left[\frac{P}{2\gamma^2}\left(1 + 2\frac{\sqrt{\gamma^2 - P}}{P-1}\gamma^2 z(t)^2\right)^{-1} \hat{x}(t),\right.$$
$$\left.\frac{P}{2\gamma^2}\left(1 - 2\frac{\sqrt{\gamma^2 - P}}{P-1}\gamma^2 z(t)^2\right)^{-1} \hat{x}(t)\right].$$

We continue with the solutions to $\bar{l}_2(t+1) = -4\frac{P}{P-1}\hat{g}^2 z(t)^2$.

$$-\frac{P}{P-1}\hat{x}(t)^2+\left(\frac{P\hat{x}(t)}{2}+\gamma^2 z(t)\right)^2/X-\left(-\frac{P\hat{x}(t)}{2}+\gamma^2 z(t)\right)^2/\gamma^2-P\hat{x}(t)^2 = -4\frac{P}{P-1}\hat{g}^2 z(t)^2$$

Using (21) we get

$$\begin{aligned}
\iff & \left(\frac{1}{X}-\frac{1}{\gamma^2}\right)\left(1-4\frac{\gamma^2-P}{(P-1)^2}\right)\left(\gamma^2 z(t)\right)^2+\left(\frac{1}{X}+\frac{1}{\gamma^2}\right)P\hat{x}(t)\gamma^2 z(t) \\
&+\left(\frac{1}{4}\left(\frac{1}{X}-\frac{1}{\gamma^2}\right)-\frac{1}{P-1}\right)\left(P\hat{x}(t)\right)^2=0 \\
\iff & \left(z(t)\right)^2-\frac{X+\gamma^2}{X-\gamma^2}\frac{(P-1)^2}{(P-1)^2-4(\gamma^2-P)}P\hat{x}(t)\gamma^2 z(t) \\
&+\frac{\frac{1}{4}-\frac{1}{P-1}\frac{1}{1/X-1/\gamma^2}}{(P-1)^2-4(\gamma^2-P)}(P-1)^2P^2\hat{x}(t)^2=0 \\
\iff & \left(\gamma^2 z(t)\right)^2-\frac{(P+2\gamma^2-1)(P-1)}{(P+1)^2-4\gamma^2}P\hat{x}(t)\gamma^2 z(t) \\
&+\frac{1}{4}\frac{(P-1)^2+4\gamma^2(P+\gamma^2-1)}{(P+1)^2-4\gamma^2}P^2\hat{x}(t)^2=0 \\
\iff & \left(\gamma^2 z(t)-\frac{1}{2}\frac{(P+2\gamma^2-1)(P-1)}{(P+1)^2-4\gamma^2}P\hat{x}(t)\right)^2 \\
&-(P+2\gamma^2-1)^2\frac{\gamma^2-P}{\left((P+1)^2-4\gamma^2\right)^2}P^2\hat{x}(t)^2=0
\end{aligned}$$

which has the solutions

$$z(t)=\frac{1}{2\gamma^2}(P+2\gamma^2-1)\frac{P-1\pm 2\sqrt{\gamma^2-P}}{(P+1)^2-4\gamma^2}P\hat{x}(t).$$

Thus for positive $\hat{x}(t)$,

$$\bar{\mathcal{I}}_{-1}=\left[\frac{1}{2\gamma^2}(P+2\gamma^2-1)\frac{P-1-2\sqrt{\gamma^2-P}}{(P+1)^2-4\gamma^2}P\hat{x}(t),\right. \\ \left.\frac{1}{2\gamma^2}(P+2\gamma^2-1)\frac{P-1+2\sqrt{\gamma^2-P}}{(P+1)^2-4\gamma^2}P\hat{x}(t)\right]$$

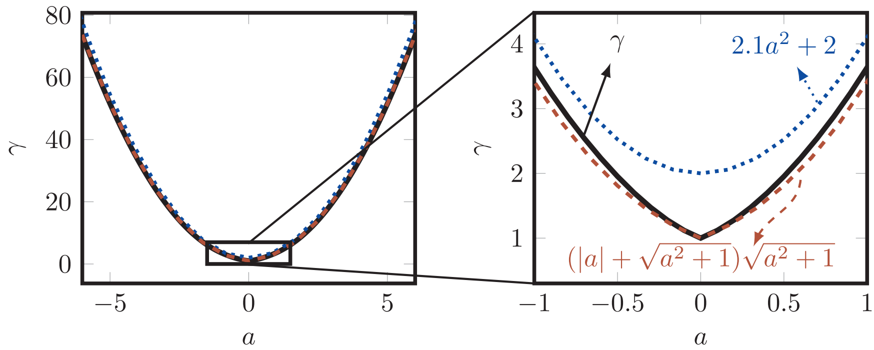


Figure 4. Guaranteed bound on the ℓ_2-gain from disturbances to error under feedback with the certainty equivalence controller with respect to a. We note that experimentally γ is lower bounded by $(|a| + \sqrt{a^2+1})\sqrt{a^2+1}$ and upper bounded by $\leq 2.1a^2 + 2$. The lower bound becomes tighter as a increases.

From the definition, it is clear that the vertex of $\bar{l}_1(t+1)$ lies closer to the origin, than that of $\bar{l}_{-1}(t+1)$. Thus $\#\left(\bar{\mathcal{I}}_1 \cap \bar{\mathcal{I}}_2\right) \leq 1$ is equivalent to

$$\frac{P}{2\gamma^2}\left(1 - 2\frac{\sqrt{\gamma^2 - P}}{P-1}\gamma^2 z(t)^2\right)^{-1} \hat{x}(t) \leq \frac{1}{2\gamma^2}(P + 2\gamma^2 - 1)\frac{P - 1 - 2\sqrt{\gamma^2 - P}}{(P+1)^2 - 4\gamma^2} P\hat{x}(t),$$

which simplifies to (18). The case when $\hat{x}(t)$ is negative is similar. □

6. Conclusions

This article presents a constructive approach to accounting for worst-case models of measurement noise, disturbance and uncertain parameters in controller design. In particular Theorem 2 shows that it is necessary and sufficient to consider feedback from the current states of a finite set of observers and cumulative performance measures. The performance measures compress the history allowing the controller to learn from past data. In Section 5, we used this constructive approach to extend the results of [Vinnicombe, 2004] to the case of noisy measurements. We focused on scalar systems, but Theorems 1 and 2 can easily be extended to MIMO systems. In particular, we are excited about the potential in extending Minimax Adaptive Control [Rantzer, 2021] to the output feedback case.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 834142 (ScalableControl). The authors are thankful to their colleagues Bo Bernhardsson and Venkatraman Renganathan (Department of Automatic Control, Lund University) for help in reviewing and revising earlier versions of the manuscript.

References

Agarwal, N., B. Bullins, E. Hazan, S. Kakade, and K. Singh (2019). "Online control with adversarial disturbances". In: Chaudhuri, K. et al. (Eds.). *Proc. 36th Int. Conf. Mach. Learn.* Vol. 97. PMLR, pp. 111–119. URL: https://proceedings.mlr.press/v97/agarwal19c.html.

Basar, T. and P. Bernhard (2008). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach.* 2nd ed. Birkhauser, Boston, MA, USA. DOI: 10.1007/978-0-8176-4757-5.

Dean, S., S. Tu, N. Matni, and B. Recht (2019). "Safely learning to control the constrained linear quadratic regulator". In: *Proc. 2019 IEEE Amer. Control Conf.* Pp. 5582–5588. DOI: 10.23919/ACC.2019.8814865.

Didinsky, G. and T. Basar (1994). "Minimax adaptive control of uncertain plants". In: *Proc. 33rd IEEE Conf. Decis. Control.* Vol. 3, pp. 2839–2844. DOI: 10.1109/CDC.1994.411368.

Matni, N., A. Proutiere, A. Rantzer, and S. Tu (2019). "From self-tuning regulators to reinforcement learning and back again". In: *Proc. 58th IEEE Conf. Decis. Control*, pp. 3724–3740. DOI: 10.1109/CDC40024.2019.9029916.

Rantzer, A. (2021). "Minimax adaptive control for a finite set of linear systems". In: *Proc. 3th Annu. Learning Dyn. Control Conf.* Vol. 144. Corrected version: https://arxiv.org/abs/2011.10814, pp. 893–904. URL: https://proceedings.mlr.press/v144/rantzer21a.html.

Simchowitz, M. (2020). "Making non-stochastic control (almost) as easy as stochastic". In: Larochelle, H. et al. (Eds.). *Adv. Neural Inf. Process. Syst.* Vol. 33. Curran Associates, Inc., pp. 18318–18329. URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/d4ca950da1d6fd954520c45ab19fef1c-Paper.pdf.

Vinnicombe, G. (2004). "Examples and counterexamples in finite ℓ_2-gain adaptive control". In: *16th Int. Symp. Mathematical Theory Networks Systems.* Leuven, Belgium.

Zhou, K. and J. C. Doyle (1998). *Essentials of Robust Control.* Prentice-Hall, Upper Saddle River, New Jersey, USA.

Paper IV

Minimax Dual Control with Finite-Dimensional Information State

Olle Kjellqvist

Abstract

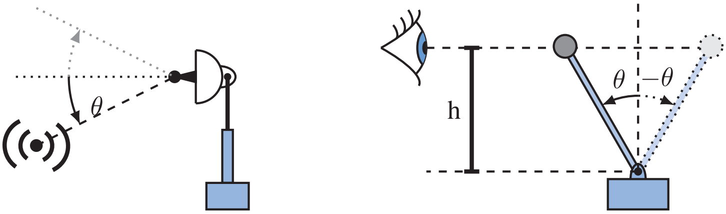
This article considers output-feedback control of systems where the function mapping states to measurements has a set-valued inverse. We show that if the set has a bounded number of elements, then minimax dual control of such systems admits finite-dimensional information states. We specialize our results to a discrete-time integrator with magnitude measurements and derive a surprisingly simple sub-optimal control policy that ensures finite gain of the closed loop. The sub-optimal policy is a proportional controller where the magnitude of the gain is computed offline, but the sign is learned, forgotten, and relearned online.

The discrete-time integrator with magnitude measurements captures real-world applications such as antenna alignment, and despite its simplicity, it defies established control-design methods. For example, whether a stabilizing linear time-invariant controller exists for this system is unknown, and we conjecture that none exists.

1. Introduction

This article concerns ouput feedback control of discrete-time systems whose measurement equations have a bounded number of solutions. As a prototype example, we consider the discrete-time integrator, where the controller only has access to the magnitude of the state. The state x_t, the control signal u_t, and disturbance w_t are real-valued scalars. The system is described by the recursion

$$x_{t+1} = x_t + u_t + w_t. \tag{1}$$



(a) Receiver **(b)** Pendulum

Figure 1. Examples of physical systems where the sign of the state is ambiguous: The left figure illustrates a receiver with an uncertain and potentially non-stationary source location. The objective is to adjust the receiver's position to an angle that maximizes signal intensity. Typically, the receiver's radiant sensitivity is symmetric relative to deviations from the incidence angle. The right figure shows an inverted pendulum, which is regulated by monitoring the pendulum's height.

We consider causal control policies, μ, that map measurements of the state magnitude

$$y_t = |x_t| \tag{2}$$

to control signals

$$u_t = \mu_t(y_0, y_1, \dots, y_t, u_0, \dots, u_{t-1}). \tag{3}$$

The uncertain sign in (2) captures some of the difficulties that may arise when optimizing a system based on measurements of some (locally) convex or concave performance quantity, as in Figure 1a. The problem is also closely related to stabilizing an inverted pendulum by feedback from height measurements rather than angular measurements, as in Figure 1b. This plain-looking problem captures a surprising amount of complexity:

1. *Exploration vs. exploitation.* The more effectively we control the system, the less confident we become about the state's sign. If the system ever reaches $y = 0$, the state's sign information is lost.

2. *No stabilizing linear time-invariant controller.* Previous work report no stabilizing linear time-invariant controller for the system (1)–(3) [Rosdahl and Bernhardsson, 2020; Alspach, 1972] and the system cannot be stabilized by proportional feedback[1]. This author *conjectures* that there exists no finite-dimensional linear time-invariant controller that stabilizes the system.

3. *Extended Kalman filter.* The extended Kalman filter (EKF) is a popular algorithm for estimating a nonlinear system's state, often coupled with

[1] A linear time-varying controller can stabilize the system. For example, $u_t = (-1)^t y_t$ will ensure $x_t = 0$ for all $t \geq 2$, for any x_0 and $w_t = 0$.

certainty-equivalence control. However, the measurement equation (2) is not differentiable at $x = 0$, and the EKF is not directly applicable. One may substitute the measurement equation with $y_t = x_t^2$ to recover differentiability, but this substitution results in an unobservable linearization.

4. *Myopic Controller.* The Myopic controller [Wittenmark, 1995] associated with minimizing the current cost $x_t^2 + u_t^2$ is not stabilizing.

In this article, we will design a control policy (3) that ensures that the induced ℓ_2-gain from w to (x, u) is less than some positive quantity γ. That is, the inequality

$$\sum_{t=0}^{N}(x_t^2 + u_t^2) \leq \gamma^2 \sum_{t=0}^{N} w_t^2 + \beta(x_0) \tag{4}$$

must be fulfilled for all $N \geq 0$, real-valued function β and realizations $w_{[0:N]} := w_0, w_1, \dots, w_N$ of the disturbance sequence. The condition (4) generalizes the classical $\mathcal{H}_\infty$-norm for linear systems. The function β is called a bias term and is used to capture the effect of the initial state. The *small-gain theorem* provides sufficient conditions for robust stability against feedback perturbations with induced ℓ_2-norm less than γ^{-1}. We refer the reader to [Khalil, 2002, Chapter 5] for a detailed discussion on finite-gain stability and the small-gain theorem. Surprisingly, we will see that it is possible to compress the observed output trajectory $(y_{[0:t]}, u_{[0:t-1]})$ using two recursively computed quantities r_t^+ and r_t^-. These quantities correspond to the smallest feasible disturbance trajectory compatible with the observed outputs and $\text{sign}(x_t) = 1$ or $\text{sign}(x_t) = -1$. Together with y_t, they make a *sufficient statistic* for optimal control of a corresponding dynamic game.

The quantities follow the recursions

$$\begin{aligned} r_{t+1}^+ &= y_t^2 + u_t^2 - \max\{r_t^+ + \gamma^2(y_{t+1} - u_t - y_t)^2, r_t^- + \gamma^2(y_{t+1} - u_t + y_t)^2\}, \\ r_{t+1}^- &= y_t^2 + u_t^2 - \max\{r_t^+ + \gamma^2(y_{t+1} + u_t + y_t)^2, r_t^- + \gamma^2(y_{t+1} + u_t - y_t)^2\}. \end{aligned} \tag{5}$$

In Section 3, we will show that these quantities are sufficient for ensuring bounded ℓ_2 gain and summarize our conclusions about the magnitude control problem in Proposition 1.

Proposition 1

An admissible policy μ exists that ensures ℓ_2-gain smaller than γ if, and only if, it is achievable with a policy of the form $u_t = \eta_t(y_t, r_t^+, r_t^-)$. Further, the controller $\eta(y, r^+, r^-) = 0.7\,\text{sign}(r^- - r^+)y$ achieves ℓ_2-gain less than 4.

We remark that η is admissible as r_t^+ and r_t^- are functions of previous measurements and control signals. Via substitution, one can recover μ.

1.1 Related work

Adaptive control From the adaptive control perspective, system (1), (2) could be interpreted as a linear system with uncertain time-varying parameters. Several methods are described in excellent textbooks like [Goodwin and Sin, 2009, Chapter 6.7] that apply uncertain linear time-varying systems. However, these methods rely on a separation of time scales between the state dynamics, the parameter adaptation, and the parameter variation. Hence, we can not expect these methods to work well in our case [Anderson and Dehghani, 2008]. Nonlinear stochastic control theory provides a framework that can, in principle, handle fast parameter variation and large uncertainties, and our problem fits well with the methodology of *dual control* [Wittenmark, 1995, Chapter 7].

Stochastic dual control has been applied to various problems with uncertain gain, as demonstrated in [Åström and Helmersson, 1986; Dumont and Åström, 1988; Allison et al., 1995]. [Alspach, 1972] considered control of an integrator based on noisy measurements of the square of the magnitude. The noise was assumed Gaussian, and the author proposed approximating the information state by a sum of Gaussians. [Rosdahl and Bernhardsson, 2020] considered a noisy version of the problem in this article but from a stochastic dual control perspective. The authors proposed to approximate the information state by a neural network.

Learning-to-control Lately, there has been a surge of interest in learning to control linear systems. Much of the work concerns the sample complexity of learning optimal controllers of linear time-invariant systems. For example, [Dean et al., 2018; Mania et al., 2019] concerns quadratic performance objectives and additive stochastic noise, [Chen and Hazan, 2021] adapts the theory of online convex optimization [Hazan, 2023] to unknown linear time-invariant systems with bounded disturbances. [Yu et al., 2023] proposed a method to control slowly varying linear systems with unknown parameters belonging to a polytope perturbed by bounded disturbances using convex body chasing.

Minimax control Minimax control for uncertain systems was introduced in the Ph.D. thesis of [Witsenhausen, 1966]. Information states, or sufficient statistics, for optimal control for output feedback minimax control, was discussed in [Bertsekas and Rhodes, 1973] based on Bertsekas's Ph.D. thesis. The game-theoretic formulation of $\mathcal{H}_\infty$-control [Basar and Bernhard, 2008] is a special case of minimax control, and the information state formulation was derived for nonlinear systems in [James and Baras, 1995] demonstrating that, in general, the information state is infinite-dimensional. The term minimax adaptive control was introduced in [Didinsky and Basar, 1994]. Recently, [Rantzer, 2021] proposed a minimax adaptive controller for uncertain linear systems with perfect state measurements. The uncertainty was assumed to belong to a finite, known set. The author proposed a finite-dimensional information state related to the empirical covariance matrix of the current state, previous state, and previous control signal. This author extended Rantzer's results to scalar linear systems with noisy measurements in [Kjellqvist and Rantzer, 2022]. Recently, [Renganathan et al., 2023] studied the regret of Rantzer's controller for linear systems with energy-bounded disturbances.

1.2 Contributions

This article identifies a class of systems where the minimax dual controller admits a finite-dimensional information state. The information state admits recursive computation, and Theorem 1 shows the equivalence between the minimax dual control problem and an information-state dynamic programming problem. We also provide a dissipativity interpretation in Theorem 2. The proofs of Theorems 1 and 2 are available in the ArXiV version of this article [Kjellqvist, 2024]. These results generalize Theorem 1 in [Rantzer, 2021] to a larger system class and specialize the results in [James and Baras, 1995] to classes of systems where the information state iteration becomes explicit. The explicit iteration results from the bounded number of solutions to the measurement equation (2) and can be exploited to obtain closed-form (suboptimal) solutions to the minimax dual control problem. We specialize these results to the magnitude control problem in the introduction and prove Proposition 1 in Section 3.

1.3 Notation

We use $\mathbb{R}$ to denote the set of real numbers, $\mathbb{R}^n$ means the set of n-dimensional real vectors, and $\mathbb{R}^{n\times m}$ means the set of $n \times m$ real matrices. The vector of ones is denoted $\mathbf{1}$. We use $y_{[0:N]}$ as shorthand for the sequence $(y_0, y_1, \ldots, y_N)$. For a matrix $A \in \mathbb{R}^{n\times m}$, we denote the transpose by A^T. For sets $A \subseteq S$ and $B \subseteq T$, and a function $f : S \to T$, the image of A is denoted $f(A)$ and the preimage of B is denoted $f^{-1}(B)$; the Cartesian product is denoted $S \times T$ and the n-ary Cartesian power $S = \underbrace{S \times S \times \ldots \times S}_{n \text{times}}$ is denoted S^n. For vectors $v, v' \in \mathbb{R}^n$, the inequality $v \leq v'$ is understood component-wise, and for functions $f, g : S \to T$, the inequality $f \leq g$ means that $f(s) \leq g(s)$ for all $s \in S$ where $\leq$ is the partial order on T. Strict inequalities are defined analogously.

2. Minimax dual control

This section introduces the minimax dual control problem, the information state, and dynamic programming. By information state, we mean an auxiliary state variable that is computable by the controller, has a recursive expression in observed quantities and is sufficient to compute the optimal control policy and the associated cost. For example, in the linear-quadratic Gaussian control problem, the information state is the conditional mean and covariance of the state given the observations—the Kalman filter estimate and the error covariance. It is well known that the "worst-case" history is an information state for the minimax control problem, and dynamic programming with this information state is pretty well understood. Unfortunately, this information state is generally infinite-dimensional and, therefore, impractical. The main contribution of this section is to show that for our class of systems, the worst-case history admits a finite-dimensional representation. This representation is, in itself, an information state. We derive a verification and an approximation theorem for value iteration specific to this finite-dimensional representation.

2.1 Problem formulation

Let $f : \mathcal{X} \times \mathcal{U} \times \mathcal{W} \to \mathcal{X}$ and $h : \mathcal{X} \to \mathcal{Y}$ describe the dynamical system

$$\begin{aligned} x_{t+1} &= f(x_t, u_t, w_t) \\ y_t &= h(x_t). \end{aligned} \tag{6}$$

The control signal, $u_t \in \mathcal{U}$ is generated by a causal control policy $\mu_t : \mathcal{Y}^t \times \mathcal{U}^{t-1} \to \mathcal{U}$, where $\mathcal{Y} = h(\mathcal{X})$ by

$$u_t = \mu_t(y_{[0:t]}, u_{[0:t-1]}). \tag{7}$$

We call the tuple $\pi = (\mu_0, \mu_1, \ldots)$ a *strategy* and the set of all such admissible strategies Π. Consider the objective function as the "worst-case" sum of stage costs $l : \mathcal{X} \times \mathcal{U} \times \mathcal{W} \to \mathbb{R}$,

$$J_\pi^N(y_0) \triangleq \sup_{w_{[0:N]}} \left\{ \sum_{t=0}^{N} l(x_t, u_t, w_t) : w_{[0:N]} \in \mathcal{W}^{N+1}, y_0 = h(x_0) \right\}. \tag{8}$$

The goal of this section is to examine the minimax optimal control problem

$$J_\star(y_0) \triangleq \inf_{\pi \in \Pi} \sup_N J_\pi^N(y_0). \tag{9}$$

We make two crucial assumptions:

ASSUMPTION 1
For all $x \in \mathcal{X}, u \in \mathcal{U}$, $\sup_w l(x, u, w) \geq 0$.

The assumption that $\sup_w l(x,u,w) \geq 0$ implies monotonicity properties of J_π^N in (8) and, as we will see later, the value iteration.

Assumption 2
For any $y \in \mathcal{Y}$, the preimage $h^{-1}\{y\} \subset \mathcal{X}$ is an indexed set of at most M elements.

This assumption relates to the dimensionality of the information state, or sufficient statistic, of the dynamic programming version of this problem. Technically, the bound M does not have to be known a priori, but we require the capability to enumerate all the solutions to $y_t = h(x_t)$ online. At first glance, this assumption may appear overly limiting, but the following examples prove otherwise.

Example 1—Magnitude control of input-output models
Consider controller design for the input-output system

$$z_{t+1} = -a_1 z_t - \cdots - a_d z_{t-d+1} + b_1 u_t + \cdots + b_d u_{t-d+1} + w_t, \tag{10}$$

where the controller has access magnitude measurements $|z_0|, |z_1|, \ldots, |z_t|$ at time t. The system (10) has a (nonminimal) state-space realization $x_{t+1} = Ax_t + Bu_t + Gw_t$, where

$$x_t = \begin{bmatrix} z_t \\ \vdots \\ z_{t-d+1} \\ u_{t-1} \\ \vdots \\ u_{t-d+1} \end{bmatrix}, \; A = \begin{bmatrix} -a_1 & \cdots & -a_d & b_2 & \cdots & b_d \\ 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix}, \; B = \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \; G = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Store the past $d-1$ inputs and outputs and define the augmented measurement

$$y_t = h(x_t) = (|z_t|, \ldots, |z_{t-n+1}|, u_{t-1}, \ldots, u_{t-n+1}).$$

Then, the preimage

$$h^{-1}\{y_t\} = \{\pm|z_t|\} \times \cdots \times \{\pm|z_{t-d+1}|\} \times \{u_{t-1}, \ldots, u_{t-d+1}\}$$

has cardinality 2^n, corresponding to the possible signs of the past measurements. A first-order difference equation can model the integrator in the introduction, so $M = 2^1 = 2$, and the inverted pendulum (linearized around its equilibrium) by a second-order difference equation, for which $M = 2^2 = 4$.

EXAMPLE 2—LINEAR SYSTEM WITH UNCERTAIN DYNAMICS
Consider the linear system $x_{t+1} = Ax_t + Bu_t + w_t$ where A, B are unknown matrices belonging to a finite set $\mathcal{M}$ of cardinality M. Then, the equivalent lifted system $x_t = (z_t, A_t, B_t)$ with

$$\begin{aligned} A_{t+1} = A_t, \quad B_{t+1} = B_t, (A_0, B_0) \in \mathcal{M} \\ z_{t+1} = Az_t + Bu_t + w_t, \quad y_t = h(x_t) = z_t \end{aligned}$$

satisfies Assumption 2 as $h^{-1}\{y_t\} = \{z_t\} \times \mathcal{M}$ has cardinality M.

EXAMPLE 3—FINITE STATE SPACE
If the state space $\mathcal{X}$ is finite, per definition $h^{-1}\{y\} \subseteq \mathcal{X}$ is finite.

REMARK 1
In our case f and h are given by (1) and (2) and the stage cost is $l(x_t, u_t, w_t) = x_t^2 + u_t^2 - \gamma^2 w_t^2$. The states, observations and inputs take values in $\mathcal{X} = \mathbb{R}, \mathcal{U} = \mathbb{R}, \mathcal{Y} = \mathbb{R}_{\geq 0}, \mathcal{W} = \mathbb{R}$. The finite-gain condition (4) then correspond to $J_\star(y_0)$ being bounded. If not for the nonlinearity $h(x_t) = |x_t|$, it would be equivalent to the standard dynamic game formulation of $\mathcal{H}_\infty$ suboptimal control [Basar and Bernhard, 2008], rather it can be seen as a special case of nonlinear $\mathcal{H}_\infty$ output feedback control [James and Baras, 1995].

2.2 An information state

Following previous work [Witsenhausen, 1966; Bertsekas and Rhodes, 1973; James and Baras, 1995; Basar and Bernhard, 2008] we consider the "worst-case history", ρ_t, that is compatible with the observations $y_{[0:t-1]}$ and inputs $u_{[0:t-1]}$ up to time $t-1$ reaching the state x at time t:

$$\rho_t(x, y_{[0:t-1]}, u_{[0:t-1]}) \triangleq \sup_{w_{[0:t-1]} \in \mathcal{W}^{t-1}} \sup_{x_0 \in \mathcal{X}} \left\{ \sum_{\tau=0}^{t-1} l(x_\tau, u_\tau, w_\tau) \right. \\ \left. : x_t = x,\ x_{\tau+1} = f(x_\tau, u_\tau, w_\tau),\ y_\tau = h(x_\tau) \right\}. \quad (11)$$

REMARK 2
We follow the convention that the supremum over the empty set is $-\infty$.

The worst-case performance of a policy $\pi \in \Pi$, $J_\pi^N(y_0)$ can be expressed in terms of ρ_t as

$$J_\pi^N(y_0) = \sup_{y_{[0:N]}, x} \rho_{N+1}(x, y_{[0:N]}, u_{[0:N]}), \quad (12)$$

where $u_{[0:N]}$ is generated by π and $y_{[0:N]}$. The functions ρ are causal functions of the measurements and control signals and obey the forward dynamic programming, [Magill, 1965], recursion:

$$\begin{aligned}&\rho_{t+1}(x, y_{[0:t]}, u_{[0:t]})\\&\quad= \sup_{\xi, w} \left\{ l(x, u_t, w) + \rho_t(\xi, y_{[0:t-1]}, u_{[0:t-1]}) : x = f(\xi_t, u_t, w), y_t = h(\xi) \right\}.\end{aligned} \tag{13}$$

Each step (13) involves extremizing over the previous state and the disturbance dependent on the current state x_t, and in general, the computational complexity of evaluating ρ_t grows with t. However, for systems satisfying Assumption 2, the set of feasible past states involved in (13) is restricted by the measurement trajectory. To exploit this restriction, we split the computation of (13) into two steps: a correction step incorporating the observation y_t and a prediction step after selecting u_t:

$$(\xi_t^i)_{i=1}^M = h^{-1}\{y_t\} \tag{14a}$$

$$r_t^i = \rho_t(\xi_t^i, y_{[0:t-1]}, u_{[0:t-1]}) \tag{14b}$$

$$\rho_{t+1}(x, y_{[0:t]}, u_{[0:t]}) = \sup_{i, w \in \mathcal{W}} \{ l(\xi_t^i, u_t, w) + r_t^i : x = f(\xi_t^i, u_t, w) \}. \tag{14c}$$

The intuition behind procedure (14) is that at time t, the realization of the state x_t must belong to the M solutions of $y_t = h(\xi)$. The value r_t^i is the worst-case performance of the system up to time t under the hypothesis that $x_t = \xi_t^i$ consistent with $y_{[0:t]}$ and $u_{[0:t-1]}$. The prediction $\rho_{t+1}(x, y_{[0:t]}, u_{[0:t]})$ is the worst-case performance of the system up to time $t+1$ under the hypothesis that $x_{t+1} = x$ consistent with $y_{[0:t]}$ and $u_{[0:t]}$. The extremization (14c) includes two terms: the stage cost $l(x, u_t, w)$ capturing the cost of transition to $x_{t+1} = x$ from x_t and the past performance r_t^i under the hypothesis $x_t = \xi_t^i$. The extremization is carried out over the hypotheses $\xi_t^1, \ldots, \xi_t^M$ and the disturbance w. Define the update functions g for the M-dimensional vector $\mathbf{r}_t = (r_t^1, \ldots, r_t^M)$

$$\begin{aligned}g_i(\mathbf{r}, y_+, y, u) \triangleq \sup_{j, w \in \mathcal{W}} \Big\{ & l(\xi^j, u, w) + r^j : \xi_+^i = f(\xi^j, u, w),\\ & (\xi_+^k)_{k=1}^M = h^{-1}\{y_+\},\ (\xi^k)_{k=1}^M = h^{-1}\{y\} \Big\}\end{aligned} \tag{15}$$

In the following proposition, we formalize the properties of the update functions g and the sequence $\mathbf{r}$.

Proposition 2

Fix N, $i = 1, \ldots, M$, y_0, a strategy π and let $\mathbf{r}_N$ be defined recursively by $\mathbf{r}_0 = 0$ and $\mathbf{r}_t = g(\mathbf{r}_{t-1}, y_t, y_{t-1}, u_{t-1})$. Then

$$\rho_{t+1}(x, y_{[0:t]}, u_{[0:t]}) = \sup_{i, w} \left\{ l(x, u_t, w) + r_t^i : x = f(\xi_t^i, u_t, w) \right\}.$$

Furthermore,

1. *There exists a sequence* $w_{[0:N-1]}$ *such that* $\max_i r_N^i \geq 0$.
2. *For fixed* $y_t, y_{t-1} \in \mathcal{Y}$ *and* $u_t \in \mathcal{U}$, *for* $\mathbf{r} \leq \mathbf{r}'$ *we have* $g(\mathbf{r}, y_t, y_{t-1}, u_t) \leq g_r(\mathbf{r}', y_t, y_{t-1}, u_t)$.
3. $g(\mathbf{r} + \mathbf{1}c, y_t, y_{t-1}, u_t) = g(\mathbf{r}, y_t, y_{t-1}, u_t) + \mathbf{1}c$ *for all* $c \in \mathbb{R}$.

Proof. *1.* follows directly from Assumption 1. *2.* follows from the monotonicity of the supremum operator. *3.* follows from that for any function f and set Z $\sup_{z\in Z}\{f(z) + c\} = \sup_x\{f(z)\} + c$ for all $c \in \mathbb{R}$. Finally, by recursion, the elements in $\mathbf{r}_t$ are equal to the ones in (14b), thus for each $i = 1, \ldots, M$ equation (14c) holds with $(r_t^i)_{i=1}^M = \mathbf{r}_t$. □

By Proposition 2, the worst-case history ρ_N, is sufficient to evaluate the objective $J_\pi^N(y_0)$. We will now study value iteration to minimize $\sup \rho_N$. Consider the time evolutions of the measurements y and representations $\mathbf{r}$:

$$y_{t+1} = v_t \tag{16a}$$

$$\mathbf{r}_{t+1} = g(\mathbf{r}_t, v_t, y_t, u_t), \quad \mathbf{r}_0 = 0, \tag{16b}$$

where the next measurement, v_t, is considered an exogenous input.

The optimization problem (9) can be expressed in terms of the worst-case history ρ_N as

$$\inf_{\eta} \sup_{N, v_{[0:N-1]}, x\in\mathcal{X}} \left\{\rho_N(x, v_{[0:N-1]}, u_{[0:N-1]})\right\}, \tag{17}$$

where an information-state feedback policy generates u_t

$$u_t = \eta_t(\mathbf{r}_t, y_t).$$

Define the set of information-state strategies $\widetilde{\Pi}$ as the set of strategies $\tilde{\pi} = (\eta_0, \eta_1, \ldots)$. As $\mathbf{r}$ is a causal function of the measurements and control signals, so is η_t (by composition) and $\widetilde{\Pi} \subset \Pi$. In other words, information-state feedback is admissible. The following examples illustrate the information-state recursions (16) for the systems in Examples 1 and 2.

Example 4—continued

In this case, it is convenient to index the hypotheses $h^{-1}\{y_t\}$ by sequences of hypothetical signs, $s_t, \ldots, s_{t-d+1}$ of the d stored measurements $|z_t|, \ldots, |z_{t-d+1}|$. The update simplifies significantly as the realizations of

$z_t, \ldots, z_{t-d+2}$ must remain unchanged between time steps t and $t+1$. Further, $w_t = z_{t+1} + a_1 z_t + \ldots + a_d z_{t-d+1} - b_1 u_t + \ldots + b_d u_{t-d+1}$ is uniquelly determined by the state trajectory, so

$$r_{t+1}^{s_{t+1},\ldots,s_{t-d+2}} = \max_{s_{t-d+1}=\pm 1} \Big\{ l(s_t|z_t|, u_t, w) + r_t^{s_t,\ldots,s_{t-d+1}}$$
$$: w = s_{t+1}|z_{t+1}| + a_1 s_t|z_t| + \ldots + a_d s_{t-d+1}|z_{t-d+1}| - b_1 u_t + \ldots + b_d u_{t-d+1} \Big\}. \tag{18}$$

Example 5—continued

Here, we index the hypotheses $h^{-1}\{y_t\}$ by the matrices A_t, B_t. The update becomes

$$r_{t+1}^{A_{t+1},B_{t+1}}$$
$$= \sup_{A_t,B_t} \Big\{ l(x_t, u_t, x_{t+1} - A_t x_t - B_t u_t) + r_t^{A_t,B_t} : A_{t+1} = A_t, B_{t+1} = B_t \Big\}.$$

By assumption $(A_t, B_t) = (0, 0)$ for all t, so the update simplifies to $r_{t+1}^{A,B} = l(x_t, u_t, x_{t+1} - Ax_t - Bu_t) + r_t^{A,B}$.

2.3 Value iteration

Towards finding (sub)optimal solutions to (9), we introduce the Bellman operators $\mathcal{B}$ and $\mathcal{B}_u$ for functions $V : (\mathbb{R} \cup \{-\infty\})^M \times \mathcal{Y} \to \mathbb{R}$

$$\mathcal{B}\, V(\mathbf{r}, y) = \min_{u \in \mathcal{U}(y)} \overbrace{\max_{v \in \mathcal{Y}} \{V(g(\mathbf{r}, v, y, u), v)\}}^{\mathcal{B}_u\, V(\mathbf{r},y)}. \tag{19}$$

and the value iteration

$$V_0(\mathbf{r}, y) = \max_{i=1,\ldots,M} \{r^i\} \tag{20a}$$
$$V_{k+1}(\mathbf{r}, y) = \mathcal{B}\, V_k(\mathbf{r}, y). \tag{20b}$$

We are ready to state the main theoretical results, justifying the value iteration algorithm (20).

Theorem 1

For the system (6) *under Assumptions 1 and 2 and strategy class* Π, *the value* (9) *is bounded for any* $x_0 \in \mathcal{X}$ *if, and only if, the sequence* $V_0, V_1, \ldots$ *defined in* (20) *is bounded. If bounded, the sequence converges to the optimal value function* $V_\star$. *The limit* $V_\star$ *is a fixed point of the Bellman operator* (19) *and the value* $J_\star(y_0) = V_\star(0, y_0)$*. If the minimum in* (19) *is attained for some* $u \in \mathcal{U}$ *for all* $y \in \mathcal{Y}$ *and* $\mathbf{r}$*, then the policy* $\eta_*(\mathbf{r}, y)$ *defined as the minimizing argument in* (19) *satisfies* $\mathcal{B}_{\eta_\star(\mathbf{r},y)} V_\star(\mathbf{r}, y) = V_\star(\mathbf{r}, y)$ *and the policy*

$$\mu_t(y_{[0:t]}, u_{[0:t-1]}) = \eta_\star(\mathbf{r}_t, y_t)$$

is optimal for (9)*.*

Proof. For any fix $N \geq 0$, the quantity $\inf_{\pi\in\Pi} J^N_\pi(y_0)$ lower bounds $J_\star(y_0)$ due to Assumption 1. By (12),

$$\inf_{\pi\in\Pi} J^N_\pi(y_0) = \inf_{\pi\in\Pi} \sup_{x,w} \rho_{N+1}(x, y_{[0:N]}, u_{[0:N]}) = \inf_{\pi\in\Pi} \sup_{i, v_{[0:N+1]}} \{r^i_{N+1} : v_t \in \mathcal{Y}\}.$$

By standard dynamic programming arguments, see for example [Bertsekas, 2005, Chapter 1.6], this is equal to

$$\begin{aligned}
&\inf_{\mu_{[0:N-1]}} \inf_{\mu_N} \sup_{v_{[0:N-1]}} \sup_{v_N} \{V_0(\mathbf{r}_{N+1}, y_{N+1}) : v_t \in \mathcal{Y}\} \\
&= \inf_{u_0\in\mathcal{U}} \sup_{v_0\in\mathcal{Y}} \cdots \inf_{u_N\in\mathcal{U}} \sup_{v_N} V_0(\mathbf{r}_{N+1}, y_{N+1}) = V_{N+1}(0, y_0).
\end{aligned}$$

This proves that the sequence $V_0(0, y_0), V_1(0, y_0), \ldots$ is bounded if $J_\star(y_0)$ is bounded. By assumption, this holds for all $y_0 \in \mathcal{Y}$.

By induction, the value iteration is non-decreasing as $\mathcal{B}$ is monotone and $V_1 \geq V_0$ follows from assumption 1. Further, $V_k(\mathbf{r}, y) \leq V_k(\max\{r_i\}\mathbf{1}, y) = V_k(0, y) + \max r_i$, proving that $V_0, V_1, \ldots$ is bounded, and since it is monotone increasing, it converges to a limit $V_\star$.

Assume that $V_0, V_1, \ldots$ is bounded towards proving the other direction. Then $V_\star$ is well-defined and satisfies $V_\star \geq V_k$ for all k. Fix an arbitrary $\epsilon > 0$ and define a policy η^ϵ_t that chooses u_t such that

$$\mathcal{B}_{u_t} V_\star(\mathbf{r}_t, y_t) \leq V_\star(\mathbf{r}_t, y_t) + \frac{1}{2}\epsilon \left(\frac{1}{2}\right)^t.$$

By the definition of the infimum, such a u_t always exists. Then, by similar arguments as above, we have

$$J_\star(y_0) \leq \sup_N J^N_{\eta^\epsilon}(y_0) \leq \sup_N V_\star(0, y_0) + \frac{1}{2}\epsilon \sum_{t=0}^{N+1} \left(\frac{1}{2}\right)^t = V_\star(0, y_0) + \epsilon.$$

So we have $V_\star(0, y_0) \leq J_\star(y_0) \leq V_\star(0, y_0) + \epsilon$. As ϵ was arbitrary, we conclude $J_\star(y_0) = V_\star(0, y_0)$. If for any $y \in \mathcal{Y}$ and $\mathbf{r}$, the minimum in (19) is attained for some $u \in \mathcal{U}$, then we can pick $\epsilon = 0$ and conclude that the minimizing argument in (19) is optimal for (9). □

Theorem 2—Approximation

For the system (6) *under Assumptions 1 and 2 and strategy class* Π*, assume that there exists a function* $\bar{V} : (\mathbb{R} \cup \{-\infty\})^M \times \mathcal{Y} \to \mathbb{R}$ *and a strategy* $\bar{\pi} = (\bar{\eta}, \bar{\eta}, \dots) \in \widetilde{\Pi}$ *such that* $\bar{V} \geq V_0$ *and*

$$\mathcal{B}_{\bar{\eta}(\mathbf{r},y)} \bar{V}(\mathbf{r}, y) \leq \bar{V}(\mathbf{r}, y).$$

Then the value iteration $V_0, V_1, \dots$ *is bounded, and* $J_{\bar{\mu}}(y_0) \leq \bar{V}(0, y_0)$ *for the policy*

$$\bar{\mu}_t(y_{[0:t]}, u_{[0:t-1]}) = \bar{\eta}(\mathbf{r}_t, y_t).$$

Proof. By monotonicity of the Bellman operator, we have that $V_k \leq \bar{V}$ for all $k = 0, 1, \dots$, implying that the value iteration $V_0, V_1, \dots$ is bounded. Further,

$$\begin{aligned} J_\star(y_0) &\leq \sup_N J_{\bar{\pi}}^N(y_0) = \sup_N \sup_{v_{[0:N]}} \{V_0(\mathbf{r}_{N+1}, y_{N+1}) : v_t \in \mathcal{Y}\} \\ &\leq \sup_N \sup_{v_{[0:N]}} \{\bar{V}(\mathbf{r}_{N+1}, y_{N+1}) : v_t \in \mathcal{Y}\} \leq \sup_N \bar{V}(\mathbf{r}_0, y_0) = \bar{V}(\mathbf{r}_0, y_0). \end{aligned}$$

□

3. Magnitude control

We now apply the above results to the example in Section 1. For any y, we denote $\xi^+ = y$ and $\xi^- = -y$. Then $h^{-1}\{y\} = \{\xi^+, \xi^-\}$. We similarly index $\mathbf{r} = [r^+, r^-]$. Then g in (15) becomes $g_s(\mathbf{r}, v, y, u) = y^2 + u^2 - \gamma^2 \min\{r^+ + (sv - u - y)^2, r^- + (sv - u + y)^2\}$ for $s = \pm 1$.

Proof of proposition 1 By the above analysis, the quantities (5) correspond to (16), and the first statement in the proposition is a direct consequence of Theorem 1. Drawing inspiration from [Rantzer, 2021], we parameterize an upper bound of the optimal value in the parameters $0 < p \leq q < \gamma^2$ by

$$\bar{V}(\mathbf{r}, y) = \max\{py^2 + r^+, py^2 + r^-, qy^2 + (r^+ + r^-)/2\}, \tag{21}$$

and a certainty equivalence policy

$$\bar{\eta}(\mathbf{r}, y) = k \,\mathrm{sign}(r^- - r^+) y \tag{22}$$

The following lemma relates the parameters of the value function approximation p, q and k to the ℓ_2 gain of the closed loop.

LEMMA 1

Given a quantity $\gamma > 0$, parameters $0 < p < q < \gamma^2$, $k \in \mathbb{R}$, and $\bar{V}$ as above. The certainty equivalency policy $\bar{\eta}$ in (22) achieves an ℓ_2-gain of at most γ for the system (1)–(3) and an objective value smaller than $\bar{V}(0, |x_0|)$ for the decision problem (9), if

$$\begin{aligned} p &> 1 + k^2 + \frac{(1-k)^2}{p^{-1} - \gamma^{-2}}, \\ q &> 1 + k^2 + \frac{(1+k)^2}{p^{-1} - \gamma^{-2}}, \\ q &> 1 + k^2 + \frac{1}{q^{-1} - \gamma^{-2}} - \gamma^2 k^2. \end{aligned} \tag{23}$$

The values $\gamma = 4$, $p = 1.7$, $q = 7$ and $k = 0.7$ satisfy the conditions of Theorem 1 and a simulation with $w_t = sin(\pi t/10)$ is shown in Figure 2.

Proof. Define

$$\begin{aligned} \bar{V}^{++}(\mathbf{r}, y) &\triangleq py^2 + r^+ \\ \bar{V}^{--}(\mathbf{r}, y) &\triangleq py^2 + r^- \\ \bar{V}^{+-}(\mathbf{r}, y) &\triangleq qy^2 + (r^+ + r^-)/2. \end{aligned} \tag{24}$$

Then, for a fixed u, we have

$$\begin{aligned} &\mathcal{B}_u\{\max\{\bar{V}^{++}(\mathbf{r}, y), \bar{V}^{--}(\mathbf{r}, y))\} \\ &\qquad = y^2 + u^2 + \max\left\{\frac{(y+u)^2}{p^{-1} - \gamma^{-2}} + r^+, \frac{(y-u)^2}{p^{-1} - \gamma^{-2}} + r^-\right\}. \end{aligned}$$

Define for $i, j \in \{+, -\}$

$$\alpha^{ij} \triangleq \sup_{v \geq 0}\{qv^2 - \frac{\gamma^2}{2}\left((v - u - iy)^2 + (-v - u - jy)^2\right) + \frac{r^i + r^j}{2}.$$

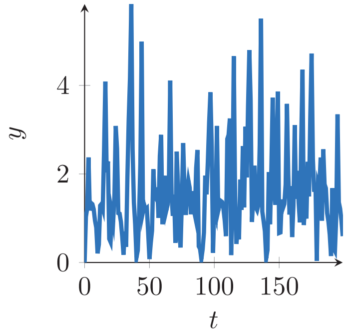
Then $\mathcal{B}_u\{\bar{V}^{+-}(\mathbf{r}, y)\} = y^2 + u^2 + \max_{i,j \in \{+,-\}} \alpha^{ij}$, where, for $i \neq j$,

$$\alpha^{ii} = r^i - \gamma^2(u + iy)^2, \quad \max\{\alpha^{ij}, \alpha^{ji}\} = \frac{y^2}{q^{-1} - \gamma^{-2}} - \gamma^2 u^2 + \frac{r^i + r^j}{2}$$

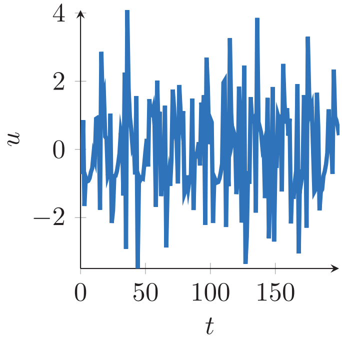
Let $l = \arg\max_{i \in \{+,-\}}\{r^i\}$, then by (23), we have

$$\begin{aligned} &\mathcal{B}_{-k_l y}\{\bar{V}(\mathbf{r}, y)\} = \max_{ij}\{\mathcal{B}_{-k_l y}\, \bar{V}^{ij}(\mathbf{r}, y)\} \\ &\qquad \leq \max\{\mathcal{B}_{-k_l y}\, \bar{V}^{il}(\mathbf{r}, y) - (r^l - r^j)/2\} \leq \max\{\bar{V}^{il}(\mathbf{r}, y)\} \leq \bar{V}(\mathbf{r}, y). \end{aligned}$$

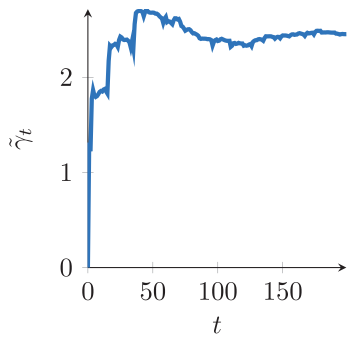
Therefore, by Theorem 2, the objective value is bounded from above by $\bar{V}(0, y_0)$. $\square$



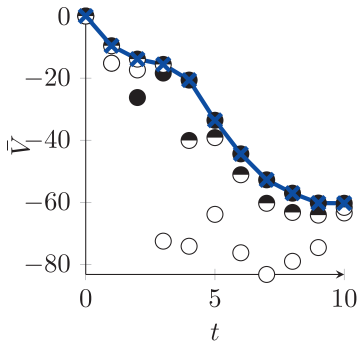
(a) Output



(b) Control signal



(c) Empirical gain



(d) Value function

Figure 2. Figures 2a and 2b contain plots of the outputs and control signal, respectively. Figure 2c shows the empirical gain from w to (y, u) and Figure 2d shows the value function approximation $\bar{V} = \max\{\bar{V}^{++}, \bar{V}^{+-}, \bar{V}^{--}\}$ defined in (21). The black marks corresponds to $\bar{V}^{++}$, the half circles to $\bar{V}^{+-}$, the white marks to $\bar{V}^{--}$ and the blue crosses to $\bar{V}$. Note that the value function approximation is monotonically decreasing.

4. Conclusion

This article demonstrated that output feedback minimax dual control possesses a finite-dimensional information state when the measurement equation has a finite number of solutions. We applied this finding to the magnitude control of an integrator, resulting in a surprisingly simple sub-optimal control policy. The controller is proportional, with the gain determined through hypothesis testing and updated online. However, the results are limited to cases where the measurement equation has a finite number of solutions. This restriction excludes scenarios where measurements are affected by real-valued sensor noise, which typically leads to an infinite-dimensional information state.

Future work will focus on extending these results to cases with noisy measurements, specifically where the dynamics are linear and uncertain but belong to a finite set. Progress has already been made for scalar systems [Kjellqvist and Rantzer, 2022], and the extension to multi-dimensional cases is currently under investigation.

Acknowledgements

I was introduced to the problem of output feedback from magnitude measurements by Bo Bernhardsson at the start of my Ph.D. studies. This manuscript has benefited from lengthy discussions with the author's colleagues at the Department of Automatic Control. In particular, I want to thank Anders Rantzer and Tore Hägglund for their encouragement and feedback along the way, Venkatraman Renganathan for reviewing an earler version of this manuscript and Jonas Hansson for tackling my conjecture with much enthusiasm.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 834142 (ScalableControl). The author is a member of the ELLIIT Strategic Research Area at Lund University.

References

Allison, B. J., J. E. Ciarniello, P. J.-C. Tessier, and G. A. Dumont (1995). "Dual adaptive control of chip refiner motor load". *Automatica* **31**:8, pp. 1169–1184. ISSN: 0005-1098. DOI: 10.1016/0005-1098(95)00030-Z.

Alspach, D. (1972). "Dual control based on approximate a posteriori density functions". *IEEE Trans. Autom. Control* **17**:5, pp. 689–693. DOI: 10.1109/TAC.1972.1100099.

Anderson, B. and A. Dehghani (2008). "Challenges of adaptive control-past, permanent and future". *Annu. Rev. Control* **32**, pp. 123–135. DOI: 10.1016/j.arcontrol.2008.06.001.

Åström, K. and A. Helmersson (1986). "Dual control of an integrator with unknown gain". *Comput. Math. Appl.* **12**:6, Part A, pp. 653–662. ISSN: 0898-1221. DOI: 10.1016/0898-1221(86)90052-0.

Basar, T. and P. Bernhard (2008). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach.* 2nd ed. Birkhauser, Boston, MA, USA. DOI: 10.1007/978-0-8176-4757-5.

Bertsekas, D. and I. Rhodes (1973). "Sufficiently informative functions and the minimax feedback control of uncertain dynamic systems". *IEEE Trans. Autom. Control* **18**:2, pp. 117–124. DOI: 10.1109/TAC.1973.1100241.

Bertsekas, D. P. (2005). *Dynamic Programming and Optimal Control Vol. 1.* 3rd. Vol. I. Athena Scientific, Belmont, MA, USA.

Chen, X. and E. Hazan (2021). "Black-box control for linear dynamical systems". In: Belkin, M. et al. (Eds.). *Proc. 34th Conf. Learn. Theory.* Vol. 134. PMLR, pp. 1114–1143. URL: https://proceedings.mlr.press/v134/chen21c.html.

Dean, S., H. Mania, N. Matni, B. Recht, and S. Tu (2018). "Regret bounds for robust adaptive control of the linear quadratic regulator". In: *Adv. Neural Inf. Process. Syst.* URL: https://api.semanticscholar.org/CorpusID:43946868.

Didinsky, G. and T. Basar (1994). "Minimax adaptive control of uncertain plants". In: *Proc. 33rd IEEE Conf. Decis. Control.* Vol. 3, pp. 2839–2844. DOI: 10.1109/CDC.1994.411368.

Dumont, G. and K. Åström (1988). "Wood chip refiner control". *IEEE Control Syst. Mag.* **8**, pp. 38–43. DOI: 10.1109/37.1872.

Goodwin, G. C. and K. S. Sin (2009). *Adaptive Filtering Prediction and Control.* Dover Publications, Inc., USA. ISBN: 0486469328.

Hazan, E. (2023). *Introduction to online convex optimization.* arXiv: 1909.05207 [cs.LG].

James, M. and J. Baras (1995). "Robust $\mathcal{H}_\infty$ output feedback control for nonlinear systems". *IEEE Trans. Autom. Control* **40**:6, pp. 1007–1017. DOI: 10.1109/9.388678.

Khalil, H. K. (2002). *Nonlinear systems; 3rd ed.* Prentice-Hall, Upper Saddle River, NJ. URL: https://cds.cern.ch/record/1173048.

Kjellqvist, O. (2024). *Minimax dual control with finite-dimensional information state.* DOI: 10.48550/arXiv.2312.05156.

Kjellqvist, O. and A. Rantzer (2022). "Learning-enabled robust control with noisy measurements". In: Firoozi, R. et al. (Eds.). *Proc. 4th Annu. Learning Dyn. Control Conf.* Vol. 168. PMLR, pp. 86–96. URL: https://proceedings.mlr.press/v168/kjellqvist22a.html.

Magill, D. (1965). "Optimal adaptive estimation of sampled stochastic processes". *IEEE Trans. Autom. Control* **10**:4, pp. 434–439. DOI: 10.1109/TAC.1965.1098191.

Mania, H., S. Tu, and B. Recht (2019). "Certainty equivalence is efficient for linear quadratic control". In: Wallach, H. et al. (Eds.). *Adv. Neural Inf. Process. Syst.* Vol. 32. URL: https://proceedings.neurips.cc/paper_files/paper/2019/file/5dbc8390f17e019d300d5a162c3ce3bc-Paper.pdf.

Rantzer, A. (2021). "Minimax adaptive control for a finite set of linear systems". In: *Proc. 3th Annu. Learning Dyn. Control Conf.* Vol. 144. Corrected version: https://arxiv.org/abs/2011.10814, pp. 893–904. URL: https://proceedings.mlr.press/v144/rantzer21a.html.

Renganathan, V., A. Iannelli, and A. Rantzer (2023). "An online learning analysis of minimax adaptive control". In: *Proc. 62nd IEEE Conf. Decis. Control*, pp. 1034–1039. DOI: 10.1109/CDC49753.2023.10384114.

Rosdahl, C. and B. Bernhardsson (2020). "Dual control of linear discrete-time systems with time-varying parameters". In: *Proc. 2020 Int. Conf. Control Automat. Diagnosis*, pp. 1–6. DOI: 10.1109/ICCAD49821.2020.9260540.

Witsenhausen, H. (1966). *Minimax Control of Uncertain Systems.* Massachusetts Institute of Technology. URL: https://books.google.se/books?id=PCxMnQEACAAJ.

Wittenmark, B. (1995). "Adaptive dual control methods: an overview". In: *Proc 5th IFAC Symp. Adaptive Syst. Control Signal Process.* Vol. 28. 13, pp. 67–72. DOI: 10.1016/S1474-6670(17)45327-4.

Yu, J., D. Ho, and A. Wierman (2023). "Online adversarial stabilization of unknown networked systems". *Proc. ACM Meas. Anal. Comput. Syst.* **7**, pp. 1–43. DOI: 10.1145/3579452.

Paper V

Output Feedback Minimax Adaptive Control

Olle Kjellqvist and Anders Rantzer

Abstract

This paper formulates adaptive controller design as a minimax dual control problem. The objective is to design a controller that minimizes the worst-case performance over a set of uncertain systems. The uncertainty is described by a set of linear time-invariant systems with unknown parameters. The main contribution is a common framework for both state feedback and output feedback control. We show that for finite uncertainty sets, the minimax dual control problem admits a finite-dimensional information state. This information state can be used to design adaptive controllers that ensure that the closed-loop has finite gain. The controllers are derived from a set of Bellman inequalities that are amenable to numerical solution. The proposed framework is illustrated on a challenging numerical example.

1. Introduction

This paper addresses the design of adaptive controllers with guaranteed performance for linear time-invariant systems with uncertain parameters. The performance index is quadratic with a soft constraint on the size of the disturbance. The performance measure quantifies transient behavior and, if finite, guarantees a bounded ℓ_2-gain. Via the small-gain theorem, we guarantee stability in the presence of unmodeled dynamics.

This property implies that the closed-loop system behaves well even if the assumptions on the model class are violated, as long as the violation is minor. Toward this end, we do not make any assumptions on the statistical

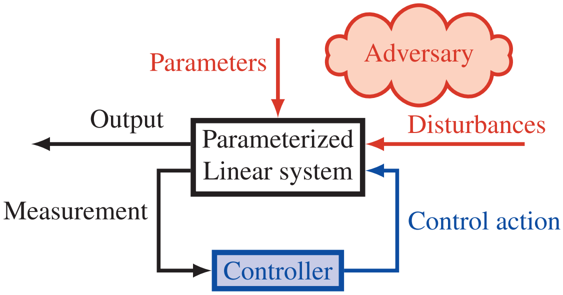


Figure 1. The minimax control problem. The controller minimizes a performance index by selecting inputs, while the adversary selects the parameter realization and disturbances to maximize it.

properties of the parameters or the exogenous signals. Instead, the underlying models are deterministic, and the uncertain parameters and signals are chosen by an adversary that seeks to maximize the performance index. See Fig. 1 for an illustration of the problem.

1.1 Contributions

To address the complexities and challenges outlined above, this paper makes several key contributions to the field of adaptive control.

Unifying State-Feedback and Output-Feedback We show that the state-feedback and output-feedback minimax dual control problems can be reduced to a minimax control problem with linear (known) dynamics and uncertain objective functions. This reduction is based on the concept of information-state feedback[Bertsekas and Rhodes, 1973; James and Baras, 1995], and is illustrated in Fig. 2. The problem with uncertain objective functions is introduced in Section 2.1. State-feedback and output-feedback minimax dual control problems and their reductions are presented in Sections 2.2 and 2.3, respectively.

Finite-Dimensional Information State We show that if the uncertain parameters belong to a finite set, the optimal output-feedback controller is observer-based and can be computed by dynamic programming. This is a specialization of the result in [James and Baras, 1995] for the nonlinear $\mathcal{H}_\infty$-control problem. However, in contrast to the general result [James and Baras, 1995], we show in Section 2.3 that the observer state is finite-dimensional, and we provide a constructive method to compute the observer state.

Heuristics for Suboptimal Controller Synthesis We provide a heuristic method to synthesize suboptimal controllers. The method is based on

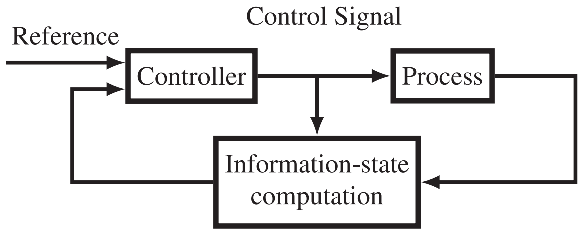


Figure 2. Closed-loop with information state-feedback control.

approximating the value function as a piecewise-quadratic function and the controllers as certainty-equivalence controllers. We introduce periodic Bellman inequalities to deal with delays and other nonminimum-phase behavior. The method generalizes [Rantzer, 2021, Theorem 3] to minimax control of linear time-invariant systems with unknown objective functions. Hence, the method is applicable to both state-feedback and output-feedback minimax dual control problems. The resulting controllers have guaranteed worst-case performance in the sense of a bounded ℓ_2-gain. The details of the method are presented in Section 3, and the use of periodicity to deal with nonminimum-phase behavior is examplified in Section 4.1.

Numerical Examples We provide a `Julia` [Bezanson et al., 2017] implementation[1] of the proposed methods and design a controller that simultaneously stabilizes G_{mp} and G_{nmp},

$$G_{\text{mp}}(z) = \frac{z_0 z - 1/z_0}{(z-1)^2}, \tag{1a}$$

$$G_{\text{nmp}}(z) = \frac{z/z_0 - z_0}{(z-1)^2}. \tag{1b}$$

Here z is the complex frequency, and z_0 is 1.01. Both systems are unstable, with a double pole in $1/(z-1)$. G_{mp} has a minimum phase zero at $1/z_0^2$ and G_{nmp} has a nonminimum-phase zero at z_0^2. The results are presented in Section 4.2.

1.2 Related work

Minimax Control Witsenhausen introduced minimax control in his thesis [Witsenhausen, 1966] as a decision-theoretic approach to control of uncertain dynamical systems. Bertsekas and Rhodes [Bertsekas and Rhodes, 1973] showed that the optimal controller can be decomposed into an estimator and an actuator. The optimal estimator can be expressed as a function of the observations, a so-called "sufficiently informative function". $\mathcal{H}_\infty$-control, both

[1] The code is available at `https://github.com/kjellqvist/MinimaxAdaptiveControl.jl`.

in the linear-quadratic case [Basar and Bernhard, 2008] and in the nonlinear case [James and Baras, 1995], has been formulated in terms of minimax control. Recently, Goel and Hassibi [Goel and Hassibi, 2023] showed that minimizing the regret or competitive ratio compared to an acausal controller with access to the future disturbance trajectory may provide excellent nominal performance at only a small robustness expense and that optimal controller synthesis can be reformulated as standard $\mathcal{H}_\infty$ controller synthesis. Karapetyan et. al [Karapetyan et al., 2022] studied the suboptimality of $\mathcal{H}_\infty$ control in the same setting.

The term “minimax adaptive” was introduced in [Didinsky and Basar, 1994] as a term for controllers that minimize the worst-case performance realization for systems with parametric uncertainty. The authors considered continuous-time state-feedback control of systems with uncertain but constant parameters and showed that the cost can be rewritten in terms of the least-squares estimate of the parameters. The reformulation using the least-squares estimate of the parameters is viable in the state-feedback case because it is sufficient to reconstruct the worst-feasible realization consistent with a model hypothesis and data, which James and Baras [James and Baras, 1995] showed is an information state—or informative statistic in Striebel's terms [Striebel, 1965]. This information state has a recursive formulation and is generally not finite-dimensional. This statistic corresponds to our function r in (6) and has a finite-dimensional representation. Pan and Başar generalized the results to of nonlinear SISO systems on “parametric strict-feedback form” in [Pan and Basar, 1998].

Vinnicombe [Vinnicombe, 2004] studied scalar systems where the parameters' signs are unknown and provided an explicit suboptimal controller based on certainty equivalence control with the least-squares parameter estimate. Megretski and Rantzer [Megretski and Rantzer, 2003] provides lower bounds on the achievable ℓ_2-gain for scalar systems with an uncertain pole belonging to an interval. Rantzer extended Vinnicombe's result to higher-order systems where the state matrix has an unknown sign [Rantzer, 2020] and to finite sets of linear systems assuming full state measurements [Rantzer, 2021]. A sufficient condition for finite ℓ_2-gain is formulated as bilinear matrix inequalities, and a controller is obtained from the solution.

Cederberg *et al.* [Cederberg et al., 2022] proposed linearizing Rantzer's inequalities to improve performance iteratively, Bencherki and Rantzer [Bencherki and Rantzer, 2023] gave conditions under which a solution to the inequalities is guaranteed to exist and Renganathan *et al.* [Renganathan et al., 2023] studied empirical performance. Kjellqvist generalized the framework to nonlinear systems where the preimage of the output under the measurement function is a finite set [Kjellqvist, 2024] assuming noise-free measurements. Kjellqvist and Rantzer [Kjellqvist and Rantzer, 2022] previously extended Vinnicombe's [Vinnicombe, 2004] controller to the one-dimensional output-feedback case.

Dual Control The controllers in Section 3 are dual controllers. In the nomenclature of Filatov and Unbehauen's survey[Filatov and Unbehauen, 2000], they are *implicit* dual controllers, as they are suboptimal solutions to dual control problems. Duality here is in the sense of Feldbaum's observation: that optimal controllers for uncertain nonlinear systems tend to have both regulating and experimenting mechanisms [Feldbâum, 1963]. This duality is known as the *exploration-exploitation trade-off* in the reinforcement learning literature [Sutton and Barto, 2018]. For further reading on dual control, see the surveys by Wittenmark [Wittenmark, 1995], Filatov and Unbehauen [Filatov and Unbehauen, 2000] and Mesbah [Mesbah, 2018].

Supervisory Control & Multiple-Model Adaptive Control Supervisory control, or multiple-model adaptive control, is a controller architecture where a supervisor selects a controller from a set of candidate controllers [Hespanha, 2001]. Supervisory controllers typically come in two flavors: *Estimator-based*, where each model has an associated estimator and control law, and the supervisor selects the model based on the estimator's output [Buchstaller and French, 2016], and *controller-based*, where the supervisor selects among control laws by disqualifying controllers that violates assumed performance guarantees [Safonov and Tsao, 1997; Patil et al., 2022]. Our certainty-equivalence controllers can be seen as an instance of the former.

Switching, even among stable subsystems, may induce instability[Liberzon, 2003], so the supervisor must ensure that the switching is safe. In estimator-based frameworks, this typically translates to dwell-time constraints. In the controller-based framework, this can be achieved by hysteresis in switching out underperforming controllers [Battistelli et al., 2010]. This switching restriction relates to the separation of time scales exploited in Ljung's averaging arguments [Ljung, 1977], and to the difficulties of fast adaptation [Anderson, 2005]. The periodicity in our certainty-equivalence controllers can be seen as a form of dwell-time constraint.

Learning Recently, there has been a surge of interest in the intersection of learning theory and control. Advances in, for example, high-dimensional statistics [Tsiamis et al., 2023] and online convex optimization [Hazan, 2016] have provided tools for the design of new adaptive algorithms and analysis of achievable performance. Most work has focused on relating the asymptotic scaling of performance bounds to assumptions on the size and number of uncertain parameters. It is based on assumptions about the statistics of the exogenous signals. We now provide a brief overview of the works most closely related to our own.

Agarwal *et al.* [Agarwal et al., 2019] considered control of linear systems with known dynamics, where the control objective and disturbances were adversarially chosen, as in our Problem 1. In contrast to our work, the cost functions were time-varying, revealed sequentially after actuation, and did not include a disturbance term. Instead of a soft constraint, they assumed that the disturbance was bounded.

Simchowitz *et al.* [Simchowitz, 2020] extended the results to output feedback and uncertain dynamics but relied on apriori knowledge of a stabilizing static output feedback controller and evaluations of the objective function.

Ghai *et al.* [Ghai et al., 2022] considered online control with model misspecification assuming perfect state measurements and provided an adaptive controller with bounded ℓ_2-gain. The bound is asymptotic and scales with the number of uncertain parameters and the size of the model mismatch. In contrast our ℓ_2-bound is specific to the problem instance, and the user specifies precisely what parameters are uncertain. Lee *et al.* [Lee et al., 2024] studied the regret of certainty-equivalence controllers with normally distributed exploration for approximately parameterized linear systems. The approximation allows the user to inject prior knowledge and identify a reduced number of parameters, but means that the true system lies outside the space spanned by the parameters. They showed an improvement over black-box adaptation in the small-data regime, as long as the model misspecification and the number of parameters are few.

1.3 Notation

The set of $n \times m$ matrices with real coefficients is denoted $\mathbb{R}^{n\times m}$. The transpose of a matrix A is denoted $A^\top$. The space of real symmetric matrices in $\mathbb{R}^{n\times n}$ is denoted $\mathbb{S}^n$. For a symmetrix matrix $H \in \mathbb{S}^{n+m}$ with blocks

$$H = \begin{bmatrix} H^{11} & H^{12} \\ H^{21} & H^{22} \end{bmatrix},$$

we denote the Schur complements of H^{11} and H^{22} in H by

$$\begin{aligned} H/H^{22} &= H^{11} - H^{12}H^{-22}H^{21}, \\ H/H^{11} &= H^{22} - H^{21}H^{-11}H^{12}. \end{aligned}$$

H^{-ii} denotes the inverse of H^{ii}, if it exists. For a symmetric matrix $H \in \mathbb{S}^{n_1+\cdots n_m}$ with blocks

$$H = \begin{bmatrix} H^{11} & \cdots & H^{1m} \\ \vdots & \ddots & \vdots \\ H^{m1} & \cdots & H^{mm} \end{bmatrix},$$

and vectors $x_1 \in \mathbb{R}^{n_1}, \ldots, x_m \in \mathbb{R}^{n_m}$, we define the quadratic form

$$\sigma_H(x_1, \ldots, x_m) = \sum_{i,j} x_i^\top H^{ij} x_j.$$

We write $H \succeq 0$ to indicate that H is positive semidefinite and $H \succ 0$ to indicate that H is positive definite. If $H \succ (\succeq) 0$, we sometimes write $|x|_H^2$ to emphasize that $|x|_H^2 = \sigma_H(x)$ is a (semi) norm. The standard euclidean norm is denoted $|x| = \sqrt{\sigma_I(x)}$, and by extension $|(x_1, \ldots, x_m)|^2 = \sum_{i=1}^m |x|^2$. We refer to the value of a signal w at time t as w_t and use the shorthand notation $w_{[0:t]}$ for the sequence $(w_0, w_1, \ldots, w_t)$. We sometimes use asterisks in matrix expressions to denote elements implied by symmetry. For two sets, X and Y, we denote the set of functions from X to Y by Y^X.

2. Exact Analysis

2.1 Principal Problem

This section introduces the principal problem of this paper and presents theory on minimax dynamic programming and value iteration. In Sections 2.2 and 2.3, we show how to reduce state feedback and output feedback adaptive control to the principal problem.

Problem 1—Principal problem

Let $\hat{A} \in \mathbb{R}^{n_z \times n_z}$, $\hat{B} \in \mathbb{R}^{n_z \times n_u}$, $\hat{G} \in \mathbb{R}^{n_z \times n_d}$, $z_0 \in \mathbb{R}^{n_z}$ and let $\mathcal{M} \subset \mathbb{S}^{n_z+n_u+n_d}$ be a compact set whose members, H, satisfy

$$H = \begin{bmatrix} H^{zz} & H^{zu} & H^{zd} \\ H^{uz} & H^{uu} & H^{ud} \\ H^{dz} & H^{du} & H^{dd} \end{bmatrix}, \qquad \begin{aligned} H^{dd} &\prec 0, \\ H/H^{dd} &\succeq 0, \\ [H/H^{dd}]^{uu} &\succ 0. \end{aligned} \tag{2}$$

Compute

$$\inf_{\mu} \sup_{H \in \mathcal{M}, d, N} \underbrace{\sum_{t=0}^{N-1} \sigma_H(z_t, u_t, d_t)}_{\hat{J}_\mu^N(z_0, H, d)}, \tag{3}$$

where $N \geq 0$ and the sequences, $z_{[0:N]}$, $u_{[0:N-1]}$ are generated by

$$z_{t+1} = \hat{A} z_t + \hat{B} u_t + \hat{G} d_t, \quad t \geq 0 \tag{4}$$

$$u_t = \mu_t(z_{[0:t]}, u_{[0:t-1]}, d_{[0:t-1]}). \tag{5}$$

Problem 1 concerns the upper value of a two-player zero-sum game, where the minimization is over the controller, μ, and the maximization is over the disturbance, d, and the realization of the cost function, $H \in \mathcal{M}$. If not for the uncertainty in the cost function, the problem would be a (nonstandard) linear-quadratic control problem, which is a well understood problem class [Basar and Bernhard, 2008].

The relation to adaptive control is as follows. In state-feedback adaptive control, with dynamics of the form $x_{t+1} = Ax_t + Bu_t + w_t$, where x_t is the state and w_t is the disturbance and the pair (A, B) is unknown, and quadratic stage costs, we let $d_t = x_{t+1}$ and $z_t = x_t$. Substituting $w_t = d_t - Az_t - Bu_t$ and z_t into the cost function gives dynamics of the form (4) and cost functions of the form (3). This is explained in more detail in Section 2.2.

For output-feedback adaptive control with a finite set of feasible models and quadratic stage costs, we quantify the worst-case accrued cost using one observer for each model. The z_t of Problem 1 is constructed by stacking the observer states, d_t is the measured output, and u_t is the control input. The matrices $\hat{A}$, $\hat{B}$ and $\hat{G}$ corresponds to aggregating the observer dynamics. We get one Hessian, H, for each model, expressing the past performance of the observer. The reformulation of output-feedback adaptive control as an instance of Problem 1 is explained in Section 2.3. The rest of this section is devoted to dynamic programming and value iteration.

Dynamic Programming Define the functions $r_t : \mathcal{M} \to \mathbb{R}$ for $t = 0, 1, 2, \ldots$ by

$$r_t(H) = \sum_{s=0}^{t-1} \sigma_H(z_s, u_s, d_s). \tag{6}$$

Then r_t satisfies the recursion

$$r_{t+1}(H) = \underbrace{r_t(H) + \sigma_H(z_t, u_t, d_t)}_{f(r_t, z_t, u_t, d_t)(H)}. \tag{7}$$

Although the controller does not know the realization of H, the functions r_t are constructed of known quantities and can be computed by the controller at time t.

Remark 1

The functions r_t take the form $r_t(H) = \langle Z_t, H\rangle$. The positive semidefinite matrix Z_t can be computed recursively by

$$Z_{t+1} = Z_t + \begin{bmatrix} z_t \\ u_t \\ d_t \end{bmatrix} \begin{bmatrix} z_t \\ u_t \\ d_t \end{bmatrix}^\top, \quad Z_0 = 0.$$

This means that the matrix Z_t compresses the information of the past states, inputs, and disturbances into a single matrix. If the cardinality of $\mathcal{M}$ is large *compared to the state dimension, n_z, then the matrix Z_t can be used to reduce the computational complexity of the problem. If the model set is finite, then one can store the function values $r_t(H)$ for each $H \in \mathcal{M}$ in an array.*

For a function $V : \mathbb{R}^{n_z} \times \mathbb{R}^{\mathcal{M}} \to \mathbb{R}$, define the Bellman operators

$$\mathcal{B}_u V(z,r) = \max_d V\left(\hat{A}z + \hat{B}u + \hat{G}d, f(r,z,u,d)\right), \tag{8a}$$

$$\mathcal{B} V(z,r) = \min_u \mathcal{B}_u V(z,r), \tag{8b}$$

and the value iteration

$$V_0(z,r) = \max_{H \in \mathcal{M}} r(H), \tag{9a}$$

$$V_{k+1}(z,r) = \mathcal{B} V_k(z,r). \tag{9b}$$

We will consider control policies $\eta_t : \mathbb{R}^{n_z} \times \mathbb{R}^{\mathcal{M}} \to \mathbb{R}^{n_u}$ of the form

$$u_t = \eta_t(z_t, r_t), \tag{10}$$

and note that this policy is admissible as r_t depends causally on the states, inputs, and measured disturbances.

THEOREM 1
The following facts holds for Problem 1, the Bellman operator $\mathcal{B}$ in (8b) and the value iteration defined in (9).

1. *$\mathcal{B}$ is monotone: $V' \geq V \implies \mathcal{B} V' \geq \mathcal{B} V$.*
2. *The value iteration is nondecreasing: $V_{k+1} \geq V_k$.*
3. *The value iteration converges if, and only if, it is bounded.*
4. *The value (3) is bounded for all $z_0 \in \mathbb{R}^{n_z}$ if, and only if, the value iteration converges.*

If the value iteration converges to a limit $V_\star$, then

5. *The value (3) is equal to $V_\star(z_0, 0)$.*
6. *$V_\star$ is a fixed point of $\mathcal{B}$, not necessarily unique.*
7. *$V_\star$ is the minimal fixed point of $\mathcal{B}$ greater than V_0.*
8. *The control law $\eta_\star$ defined as the minimizer in (8b) achieves $\mathcal{B}_{\eta_\star(x,r)} V_\star(x,r) = \mathcal{B} V(x,r)$, and the policy*

$$\mu_t(z_{[0:t]}, u_{[0:t-1]}, d_{[0:t-1]}) = \eta_\star(z_t, r_t)$$

is optimal for Problem 1.

Proof. The proofs of the statements in the theorem are standard but we include them here for completeness.

1. As the maximization in (8a) and minimization in (8b) are monotone operations, so is their composition $\mathcal{B}$.

2. Assume that $V_k \geq V_{k-1}$. By monotonicity of $\mathcal{B}$, $V_{k+1} = \mathcal{B}V_k \geq \mathcal{B}V_{k-1} = V_k$. We now consider the base case, and prove that $V_1 \geq V_0$. By the minmax inequality

$$\begin{aligned} V_1(z, r) &= \min_u \max_d V_0(\hat{A}z + \hat{B}u + \hat{G}d, f(r, z, u, d)) \\ &\geq \max_{H \in \mathcal{M}} \min_u \max_d \Big\{ \sigma_H(\hat{A}z + \hat{B}u + \hat{G}d, u, d) \\ &\quad + r(H) \Big\} \end{aligned}$$

   By (2), $\max_d \sigma_H(z, u, d) \geq 0$ for all $H \in \mathcal{M}$, we have that $V_1(z, r) \geq \max_{H \in \mathcal{M}} r(H) = V_0(z, r)$.

3. Pointwise convergence of the value iteration is equivalent to convergence of the monotone sequence of real numbers $V_0(z, r), V_1(z, r), \ldots$.

We first show that $V_k(z, r)$ converges if, and only if $V_k(z, 0)$ converges. Let $a \in \mathbb{R}$ be a constant. By induction $V_k(z, r + a) = V_k(z, r) + a$. Further, as f and V_0 are monotone in r, so is V_k: $r' \geq r \implies V_k(z, r') \geq V_k(z, r)$. Thus, for all $k = 1, 2 \ldots$

$$V_k(z, 0) + \min\{r\} \leq V_k(z, r) \leq V_k(z, 0) + \max\{r\}.$$

For any N, standard dynamic programming arguments show that $V_N(z_0, 0) = \inf_\mu \sup_{d, H \in \mathcal{M}} \hat{J}_\mu^N(z_0, H, d)$, which is a lower bound for (3). Thus (3) is bounded *only if* the value iteration is bounded.

Now, fix an $\epsilon > 0$ and let $\tilde{\eta}_t$ be a policy that achieves $\mathcal{B}_{\eta_t} V_\star(z, r) \leq V_\star(z, r) + \epsilon^{t+1}$. Then (3) is bounded from above by the expression

$$\begin{aligned} \sup_{N, d, H \in \mathcal{M}} \hat{J}_{\tilde{\eta}}^N(z_0, H, d) &\leq \sup_N \left\{ V_\star(z, r) + \epsilon \sum_{t=0}^{N} \epsilon^t \right\} \\ &= V_\star(z, 0) + \frac{\epsilon}{1 - \epsilon}. \end{aligned}$$

Thus we conclude statements 4, 5, and 8.

As $V_\star$ is the limit of the value iteration, it is a fixed point of $\mathcal{B}$ (otherwise, the limit would not exist). To see that it is minimal, assume that V' is another fixed point of $\mathcal{B}$ such that $V' \geq V_0$ but that $V'(z_0, r_0) < V_\star(z, r)$ for some z_0 and r_0. Define the function $\hat{V}_0$ by

$$\hat{V}_0(z,r) = \begin{cases} V'(z,r) & \text{if } (z,r) = (z_0, r_0), \\ V_0(z,r) & \text{otherwise.} \end{cases}$$

Then $V_0 \leq \hat{V}_0 \leq V'$, so the value iteration $\hat{V}, \mathcal{B}\hat{V}, \ldots$ converges to some limit $\hat{V}_\star$. But by monotonicity, $V_\star \leq \hat{V}_\star \leq V'$, which is a contradiction. □

2.2 State Feedback

Problem 2—State-feedback minimax adaptive control

Let $Q \in \mathbb{S}^{n_x}$, $R \in \mathbb{S}^{n_u}$, be positive definite. Given a compact set $\mathcal{M} \subset \mathbb{R}^{n_x \times n_x} \times \mathbb{R}^{n_x \times n_u}$, initial state $x_0 \in \mathbb{R}^{n_x}$, and a positive quantity $\gamma > 0$, compute

$$\inf_{\mu} \sup_{w,N,M \in \mathcal{M}} \underbrace{\sum_{t=0}^{N-1} \left(|x_t|_Q^2 + |u_t|_R^2 - \gamma^2 |w_t|^2\right)}_{J_\mu^N(x_0, M, w)} \tag{11}$$

where $M = (A, B) \in \mathcal{M}$, $w_t \in \mathbb{R}^{n_w}$, $N \geq 0$, and the sequences $x_{[0:N]}$, $u_{[0:N-1]}$ are generated by

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad t \geq 0 \tag{12}$$

$$u_t = \mu_t(x_{[0:t]}, u_{[0:t-1]}). \tag{13}$$

The state-feedback minimax adaptive control problem, Problem 2, is similar to a standard $\mathcal{H}_\infty$ control problem, but differs in that the dynamics are uncertain and chosen by the adversary. The problem differs from the principal problem in that the realization of the objective function is known, but the dynamics are not. To relate the two problems, we introduce $z_t = x_t$, $d_t = x_{t+1}$. Substituting $w_t = d_t - Ax_t - Bu_t$ into the dynamics (12), we get

$$z_{t+1} = \hat{A}z_t + \hat{B}u_t + \hat{G}d_t, \tag{14}$$

where $\hat{A} = 0$, $\hat{B} = 0$ and $\hat{G} = I$. For $M = (A, B) \in \mathcal{M}$, let

$$H_M = \begin{bmatrix} Q & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & 0 \end{bmatrix} - \gamma^2 \begin{bmatrix} -A^\mathsf{T} \\ -B^\mathsf{T} \\ I \end{bmatrix} \begin{bmatrix} -A^\mathsf{T} \\ -B^\mathsf{T} \\ I \end{bmatrix}^\mathsf{T}. \tag{15}$$

Then, the objective (11) becomes

$$\inf_{\mu} \sup_{d,N,M \in \mathcal{M}} \sum_{t=0}^{N-1} \sigma_{H_M}(z, u, d). \tag{16}$$

$$H_M/H_M^{dd} = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \succeq 0.$$

Finally, note that H_M fulfills (2) as $R \succ 0$. We summarize the result in the following theorem.

THEOREM 2—STATE-FEEDBACK REDUCTION
The value of Problem 2 is equal to the value of Problem 1 with $\mathcal{M} = \{H_M : M \in \mathcal{M}\}$ with H_M as in (15), $z_t = x_t$, $d_t = x_{t+1}$, and $\hat{A} = 0$, $\hat{B} = 0$, and $\hat{G} = I$. Further, given a policy $\hat{\mu} : (z_{[0:t]}, u_{[0:t-1]}, d_{[0:t-1]}) \mapsto u_t$,

$$\sup_{d,N,H \in \widehat{\mathcal{M}}} \hat{J}_{\hat{\mu}}^N(z_0, H, d) = \sup_{w,N,M \in \mathcal{M}} J_\mu^N(x_0, M, w),$$

where

$$\mu(x_{[0:t]}, u_{[0:t]}) = \hat{\mu}(z_{[0:t]}, u_{[0:t-1]}, d_{[0:t-1]})$$

is feasible for Problem 3.

2.3 Output Feedback

This section presents how to rewrite the output-feedback minimax adaptive control problem formalized below as an instance of the principal problem, Problem 1.

PROBLEM 3—OUTPUT-FEEDBACK MINIMAX ADAPTIVE CONTROL
Let $Q \in \mathbb{S}^{n_x}$, $R \in \mathbb{S}^{n_u}$, be positive definite. Given a compact set $\mathcal{M} \subset \mathbb{R}^{n_x \times n_x} \times \mathbb{R}^{n_x \times n_u} \times \mathbb{R}^{n_x \times n_w} \times \mathbb{R}^{n_y \times n_x} \times \mathbb{R}^{n_y \times n_v}$, and a positive quantity $\gamma > 0$, consider

$$\underbrace{-|x_0 - \hat{x}_0|^2_{S_{M,0}} + \sum_{t=0}^{N-1} \left(|x_t|_Q^2 + |u_t|_R^2 - \gamma^2 |(w_t, v_t)|^2\right)}_{J_\mu^N(\hat{x}_0, M, y, w, v, x_0)}. \tag{17}$$

where $S_{M,0} \succeq 0$, $M = (A, B, G, C, D) \in \mathcal{M}$, $w_t \in \mathbb{R}^{n_w}$, $v_t \in \mathbb{R}^{n_v}$, $N \geq 0$, and the sequences, $x_{[0:N]}$, $y_{[0:N-1]}$ and $u_{[0:N-1]}$ are generated by

$$x_{t+1} = Ax_t + Bu_t + Gw_t, \quad t \geq 0 \tag{18a}$$
$$y_t = Cx_t + Dv_t, \tag{18b}$$
$$u_t = \mu_t(y_{[0:t-1]}, u_{[0:t-1]}). \tag{18c}$$

Compute

$$\inf_\mu \sup_{y, M \in \mathcal{M}, N, w, v} J_\mu^N(\hat{x}_0, M, y, w, v). \tag{19}$$

Remark 2

We assume that all members of $\mathcal{M}$ have the same order, i.e., n_x, n_w and n_v are constant for all $M \in \mathcal{M}$. This is for notational simplicity only and is not necessary for the theoretical development in this section.

We make the following assumptions on the problem parameters.

Assumption 1—Problem parameters

For each $M \in \mathcal{M}$, $\begin{bmatrix} A & G \end{bmatrix}$ and D have full column rank, and $\begin{bmatrix} A^\mathsf{T} & \sqrt{Q} \end{bmatrix}^\mathsf{T}$ has full row rank.

We follow the approaches of [James and Baras, 1995] and [Basar and Bernhard, 2008], and split the optimization problem (19) into three steps,

$$\inf_{\mu} \sup_{y, M \in \mathcal{M}, w, v, N, x_0} J_\mu^N(\hat{x}_0, M, y, w, v, x_0) = \inf_{\mu} \sup_{y, M \in \mathcal{M}, N} \sup_{x} \underbrace{\left[\sup_{w,v,x_0} \{ J_\mu^N(\hat{x}_0, M, y, w, v, x_0) : x_N = x \} \right]}_{W_M^N(x, u_{[0:N-1]}, y_{[0:N-1]})}. \tag{20}$$

The supremum in $W_M^N(x, u_{[0:N-1]}, y_{[0:N-1]})$ is taken subject to the inputs $w_{[0:N-1]}$, $v_{[0:N-1]}$ and $u_{[0:N-1]}$, the observed output $y_{[0:N-1]}$, final state $x_N = x$ and model M being feasible. Feasibility means that the state sequence $x_{[0:N]}$ is generated by the dynamics (18a) and the output sequence $y_{[0:N-1]}$ is generated by the output equation (18b) under the model M.

Note that $W_M^N(x, u, y)$ is a function of the trajectory $u_{[0:N-1]}$ and not the control law μ. This is because the outer optimization steps determine the control law and the output sequence. The control law and the output sequence in turn determine the control signal trajectory.

The reformulation has two major benefits. The first is that minimizing over μ_t and maximizing over $y_{[0:t-1]}$ commutes as μ_t is a function of $y_{[0:t-1]}$. This interchange of extremization leads to a sequential optimization problem that can be solved by backwards dynamic programming. The second benefit is that $W_M^N(x, u_{[0:N-1]}, y_{[0:N-1]})$ can be characterized using standard forward Riccati recursions and observer equations of $\mathcal{H}_\infty$-control theory.

The rest of the section is organized as follows. Section 2.3 shows how to rewrite Problem 3 as an instance of Problem 1 in the case where the dynamics are known. Section 2.3 modifies the approach to the case where $\mathcal{M}$ is a finite set of models.

Known Dynamics In this case, $\mathcal{M} = \{M\}$ for some $M = (A, B, G, C, D)$, and we will drop the subscript M from the notation. The value (19) is then equal to

$$\inf_{\mu} \sup_{N, y_{[0:N-1]}, x} W^N(x, u_{[0:N-1]}, y_{[0:N-1]}).$$

Consider the forward Riccati recursions:

$$\begin{aligned} S_{t+1} &= (AX_t^{-1}A^\mathsf{T} + \gamma^{-2}GG^\mathsf{T})^{-1}, \\ X_t &= S_t + \gamma^2 C^\mathsf{T}(DD^\mathsf{T})^{-1}C - Q, \\ L_t &= \gamma^2 AX_t^{-1}C^\mathsf{T}(DD^\mathsf{T})^{-1}, \end{aligned} \tag{21}$$

where $S_t \in \mathbb{S}^{n_x}$, $X_t \in \mathbb{S}^{n_y}$, $L_t \in \mathbb{R}^{n_x \times n_y}$. The $\mathcal{H}_\infty$-observer states obey the dynamics

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + L_t(y_t - C\hat{x}_t) + AX_t^{-1}Q\hat{x}_t. \tag{22}$$

The initial S_0 and $\hat{x}_0$ are provided by the designer in (17). The following lemma summarizes the recursive computation of W^N, we refer the reader to [Basar and Bernhard, 2008, Chapter 6] for a proof.

Lemma 1
For Problem 3 with $\mathcal{M} = \{(A, B, C, G, D)\}$, let $S_{t+1} \in \mathbb{S}^{n_x}$, $X_t \in \mathbb{S}^{n_y}$, and $L_t \in \mathbb{R}^{n_x \times n_y}$ be defined recursively for $t = 0, 1, \ldots$ by (21). If $S_t \succ Q$ for all $t = 0, \ldots, N$, then W^N defined in (20) satisfies

$$W^N(x, u_{[0:N-1]}, y_{[0:N-1]}) = -|x - \hat{x}_N|^2_{S_N} + r_N. \tag{23}$$

The observer states $\hat{x}_t$ follow the observer equation (22), $r_0 = 0$ and

$$\begin{aligned} r_{t+1} = r_t + |\hat{x}_t|^2_{(Q^{-1} - S_t^{-1})^{-1}} + |u_t|^2_R \\ - |y_t - C(S-Q)^{-1}S\hat{x}_t|^2_{(DD^\mathsf{T}/\gamma^2 + C(S-Q)^{-1}C^\mathsf{T})^{-1}}. \end{aligned} \tag{24}$$

If, for any $t \geq 0$, $S_t \not\succ Q$, then the value (19) is unbounded.

It is not obvious how the designer should choose the initial S_0. The Riccati recursions (21) are known to admit positive definite fixed points S if γ is sufficiently large, and the minimal fixed point leads to stable observer dynamics (22). Thus, we can choose $S_0 = S$, a stabilizing positive definite fixed point of the Riccati recursions, and will assume so for the rest of the section.

Assumption 2
γ is large enough so that there exists a stabilizing positive definite fixed point of the Riccati recursions (21). Denote by S the minimal fixed point and assume $S_0 = S_M$.

Remark 3
Note that under Assumption 2

$$r_t = \sum_{s=0}^{t-1} \sigma_{\hat{Q}}(\hat{x}_s, u_s, y_s)$$

where

$$\begin{aligned} &\hat{Q}^{xx} = SX^{-1}S - S \quad \hat{Q}^{xu} = 0, \\ &\hat{Q}^{uu} = R, \qquad \hat{Q}^{yy} = -(DD^{\mathsf{T}}/\gamma^2 + C(S-Q)^{-1}C^{\mathsf{T}})^{-1}, \\ &\hat{Q}^{uy} = 0, \quad \hat{Q}^{xy} = \gamma^2 SX^{-1}C^{\mathsf{T}}(DD^{\mathsf{T}})^{-1}. \end{aligned} \tag{25}$$

Remark 4
The expression of r_{t+1} is unnecessary when one employs the certainty equivalence principle of [Basar and Bernhard, 2008], but is crucial to our theory of adaptive control.

Theorem 3
For Problem 3 with $\mathcal{M} = \{(A, B, C, G, D)\}$, under assumptions 1 and 2, let $\hat{A} = AX^{-1}S$, $\hat{B} = B$, $\hat{G} = L$, where (S, X, L) is a fixed point of the Riccati recursions (21). Then, the optimal value of Problem 3 is equal to the optimal value of Problem 1 where H is replaced by $\hat{Q}$ in (25) and $z_t = \hat{x}_t$. Further, given a policy $\hat{\mu} : (z_{[0:t]}, u_{[0:t-1]}, y_{[0:t-1]})) \mapsto u_t$,

$$\sup_{d,N} \hat{J}^N_{\hat{\mu}}(z_0, H, d) = \sup_{y,N,w,v,x_0} J^N_\mu(\hat{x}_0, M, y, w, v, x_0),$$

where

$$\mu(y_{[0:t-1]}, u_{[0:t-1]}) = \hat{\mu}(z_{[0:t]}, u_{[0:t-1]}, y_{[0:t-1]})$$

is feasible for Problem 3.

Proof. From Lemma 1, we know that for a fixed policy μ and horizon $N \geq 0$, the value of the inner optimization problem in (19),

$$\sup_{w,v} \left\{ \sum_{t=0}^{N-1} \left(|x_t|^2_Q + |u_t|^2_R - \gamma^2 |(w_t, v_t)|^2 \right) - |x_0 - \hat{x}_0|^2_{S_0} \right\}$$
$$= \underbrace{\sup_{y_{[0:N-1]}, x} \left\{ -|x - \hat{x}_N|_{S_N} + \sum_{t=0}^{N-1} \sigma_{\hat{Q}}(\hat{x}_t, u_t, y_t) \right\}}_{J^N_\mu(\hat{x}_0)},$$

where $\hat{Q}$ is defined in (25) and the sequences $\hat{x}_t$ and u_t are generated by (22) and (18c), respectively. As S_N is positive definite, the unique maximizing argument is $x_\star = \hat{x}_N$. By assumption, S_0 is a fixed point of the Riccati recursion (21) and therefore $S_t = S = S_0$. The corresponding matrices $\hat{Q}_t$, X_t, L_t, $\hat{A}$, $\hat{B}$ and $\hat{G}$ are also stationary. Finally, as $\hat{x}_t$ is a function of $y_{[0:t-1]}$ and $u_{[0:t-1]}$, the sets of feasible controllers μ in Problem 3 and $\hat{\mu}$ in this theorem are equal. We conclude that the optimal values are equal. □

Main Result: Output Feedback Adaptive Control We now consider the case where the dynamics of the system are unknown and the controller must adapt to the system, but assume that the model set $\mathcal{M}$ is finite. Let $\hat{x}_{M,t}$, $\hat{A}_M$, $\hat{B}_M$, $\hat{G}_M$ and $\hat{Q}_M$ be as in Theorem 3 for each $M \in \mathcal{M}$ and define

$$z_t = \begin{bmatrix} \hat{x}_{1,t} \\ \vdots \\ x_{|\mathcal{M}|,t} \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} \hat{B}_1 \\ \vdots \\ \hat{B}_{|\mathcal{M}|} \end{bmatrix}, \quad \hat{G} = \begin{bmatrix} \hat{G}_1 \\ \vdots \\ \hat{G}_{|\mathcal{M}|} \end{bmatrix}, \tag{26}$$

$$\hat{A} = \mathrm{BlockDiag}\{A_M X_M^{-1} S_M : M \in \mathcal{M}\},$$

and $d_t = y_t$. Further, let $H_i \in \mathbb{S}^{|\mathcal{M}|n_x \times n_u \times n_y}$ be given by

$$\begin{array}{c} \\ \\ \\ \\ \text{row } i \\ \\ \\ \\ \\ \\ \end{array}
\overset{\displaystyle \begin{array}{c}\text{column } i\\ \uparrow\end{array}}{
\left[\begin{array}{ccccccc|c|c}
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \ddots & & \vdots & & & \vdots & \vdots & \vdots \\
0 & & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & \cdots & 0 & \hat{Q}_i^{xx} & 0 & \cdots & 0 & \hat{Q}_i^{xu} & \hat{Q}_i^{xy} \\
0 & & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & & & \vdots & & \ddots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
\hline
0 & \cdots & 0 & \hat{Q}_i^{ux} & 0 & \cdots & 0 & \hat{Q}_i^{uu} & \hat{Q}_i^{uy} \\
\hline
0 & \cdots & 0 & \hat{Q}_i^{yx} & 0 & \cdots & 0 & \hat{Q}_i^{yu} & \hat{Q}_i^{yy}
\end{array}\right]}. \tag{27}$$

The following theorem shows that the output-feedback adaptive control problem can be reduced to an instance of Problem 1.

THEOREM 4—REDUCTION
Under Assumptions 1 and 2, the optimal value of Problem 3 is equal to the optimal value of Problem 1 where $\mathcal{M}$ is replaced by $\{H_M : M \in \mathcal{M}\}$ with H_M as in (27) and $\hat{A}$, $\hat{B}$ and $\hat{G}$ in (26), and z_0 in (26). Further, given a policy $\hat{\mu} : (z_{[0:t]}, u_{[0:t-1]}, y_{[0:t-1]}) \mapsto u_t$,

$$\sup_{d,N,H\in\widehat{\mathcal{M}}} \hat{J}_{\hat{\mu}}^N(z_0, H, d) = \sup_{y,M\in\mathcal{M},N,w,v,x_0} J_\mu^N(\hat{x}_0, M, y, w, v, x_0),$$

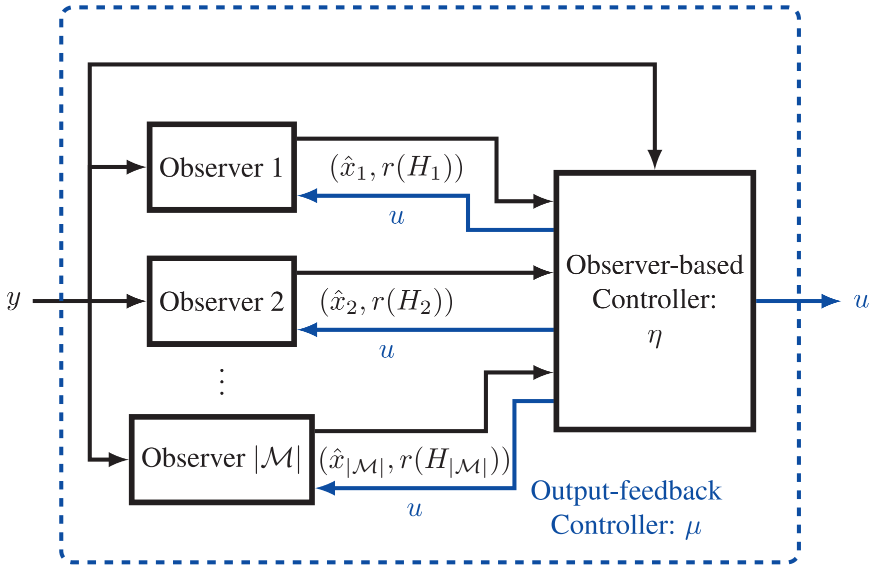


Figure 3. Observer-based adaptive controller resulting from combining Theorem 4 and Theorem 1.

where

$$\mu(y_{[0:t-1]}, u_{[0:t-1]}) = \hat{\mu}(z_{[0:t]}, u_{[0:t-1]}, y_{[0:t-1]})$$

is feasible for Problem 3.

Proof. The proof is identical to that of Theorem 3, but with the additional steps mentioned below. With W as in (20), we have that the cost (19) is equal to

$$\inf_{\mu} \sup_{M,N,y_{[0:N-1]},x} W_M^N(x, u_{[0:N-1]}, y_{[0:N-1]}).$$

Let $\hat{S}_M \in \mathbb{R}^{|\mathcal{M}|n_x \times |\mathcal{M}|n_x}$ be the matrix that is zero except on the M-th diagonal block, which is equal to S_M. By Lemma 1, we have that

$$\begin{aligned}\sup_x W_M^N(x, u_{[0:N-1]}, y_{[0:N-1]}) &= \sup_x \{-|x - \hat{x}_{M,N}|^2_{\hat{S}_M} + r_{M,N}\} \\ &= \sum_{t=0}^{N-1} \sigma_{\hat{Q}_M}(\hat{x}_{M,t}, u_t, y_t) = \sum_{t=0}^{N-1} \sigma_{H_M}(z_t, u_t, y_t).\end{aligned}$$

where $\hat{Q}_M$ is as in (25), z_t is as in (26), and H_M is as in (27). □

Together with Theorem 1, Theorem 4 can be realized by a bank of observers as in Fig. 3.

3. Explicit Controller Synthesis with Performance Bounds

3.1 Bellman Inequalities

Sometimes, it may be difficult to compute the recursion (9) or to find the minimal nonnegative fixed point of the Bellman operator $\mathcal{B}$ in (8b) and the corresponding optimal control law. This is typically the case in our setting. Some exceptions are the case of $\mathcal{M}$ being a singleton or the case where the uncertainty set $\mathcal{M}$ contains an element H_M that dominates all other models. By dominate, we mean that a control policy that is optimal for H_M achieves a lower cost for all other models in $\mathcal{M}$.

This section presents theory related to bounding the value function that relies on periodic compensation. As we will see in Section 4, delays in the control signal significantly complicates the computation of the value-function approximation. Periodic compensation is a powerful method to handle this problem.

Let τ be a positive integer. We model τ-periodic compensation as a control law that contains a supervisor, $\bar{\eta}$, that periodically selects a sequence of τ control components,

$$\bar{\eta} : (z_{\tau k}, r_{\tau k}) \mapsto (\bar{\eta}_{\tau k}, \bar{\eta}_{\tau k+1}, \dots, \bar{\eta}_{\tau k+\tau-1}). \tag{28}$$

During this period, the control signal is computed by the component control laws

$$u_{\tau k+s} = \bar{\eta}_{\tau k+s}(z_{\tau k+s}, r_{\tau k+s}). \tag{29}$$

The connection to dynamic programming lies in compositions of the Bellman operator. The periodic versions of the operators $\mathcal{B}$ and $\mathcal{B}_u$ in (8) are

$$\mathcal{B}^{\tau}\bar{V} \triangleq \underbrace{\mathcal{B}\mathcal{B}\cdots\mathcal{B}}_{\tau\text{-times composition}}\bar{V}, \tag{30a}$$

$$\begin{aligned}\bar{\mathcal{B}}_{\bar{\eta},\tau}\bar{V}(z,r) &\triangleq \mathcal{B}_{\bar{\eta}_1}\mathcal{B}_{\bar{\eta}_2}\cdots\mathcal{B}_{\bar{\eta}_\tau}\bar{V}(z,r),\\ \bar{\eta}(z,r) &= (\bar{\eta}_1, \bar{\eta}_2, \dots, \bar{\eta}_\tau).\end{aligned} \tag{30b}$$

THEOREM 5

For Problem 1, let V_0 be as in (9), and let $\mathcal{B}^{\tau}$ and $\bar{\mathcal{B}}_{\bar{\eta},\tau}$ be as in (30). If there exists a function $\bar{V} : \mathbb{R}^{n_z} \times \mathbb{R}^{\mathcal{M}} \to \mathbb{R}$ and positive integer τ such that

$$\bar{V} \geq V_0, \quad \mathcal{B}^{\tau}\bar{V} \leq \bar{V},$$

then the value iteration $V_0, V_1, \dots$ converges to a limit $V_\star \leq \bar{V}$.

If there exists a τ-periodic control law $\bar{\eta}$ as in (28) and (29) such that

$$\bar{V} \geq V_0, \quad \bar{\mathcal{B}}_{\bar{\eta},\tau}\bar{V} \leq \bar{V},$$

then $\mathcal{B}^\tau \bar{V} \leq \bar{V}$ *and the policy* $\bar{\eta}$ *achieves a cost,* (3)*, no greater than*

$$\max\{V(z_0,0), \mathcal{B}_{\bar{\eta}_\tau} \bar{V}(z_0,0), \ldots, \mathcal{B}_{\bar{\eta}_2} \cdots \mathcal{B}_{\bar{\eta}_\tau} \bar{V}(z_0,0)\}, \tag{31}$$

where $(\bar{\eta}_1, \bar{\eta}_2, \ldots, \bar{\eta}_\tau) = \bar{\eta}(z_0, 0)$*.*

Corollary 1
If $\bar{V}$ *is the smallest fixed point of* $\mathcal{B}^\tau$ *greater than* V_0*, then* $\bar{V} = V_\star$*.*

Corollary 2
If one has found a $\bar{V} \geq V_0$ *that satisfies the periodic Bellman inequality,* $\mathcal{B}^\tau \bar{V} \leq \bar{V}$*, then the control law*

$$\bar{\eta}_{\tau k+s} = \arg\min_u \mathcal{B}_u \mathcal{B}^{\tau-1-s} \bar{V}(z_{\tau k+s}, r_{\tau k+s})$$

achieves a cost, (3)*, no greater than*

$$\max_{s=0,1,\ldots,\tau-1} \{\mathcal{B}^s \bar{V}(z_0,0)\}.$$

Corollary 3—1-Periodic
If there exists a function $\bar{V} : \mathbb{R}^{n_z} \times \mathbb{R}^{\mathcal{M}} \to \mathbb{R}$ *and control law* $\bar{\eta} : \mathbb{R}^{n_z} \times \mathbb{R}^{\mathcal{M}} \to \mathbb{R}^{n_u}$ *such that* $\mathcal{B}_{\bar{\eta}} \bar{V} \leq \bar{V}$*, then the value iteration* $V_0, V_1, \ldots$ *is bounded, and the control policy* $u_t = \bar{\eta}(z_t, r_t)$ *achieves a cost no greater than* $\bar{V}(z_0, 0)$ *for Problem 1.*

Proof. Let $\bar{V} \geq V_0$ be satisfy $\mathcal{B}^\tau \bar{V} \leq \bar{V}$ for some τ. As $\mathcal{B}$ is monotone, so is $\mathcal{B}^\tau$. Then, for any $k = 0, 1, \ldots$, by monotonicity, $\bar{V} \geq \mathcal{B}^{k\tau} \bar{V} \geq \mathcal{B}^{k\tau} V_0 = V_{k\tau}$. By Theorem 1, the value iteration is monotone, so for any $l = 0, 1, \ldots, \tau - 1$, $\bar{V} \geq V_{k(\tau+1)} \geq V_{k\tau+l}$. Thus, the value iteration is bounded by $\bar{V}$.

For the second part, assume that there exists a $\bar{V}$ and $\bar{\eta}$ such that $\mathcal{B}_{\bar{\eta},\tau} \bar{V} \leq \bar{V}$. As $\mathcal{B}_{\bar{\eta},\tau} \bar{V} \geq \mathcal{B}^\tau \bar{V}$, we have that $\mathcal{B}^\tau \bar{V} \leq \bar{V}$. It remains to show that the controller $\bar{\eta}$ achieves a cost no greater than (31).

Consider the value iteration starting with $\bar{V}_0 = \bar{V}$ and

$$\bar{V}_{(k+1)\tau} = \mathcal{B}_{\bar{\eta},\tau} \bar{V}_k,$$
$$\bar{V}_{k\tau+s+1} = \mathcal{B}_{\bar{\eta}_s} \bar{V}_{k\tau+s}, \quad s = 0, 1, \ldots, \tau - 2.$$

Fix some $N = 0, 1, \ldots$ and consider the factorization $N = k\tau + s$ where $s = 0, 1, \ldots, \tau - 1$. We have that

$$\begin{aligned}\sup_{H \in \mathcal{M}, d} J_{\bar{\eta}}^N(z_0, H, d) &= \bar{\mathcal{B}}_{\eta_1} \bar{\mathcal{B}}_{\eta_2} \cdots \bar{\mathcal{B}}_{\eta_s} \bar{\mathcal{B}}_{\bar{\eta},\tau}^k V_0(z_0, 0) \\ &\leq \bar{\mathcal{B}}_{\eta_1} \bar{\mathcal{B}}_{\eta_2} \cdots \bar{\mathcal{B}}_{\eta_s} \bar{\mathcal{B}}_{\bar{\eta},\tau}^k \bar{V}(z_0, 0) \\ &\leq \bar{\mathcal{B}}_{\eta_1} \bar{\mathcal{B}}_{\eta_2} \cdots \bar{\mathcal{B}}_{\eta_s} \bar{V}(z_0, 0).\end{aligned}$$

The bound (31) follows from taking the supremum over $N = 0, 1, \ldots$ on both sides. □

3.2 Solution to the Bellman Inequality

This section is devoted to an explicit solution to the periodic Bellman inequality, $\mathcal{B}_{\bar{\eta},\tau}\bar{V} \leq \bar{V}$ in Theorem 5, in the case of a finite model set $\mathcal{M} = \{H_1, \dots, H_N\}$. We parameterize an upper bound of the value function in a set of positive definite matrices $P_{ij} \in \mathbb{S}^{n_z}$,

$$\bar{V}(z,r) = \max_{i,j} \bar{V}_{ij}(z,r), \tag{32a}$$

$$\bar{V}_{ij}(z,r) = |z|^2_{P_{ij}} + \left(r(H_i) + r(H_j)\right)/2, \tag{32b}$$

where $i, j = 1, \dots, N$. We restrict out attention to τ-periodic certainty-equivalence controllers of the form[2]

$$\begin{aligned} k(n\tau) &= \underset{i\in\{1,\dots,N\}}{\arg\max}\ r_{n\tau}(H_i), \\ u_{n\tau+s} &= -K_{k(n\tau)} z_{n\tau+s}, \quad s = 0, \dots, \tau-1, \end{aligned} \tag{33}$$

for some matrices $K_1, \dots, K_N \in \mathbb{R}^{n_u \times n_z}$. In the language of Section 3.1, the supervisor, $\bar{\eta}$, is executed at each τ-th time step and generates the feedback control law to be used over the next τ time steps:

$$\bar{\eta}(z_{n\tau}, r_{n\tau}) = (\bar{\eta}_{k(n\tau)}, \bar{\eta}_{k(n\tau)}, \dots, \bar{\eta}_{k(n\tau)}),$$

where $\bar{\eta}_k(z,r) = -K_k z$ is the component control law.

REMARK 5
The theoretical development in this section does not rely on the gain matrices K_k being constant over each period. One could let the supervisor, $\bar{\eta}$, predetermine a sequence of gain matrices for the next period.

The Bellman operator acting on a function $\bar{V}_{ij}$ is the supremum of the quadratic form of the operator $\mathcal{G}$ acting on the state and the disturbance, d:

$$\mathcal{B}_{-K_k z}\bar{V}_{ij}(z,r) = \sup_d \Big\{ \sigma_{\mathcal{G}(P_{ij},K_k,(H_i+H_j)/2)}(z,d) + (r(H_i) + r(H_j))/2 \Big\}, \tag{34}$$

where

$$\mathcal{G}(P,K,H) \triangleq \begin{bmatrix} A - BK & G \end{bmatrix}^\mathsf{T} P \begin{bmatrix} A - BK & G \end{bmatrix} + \begin{bmatrix} I & 0 \\ -K & 0 \\ 0 & I \end{bmatrix}^\mathsf{T} H \begin{bmatrix} I & 0 \\ -K & 0 \\ 0 & I \end{bmatrix}. \tag{35}$$

[2] If the max is achieved on a set, any selection mechanism will work.

We parameterize a bound of the temporal evolution over a period in a sequence of matrices $P_{ij,k}^1, \dots, P_{ij,k}^\tau \in \mathbb{S}^{n_z}$, so that $\mathcal{B}_{\bar{\eta}_k}^s V_{ij}(z, r) \leq |z|^2_{P_{ij,k}^s} + (r(H_i) + r(H_j))/2$ for $s = 1, \dots, \tau$. This requirement is equivalent to the set of matrix inequalities

$$\begin{aligned} \begin{bmatrix} P_{ij,k}^1 & 0 \\ 0 & 0 \end{bmatrix} &\succeq \mathcal{G}(P_{ij}, K_k, (H_i + H_j)/2), \\ \begin{bmatrix} P_{ij,k}^{s+1} & 0 \\ 0 & 0 \end{bmatrix} &\succeq \mathcal{G}(P_{ij,k}^s, K_k, (H_i + H_j)/2). \end{aligned} \tag{36}$$

By the choice of k in (33), we have that $r(H_i) \leq r(H_k)$ for all i. The following theorem formalizes the sufficient condition that if $P_{ij,k}^\tau \preceq P_{jk}$, then $\bar{V}$ and $\bar{\eta}$ fulfills the τ-periodic Bellman inequality.

Theorem 6—Explicit solution

For Problem 1 where the model set is finite, $\mathcal{M} = \{H_1, \dots, H_N\}$. Assume there exist

- *a positive integer τ,*
- *matrices $P_{ij,k}^s = P_{ji,k}^s \in \mathbb{S}^{n_z}$ for $i, j, k = 1, \dots, N$ and $s = 1, \dots, \tau$,*
- *positive semidefinite matrices $P_{ij} = P_{ji} \in \mathbb{S}^{n_z}$ for $i, j = 1, \dots, N$,*
- *gain matrices $K_1, \dots, K_N \in \mathbb{R}^{n_u \times n_z}$.*

If $P_{ij,k}^\tau \preceq P_{jk}$ and (36) are fulfilled for all $s = 1, \dots, \tau - 1$ and $i, j, k = 1, \dots, N$ except for $i \neq j = k$, then the value approximation $\bar{V}$ in (32) with the certainty-equivalence control law $\bar{\eta}$ in (33) fulfills the periodic Bellman inequality, $\mathcal{B}_{\bar{\eta},\tau} \bar{V} \leq \bar{V}$.

Remark 6

Theorem 3 in [Rantzer, 2021] is obtained as a corollary of Theorem 6 by substituting $\hat{A} = 0$, $\hat{B} = 0$ and $\hat{G} = I$ and H_M from (15) into $\mathcal{G}$, taking $\tau = 1$ and replacing (36) with their lower Schur complements.

Remark 7

The inequalities (36) are affine in P_{ij} and $P_{ij,k}^s$, but are not convex in K_i. One heuristic approach to solve the inequalities is to first solve the linear-quadratic problem associated with each model i to obtain K_i,

$$\inf_{\mu} \sup_{d,N} \left\{ \sum_{t=0}^{N-1} \sigma_{H_i}(z_t, u_t, d_t) \right\}.$$

Then use standard optimization software for semidefinite programming to search for P_{ij} and $P_{ij,k}^s$, holding K_i fixed. This approach was suggested in [Rantzer, 2021] and is also used in the examples in Section 4.

4. Examples

4.1 State-Feedback, Delays and Periodic Compensation

This section studies state-feedback control of the delayed discrete-time integrator where the sign of the gain is unknown. The dynamics can be modeled in two ways, either the sign uncertainty is incorporated into the state matrix or the input:

$$x_{t+1} = \underbrace{\begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}}_{A_i} x_t + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{B} u_t + w_t, \quad i = \pm 1 \tag{37}$$

$$x_{t+1} = \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}}_{A} x_t + \underbrace{\begin{bmatrix} 0 \\ i \end{bmatrix}}_{B_i} u_t + w_t, \quad i = \pm 1, \tag{38}$$

with $Q = I$ and $R = I$. Although the input to output (x_1) behavior of the systems (37) and (38) are identical, from a control perspective, they are significantly different. The difference lies in that an impulse in the controlled input at time t will reveal information about the sign of B_i in (38) at time $t+1$ but information about A_i in (37) not until $t+2$. This difference is reflected in the smallest period, τ, for which the periodic Bellman inequality can be satisfied.

We computed K_1 and K_2 according to Remark 7 and solved the conditions in Theorem 6 using `MOSEK`. We find that for (38), the conditions are satisfied for $\tau = 1$ and $\gamma = 6$. Our software implementation cannot find a solution for $\tau = 1$ for the system in (37) and it is not until $\tau = 2$ that a solution, with $\gamma = 11.2$, is found.

Proposition 1

Given A_i and B in (37), $\gamma > 0$, $Q \in \mathbb{R}^{2\times 2}$ be positive definite and $R > 0$. Then there does not exist matrices $K_i \in \mathbb{R}^{1\times 2}$ and positive definite $P \in \mathbb{R}^{2\times 2}$ such that $A_i - BK_i$ are Schur stable and,

$$\begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} \geq \mathcal{G}(P, K_k, (H_1 + H_2)/2) \tag{39}$$

for both $(i, j, k) = (1, -1, 1)$, and $(-1, 1, -1)$.

Proof. Taking the Schur complement of (39) we get the equivalent condition that for all x

$$|x|_P^2 \geq |x|_Q^2 + |K_k x|_R^2 - \gamma^2 |(A_i - B_i K_k - A_j + B_j K_k)x/2|^2 + |(A_i - B_i K_k + A_j - B_j K_k)x/2|^2_{(P^{-1} - \gamma^{-2} I)^{-1}}.$$

Note that $(P^{-1} - \gamma^{-2}I)^{-1} \succ P$. Let $K_l = \begin{bmatrix} k_l^1 & k_l^2 \end{bmatrix}$. By Lemma 2 in the Appendix, that $lk_l^1 > 0$. Furthermore, we have for $i \neq j$

$$A_i - B_i K_l + A_j - B_j K_l = \begin{bmatrix} 1 & 0 \\ -k_l^1 & -k_l^2 \end{bmatrix}.$$

Thus,

$$\begin{aligned}(A_i - B_i K_l + A_j - B_j K_l)^\mathsf{T} P (A_i - B_i K_l + A_j - B_j K_l) \\ = \begin{bmatrix} p_{11} - 2k_l^1 p_{21} + (k_l^1)^2 p_{22} & * \\ * & * \end{bmatrix},\end{aligned}$$

where $p_{12} = p_{21}$ and

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}.$$

We also note that

$$-\gamma^2 |(A_i - B_i K_l - A_j + B_j K_l)x/2|^2 = -\gamma^2 x_2^2.$$

For $x = \begin{bmatrix} x_1 & 0 \end{bmatrix}^\mathsf{T}$, we get

$$\begin{aligned}&|x|_P^2 - |x|_Q^2 + |K_k x|_R^2 - \gamma^2 |(A_i - B_i K_l - A_j + B_j K_l)x/2|^2 \\ &\quad + |(A_i - B_i K_l + A_j - B_j K_l)x/2|^2_{(P^{-1} - \gamma^{-2}I)^{-1}} \\ &\leq |x|_P^2 - |x|_Q^2 + |K_k x|_R^2 - \gamma^2 |(A_i - B_i K_l - A_j + B_j K_l)x/2|^2 \\ &\quad + |(A_i - B_i K_l + A_j - B_j K_l)x/2|_P^2 \\ &= 2k_l^1 p_{21} - (k_l^1)^2 p_{22} - q_{11} - (k_l^1)^2 R\end{aligned}$$

As $(k_l^1)^2 p_{22} + q_{11} + (k_l^1)^2 R > 0$ and as k_1^1 and k_{-1}^1 have opposite signs, we have that the last line is smaller than zero for $l = 1$, $l = -1$ or both. □

4.2 Approximate Unstable Pole Cancellation

We conclude the examples by synthesizing a controller for the double integrator with uncertain approximate pole cancellation. See the pole-zero map in Fig. 5. This corresponds to an approximate cancellation of the unstable pole. The step responses of the system in Fig. 4, and Bode plots in Fig. 6 indicate that the high-frequency behavior of G_{mp} and G_{nmp} are similar to an integrator, but that the low-frequency asymptotes are different.

The minimum phase system, G_{mp}, and the nonminimum-phase system, G_{nmp}, have state-space realizations $(A, B, C_{\text{mp}}, D, G)$ and $(A, B, C_{\text{nmp}}, D, G)$

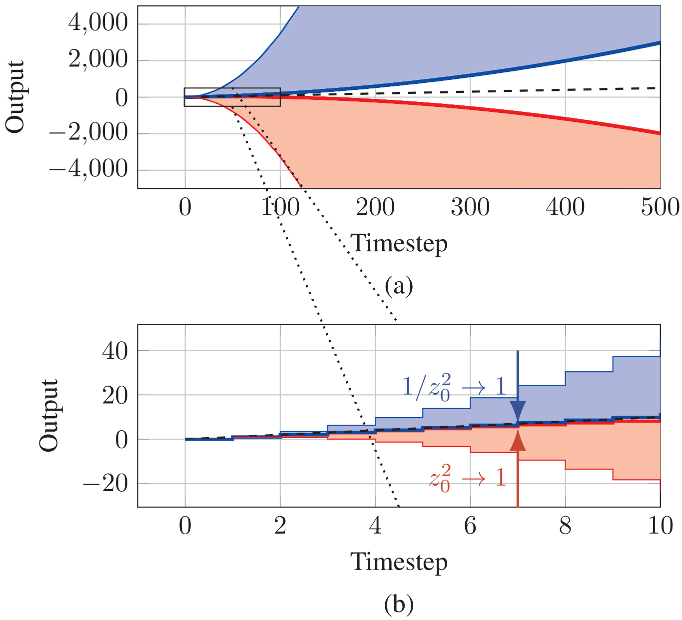


Figure 4. Step responses for the systems G_{mp} (blue) and G_{nmp} (red) as z_0 varies continuously from 1.4 to 1.01 (thick line). The dashed line $y = t$ corresponds to an integrator. As $z_0 \to 1$, note that the short-term behavior of both realizations resemble that of the integrator. The asymptotic behavior, however, do not.

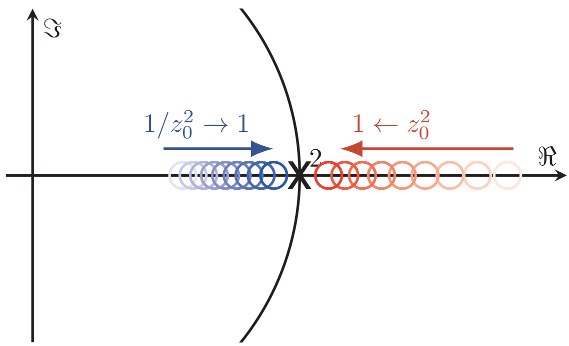


Figure 5. Pole-zero map of the systems G_{mp} (blue) and G_{nmp} (red) as z_0 approaches 1.0. The double pole at $z = 1$ is unstable, and the zero at $z = 1/z_0^2$ (blue) is minimum phase. The zero at $z = z_0^2$ (red) is nonminimum phase. The zeros are reflections in the unit circle.

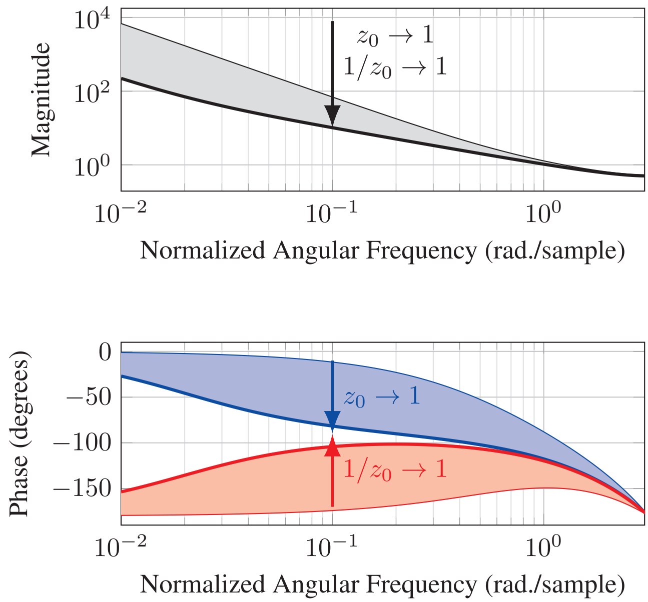


Figure 6. Bode plot of the systems G_{mp} (blue) and G_{nmp} (red) as z_0 varies continuously from 1.4 to 1.01 (thick). The systems' magnitude responses are equal for a fixed z_0, but they differ in phase. This difference is negligable for high frequncies.

respectively, where

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$C_{\text{mp}} = \begin{bmatrix} -1/z_0 + z_0 \\ z_0 \end{bmatrix}^\mathsf{T}, \qquad C_{\text{nmp}} = \begin{bmatrix} -z_0 + 1/z_0 \\ 1/z_0 \end{bmatrix}^\mathsf{T}.$$

Here $z_0 = 1.01$. The LMIs (36) have solutions for $\tau = 4$ and $\gamma = 20$ with $G = I/100$, $Q = I/100$, $R = I/100$ and $D = 1/10$. The matrices were scaled down for numerical stability in the optimization. We evaluate the performance of the periodic certainty-equivalence controller and compare to self-tuning LQG controller described in [Åström and Wittenmark, 2008, Chapter 4] by simulating the nonminimum-phase system with w_t and v_t normally distributed with zero mean and unit variance.

Time series of the output signal are shown in Fig. 7. The self-tuning controller has a spike at time-step 355 which is due to its inability to act against the growth of the mode associated with the second integrator before it starts dominating the output. The minimax adaptive controller does not lead to such spikes. We also show the evolution of γ_t in Fig. 8 and the evolution of the value function in Fig. 9.

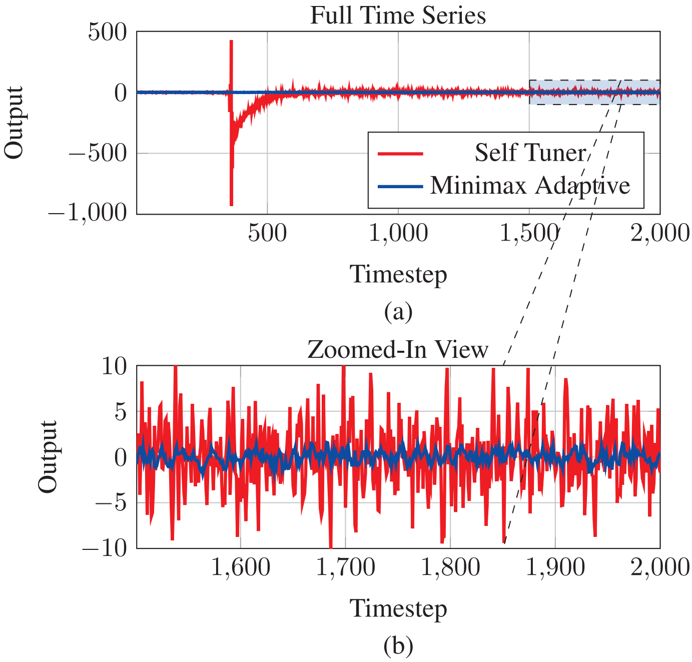


Figure 7. Simulations of the double integrator with uncertain approximate pole cancellation. The system is in feedback with the periodic certainty-equivalence controller of Section 3 (Minimax Adaptive, blue) and the self-tuning LQG controller described in [Åström and Wittenmark, 2008, Chapter 4] (Self tuner, red). The output is shown the entire duration in (a), and a shorter snapshot in (b). Note the spike for the self-tuning regulator at time-step 355. These spikes do not occur with the periodic controller.

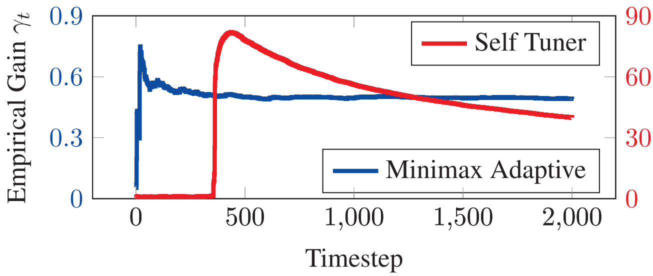


Figure 8. Evolution of γ_t for the double integrator with uncertain approximate pole cancellation. The Minimax Adaptive controller (blue, left scale) initially has a higher γ_t than the self-tuning LQG controller (red, right scale), but due to the build up in the second-order mode, the self tuner spikes at $t = 355$.

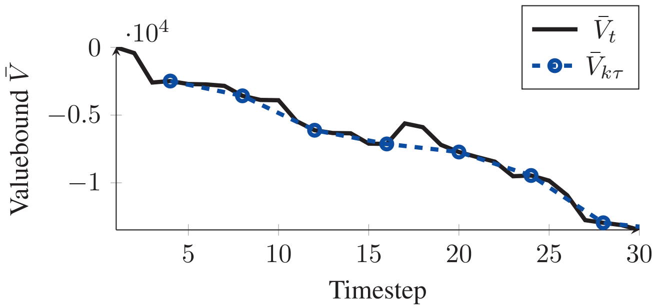


Figure 9. Evolution of the value function approximation for the double integrator with uncertain approximate pole cancellation. The value function is shown for the Minimax Adaptive controller and the periodic decrease, with period $\tau = 4$, is highlighted with circles.

5. Conclusions

We conclude with a few words about the limitations of this work and promising research directions.

1. Including pathological hypotheses in the model set, such as G_{nmp} in (1b) severely impacts the performance guarantee, even when the actual realization of the system is well-behaved. Integrating the framework of Goel *et al.* [Goel and Hassibi, 2023] with the results of this article seems promising to address the pathological hypotheses in the model set. As the authors reformulate regret optimization and competitive ratio optimization as $\mathcal{H}_\infty$ synthesis problems, our results could be integrated to compute suboptimal control policies in the case of parametric uncertainty.

2. We assumed that the model set was finite. Even though finite sets can approximate compact sets of models, our results do not inform how to choose the approximate models and how to quantify the approximation error. Theorem 4 shows that by constructing one $\mathcal{H}_\infty$-observer for each model in the model set, the output feedback problem can be exactly reduced to an instance of Problem 1. If the model set is infinite but compact, one could instead construct a robust $\mathcal{H}_\infty$-observer for each element in a finite cover and approximate the output feedback problem with an instance of Problem 1.

3. Section 4.1 demonstrates that when delays are present, the value function approximation (32) does not capture the probing effect of the control policy. The introduction of periodicity in the control policy, as in Section 3, mitigates this problem as it allows for information gathering

over a longer time horizon. Capturing the probing effect of the control policy in the value approximation is crucial for obtaining tighter performance bounds. Numerical studies of the value iteration could provide insight into this. One could also investigate using reinforcement learning to approximate the value function.

Acknowledgment

The authors would like to thank their colleague Venkatraman Renganathan for valuable feedback on the manuscript. ChatGPT [OpenAI, 2024] was used for light editing of this text. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under grant agreement No 834142 (ScalableControl).

LEMMA 2
The system

$$x_{t+1} = \begin{bmatrix} 1 & i \\ 0 & 1 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_t, \tag{40}$$

where $i \in \pm 1$ and

$$u[t] = K_i x[t]$$

is asymptotically stable if, and only if

$$ik_1^i > 0, \tag{41a}$$

and

$$ik_1^i - 1 < k_2^i < 1 + \frac{ik_1^i}{2}. \tag{41b}$$

Proof. The characteristic polynomial of the closed-loop system is

$$(z-1)(z+k_2^i) + ik_1^i = z^2 + (k_2^i - 1)z + ik_1^i - k_2^i.$$

The inequalities (41) follow by the Jury stability criterion. □

References

Agarwal, N., B. Bullins, E. Hazan, S. Kakade, and K. Singh (2019). "Online control with adversarial disturbances". In: Chaudhuri, K. et al. (Eds.). *Proc. 36th Int. Conf. Mach. Learn.* Vol. 97. PMLR, pp. 111–119. URL: https://proceedings.mlr.press/v97/agarwal19c.html.

Anderson, B. D. O. (2005). "Failures of Adaptive Control Theory and their Resolution". *Commun. Inf. Syst.* **5**:1, pp. 1–20. DOI: 10.4310/CIS.2005.v5.n1.a1.

Åström, K. and B. Wittenmark (2008). *Adaptive Control (2 ed.)* English. Dover, Mineola, NY, USA. ISBN: 0-201-55866-1.

Basar, T. and P. Bernhard (2008). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach*. 2nd ed. Birkhauser, Boston, MA, USA. DOI: 10.1007/978-0-8176-4757-5.

Battistelli, G., E. Mosca, M. G. Safonov, and P. Tesi (2010). "Stability of unfalsified adaptive switching control in noisy environments". *IEEE Trans. Autom. Control* **55**:10, pp. 2424–2429. DOI: 10.1109/TAC.2010.2056473.

Bencherki, F. and A. Rantzer (2023). "Robust simultaneous stabilization via minimax adaptive control". In: *Proc. 62nd IEEE Conf. Decis. Control*, pp. 2503–2508. DOI: 10.1109/CDC49753.2023.10384134.

Bertsekas, D. and I. Rhodes (1973). "Sufficiently informative functions and the minimax feedback control of uncertain dynamic systems". *IEEE Trans. Autom. Control* **18**:2, pp. 117–124. DOI: 10.1109/TAC.1973.1100241.

Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah (2017). "Julia: a fresh approach to numerical computing". *SIAM Rev.* **59**:1, pp. 65–98. DOI: 10.1137/141000671.

Buchstaller, D. and M. French (2016). "Robust stability for multiple model adaptive control: part i—the framework". *IEEE Trans. Autom. Control* **61**:3, pp. 677–692. DOI: 10.1109/TAC.2015.2492518.

Cederberg, D., A. Hansson, and A. Rantzer (2022). "Synthesis of minimax adaptive controller for a finite set of linear systems". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 1380–1384. DOI: 10.1109/CDC51059.2022.9993060.

Didinsky, G. and T. Basar (1994). "Minimax adaptive control of uncertain plants". In: *Proc. 33rd IEEE Conf. Decis. Control*. Vol. 3, pp. 2839–2844. DOI: 10.1109/CDC.1994.411368.

Feldbâum, A. (1963). "Dual control theory problems". In: *Proc. 2nd (IFAC) Congr. Automat. Remote Control Theory*. Vol. 1. 2, pp. 541–550. DOI: 10.1016/S1474-6670(17)69687-3.

Filatov, N. and H. Unbehauen (2000). "Survey of adaptive dual control methods". *IEE Proc. Control Theory Appl.* **147**, pp. 118–128. DOI: 10.1049/ip-cta:20000107.

Ghai, U., X. Chen, E. Hazan, and A. Megretski (2022). "Robust online control with model misspecification". In: Firoozi, R. et al. (Eds.). *Proc. 4th Annu. Learning Dyn. Control Conf.* Vol. 168. PMLR, pp. 1163–1175. URL: https://proceedings.mlr.press/v168/ghai22a.html.

Goel, G. and B. Hassibi (2023). "Competitive control". *IEEE Trans. Autom. Control* **68**:9, pp. 5162–5173. DOI: 10.1109/TAC.2022.3218769.

Hazan, E. (2016). "Introduction to online convex optimization". *Found. Trends Optim.* **2**:3–4, pp. 157–325. ISSN: 2167-3888. DOI: 10.1561/2400000013.

Hespanha, J. P. (2001). "Tutorial on supervisory control". In: *40th IEEE Conf. Decis. Control.* Lecture notes for the workshop *Control using Logic and Switching.* Orlando, FL, USA.

James, M. and J. Baras (1995). "Robust $\mathcal{H}_\infty$ output feedback control for nonlinear systems". *IEEE Trans. Autom. Control* **40**:6, pp. 1007–1017. DOI: 10.1109/9.388678.

Karapetyan, A., A. Iannelli, and J. Lygeros (2022). "On the regret of $\mathcal{H}_\infty$ control". In: *Proc. 61st IEEE Conf. Decis. Control*, pp. 6181–6186. DOI: 10.1109/CDC51059.2022.9992705.

Kjellqvist, O. (2024). "Minimax dual control with finite-dimensional information state". In: Abate, A. et al. (Eds.). *Proc. 6th Annu. Learning Dyn. Control Conf.* Vol. 242. PMLR, pp. 299–311. URL: https://proceedings.mlr.press/v242/kjellqvist24a.html.

Kjellqvist, O. and A. Rantzer (2022). "Learning-enabled robust control with noisy measurements". In: Firoozi, R. et al. (Eds.). *Proc. 4th Annu. Learning Dyn. Control Conf.* Vol. 168. PMLR, pp. 86–96. URL: https://proceedings.mlr.press/v168/kjellqvist22a.html.

Lee, B., A. Rantzer, and N. Matni (2024). "Nonasymptotic regret analysis of adaptive linear quadratic control with model misspecification". In: *Proc. 6th Annu. Learning Dyn. Control Conf.* Vol. 242, pp. 980–992. URL: https://proceedings.mlr.press/v242/lee24a.html.

Liberzon, D. (2003). *Switching in Systems and Control.* Birkhäuser, Boston, MA, USA. ISBN: 978-1-4612-0017. DOI: 10.1007/978-1-4612-0017-8.

Ljung, L. (1977). "Analysis of recursive stochastic algorithms". *IEEE Trans. Autom. Control* **22**:4, pp. 551–575. DOI: 10.1109/TAC.1977.1101561.

Megretski, A. and A. Rantzer (2003). *Bounds on the optimal ℓ_2-gain in adaptive control of a first order linear system.* Tech. rep. Mittag-Leffler, Stockholm, Sweden. ISRN: IML-R-41-02/03-SE+spring.

Mesbah, A. (2018). "Stochastic model predictive control with active uncertainty learning: a survey on dual control". *Annu. Rev. Control* **45**, pp. 107–117. ISSN: 1367-5788. DOI: 10.1016/j.arcontrol.2017.11.001.

OpenAI (2024). *Chatgpt: a large language model by openai.* URL: https://www.openai.com/chatgpt.

Pan, Z. and T. Basar (1998). "Adaptive controller design for tracking and disturbance attenuation in parametric strict-feedback nonlinear systems". *IEEE Trans. Autom. Control* **43**:8, pp. 1066–1083. DOI: 10.1109/9.704978.

Patil, S. V., Y.-C. Sung, and M. G. Safonov (2022). "Unfalsified adaptive control for nonlinear time-varying plants". *IEEE Trans. Autom. Control* **67**:8, pp. 3892–3904. DOI: 10.1109/TAC.2021.3110434.

Rantzer, A. (2020). "Minimax adaptive control for state matrix with unknown sign". In: *Proc. 21st IFAC World Congr.* Vol. 53. 2, pp. 58–62. DOI: 10.1016/j.ifacol.2020.12.049.

Rantzer, A. (2021). "Minimax adaptive control for a finite set of linear systems". In: *Proc. 3th Annu. Learning Dyn. Control Conf.* Vol. 144. Corrected version: https://arxiv.org/abs/2011.10814, pp. 893–904. URL: https://proceedings.mlr.press/v144/rantzer21a.html.

Renganathan, V., A. Iannelli, and A. Rantzer (2023). "An online learning analysis of minimax adaptive control". In: *Proc. 62nd IEEE Conf. Decis. Control*, pp. 1034–1039. DOI: 10.1109/CDC49753.2023.10384114.

Safonov, M. and T.-C. Tsao (1997). "The unfalsified control concept and learning". *IEEE Trans. Autom. Control* **42**:6, pp. 843–847. DOI: 10.1109/9.587340.

Simchowitz, M. (2020). "Making non-stochastic control (almost) as easy as stochastic". In: Larochelle, H. et al. (Eds.). *Adv. Neural Inf. Process. Syst.* Vol. 33. Curran Associates, Inc., pp. 18318–18329. URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/d4ca950da1d6fd954520c45ab19fef1c-Paper.pdf.

Striebel, C. (1965). "Sufficient statistics in the optimum control of stochastic systems". *J. Math. Anal. Appl.* **12**:3, pp. 576–592. ISSN: 0022-247X. DOI: 10.1016/0022-247X(65)90027-2.

Sutton, R. S. and A. G. Barto (2018). *Reinforcement Learning: An Introduction.* Second. The MIT Press, Cambridge, MA, USA. URL: http://incompleteideas.net/book/the-book-2nd.html.

Tsiamis, A., I. Ziemann, N. Matni, and G. J. Pappas (2023). "Statistical learning theory for control: a finite-sample perspective". *IEEE Control Syst. Mag.* **43**:6, pp. 67–97. DOI: 10.1109/MCS.2023.3310345.

Vinnicombe, G. (2004). "Examples and counterexamples in finite ℓ_2-gain adaptive control". In: *16th Int. Symp. Mathematical Theory Networks Systems*. Leuven, Belgium.

Witsenhausen, H. (1966). *Minimax Control of Uncertain Systems*. Massachusetts Institute of Technology. URL: https://books.google.se/books?id=PCxMnQEACAAJ.

Wittenmark, B. (1995). "Adaptive dual control methods: an overview". In: *Proc 5th IFAC Symp. Adaptive Syst. Control Signal Process.* Vol. 28. 13, pp. 67–72. DOI: 10.1016/S1474-6670(17)45327-4.