LUND UNIVERSITY

On Control and Estimation of Large and Uncertain Systems

Kjellqvist, Olle

2022

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

Link to publication

LUND UNIVERSITY

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

On Control and Estimation of Large and Uncertain Systems

Olle Kjellqvist

Lund University

Department of Automatic Control

Licentiate Thesis TFRT-3277
ISSN 0280–5316

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

Printed in Sweden by Media-Tryck.
Lund 2022

Abstract

This thesis contains an introduction and six papers about the control and estimation of large and uncertain systems.

The first paper poses and solves a deterministic version of the multiple-model estimation problem for finite sets of linear systems. The estimate is an interpolation of Kalman filter estimates. It achieves a provided energy gain bound from disturbances to the point-wise estimation error, given that the gain bound is feasible. The second paper shows how to compute upper and lower bounds for the smallest feasible gain bound. The bounds are computed via Riccati recursions. The third paper proves that it is sufficient to consider observer-based feedback in output-feedback control of linear systems with uncertain parameters, where the uncertain parameters belong to a finite set. The paper also contains an example of a discrete-time integrator with unknown gain.

The fourth paper argues that the current methods for analyzing the robustness of large systems with structured uncertainty do not distinguish between sparse and dense perturbations and proposes a new robustness measure that captures sparsity. The paper also thoroughly analyzes this new measure. In particular, it proposes an upper bound that is amenable to distributed computation and valuable for control design. The fifth paper solves the problem of localized state-feedback $\mathcal{H}_2$ control with communication delay for large discrete-time systems. The synthesis procedure can be performed for each node in parallel. The paper combines the localized state-feedback controller with a localized Kalman filter to synthesize a localized output feedback controller that stabilizes the closed-loop subject to communication constraints.

The sixth paper concerns optimal linear-quadratic team-decision problems where the team does not have access to the model. Instead, the players must learn optimal policies by interacting with the environment. The paper contains algorithms and regret bounds for the first- and zeroth-order information feedback.

Acknowledgements

"There are so many people that have helped me during these years... I will probably forget to mention most of them; I will have to apologize instead."

"The department is such a lovely place to be; everyone is so open, supportive and things just work."

"Oh, I must thank Anders Rantzer, my supervisor, of course. Every PhD student must thank their supervisor... But what has he done? -well. He gave me some initial tasks to get me started, which was great. Then, when I started having some ideas of my own he encouraged me to investigate those, and suggested literature and people to talk to. He has an uncanny ability to figure out the theorem before the proof is even written, I wish I knew how to do that. It's like cheating on the exam. He's proofread most of my manuscripts, supported me by writing recommendation letters, introduced me to other researchers, and been a fantastic boss."

"I also have co-supervisors, Venkatraman (Venkat) Renganathan, and Bo (Bosse, Bob) Bernhardsson. They have also been very supportive, and we have had so many interesting discussions along the way. Bosse also helped me out a lot with the first course in Linear Functional analysis. I was quite underprepared... It was great to have someone who could explain what was actually going on."

"I wouldn't have applied for the position if not for the encouragement of Anton Cervin and Olof Troeng, I need to thank them too... And Anton has organized many of the after works and has a great taste in beer, this should also be acknowledged. Especially the latter part, which is a great quality in a coworker."

"Many of my mornings start with an invasion of Eva's office, where we gossip about the latest news, drinking the morning coffee that Eva, Nils, Anders or someone else has brewed. That's a fantastic way to start a day."

"Oh, and Emma helped me review the application for the Sweden-America foundation stipend. That was of great help. I got that stipend and could go to Caltech to take some courses. Richard Murray invited me to many of his group meetings, I learned so many new and interesting things there. And John Doyle and his students Lisa Li and Jing Yu were great collaborators in discussing the theory of large systems. John also taught me a lot about asking good research questions and cooking up counterexamples."

"Also, I need to thank Anders Nilsson for keeping my computer functional and Eva, Mika, Monika, and Cecilia for keeping the department functional."
"My family, especially my sambo, Ingrid, has shown great emotional support."
"I also need to disclose funding... but it's probably a bad idea to enclose that in quotation marks."

The author is a member of the ELLIIT Strategic Research Area at Lund University. 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

At the United Nations Summit 2015, our world leaders adopted 17 Sustainable Development Goals. Reaching these goals require efficient, reliable, and safe infrastructure. For example, Goal 7: Affordable and Clean Energy requires infrastructure insensitive to the loss of the inertia prevalent in conventional power plants, such as plants based on coal, gas, and nuclear power. As the nature of consumption and production changes, the networks' structures and underlying control mechanisms must keep up. Unfortunately, many anticipated changes increase the load and introduce additional complexity. Examples are micro-producers of electricity, autonomous vehicles in transportation networks, and increased nodes in communication networks. As complexity can increase by orders of magnitude, controlling these networks requires models at an entirely new scale; manually sustaining accurate models of all the individual components is infeasible. A solution is to use adaptation and learning to automatically learn and sustain models, taking care to do so in a reliable and scalable way.

The overarching topic of this thesis is the control and estimation of uncertain and large-scale dynamical systems. Our main tools in handling uncertainty are robustness and adaptation. Robustness concerns the resilience to deviations from initial assumptions, like a mismatch between the mathematical model and reality, on the system's behavior. Adaptation in a dynamical system means that past measurements are used to reduce uncertainty. However, adaptive components in control systems come with assumptions that may be difficult or impossible to validate and enforce.

Notation

This subsection briefly explains the mathematical notation used in the introduction. The transpose of a matrix A is denoted $A^\top$. For a vector x and matrix symmetric matrix Q of appropriate dimensions, $|x|_Q^2 = x^\top Q x$.

1.1 Estimation

Many interesting quantities cannot be directly measured online, such as the state of charge in a battery. One could, in principle, cut open the battery and measure the ion concentration, but that may have adverse effects on its life span. Differential equations can describe many physical systems, and if those differential equations are known, we can sometimes estimate the system's internal states from other observations and auxiliary quantities. In our battery example, that may be the resistance, capacitance, energy and temperature. Below we give an account of some estimation techniques related to this thesis, both for the case where the governing equations are fully known and when they are not.

The Kalman Filter

A clear majority of the work of estimation in dynamical systems is concerned with the stochastic framework. Let us recall the celebrated Kalman filter for the problem with known dynamics. Consider a linear system on state-space form

$$\begin{aligned} x_{t+1} &= Fx_t + Gu_t + w_t \\ y_t &= Hx_t + v_t. \end{aligned} \tag{1.1}$$

Here w_t, v_t are assumed zero mean, i.i.d. random variables, drawn from normal distributions with covariance matrices Σ_W and Σ_V, respectively. The initial state is assumed to be normally distributed, centered around x_0 with covariance P_0. The Kalman filter estimate of the state x_N then becomes $L\check{x}_N$ where $\check{x}_t$ is given by the recursion

$$\begin{aligned} P_0 &= P_0 \\ P_{t+1} &= \Sigma_w + FP_tF^\top - FP_tH^\top(\Sigma_V + HP_tH^\top)^{-1}HP_tF^\top \\ \check{x}_{t+1} &= F\check{x}_t + K_t(y_t - H\check{x}_t) + Gu_t, \quad \check{x}_0 = \hat{x}_0 \\ K_t &= FP_tH^\top(\Sigma_V + HP_tH^\top)^{-1}. \end{aligned}$$

Here K_t is called the Kalman filter gain and defines together with F, H, and G a state observer. The Kalman filter has numerous desirable properties; we will discuss some. Firstly, if Σ_v is positive definite, then the error dynamics $(x_t - \check{x}_t)$ are stable [Crassidis and Junkins, 2011]. Secondly, the estimate $\check{x}$ is the minimum-variance estimate, $\mathbb{E}[\check{x}_t] = x_t$ and $\mathbb{E}[(\check{x}_t - x_t)(\check{x}_t - x_t)^\top] = P_t$. Under the assumption of Gaussian noise, $\check{x}_t$ and x_t will be a sum of independent Gaussian variables and is therefore also normally distributed, or

$$p(x_t | y_0, \ldots, y_{t-1}, u_0, \ldots, u_{t-1}) \sim \mathcal{N}(\check{x}_t, P_t).$$

It is also least-squares optimal with respect to the prediction error and for the "smallest" disturbances that can explain the data. More formally, given an observed trajectory $(y_0, u_0), \dots, (y_{N-1}, u_{N-1})$, we have that, as a function of the final state x_T,

$$\inf_{w,v} \left\{ |x_0 - \hat{x}_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1} \left(|w_t|^2_{Q^{-1}} + |v_t|^2_{R^{-1}} \right\} \right.$$
$$= |x_N - \check{x}_N|^2_{P_N} + \sum_{t=0}^{N-1} |H\check{x}_t - y_t|^2_{(R+HP_tH^\top)^{-1}},$$

where the infimum is taken over arbitrary noise trajectories consistent with the observed data, the dynamics (1.1), and the final state. Note that in this *deterministic* setting, the trajectories $(w_t)_{t=0}^{N-1}$, $(v_t)_{t=0}^{N-1}$ consist of real-valued vectors but are otherwise arbitrary. The claim can be proven by forward dynamic programming [Simon, 2006] or by invoking results of optimal control for the tracking problem [Athans and Falb, 2007, Chapter 9.9] and applying them in reverse time.

Nonlinear filtering

The Kalman filter gives the minimum variance estimate for systems with linear dynamics and zero-mean disturbances with known covariance matrices. To compute the minimum variance estimate for a nonlinear system, we would need to propagate the state mean and variance through the dynamics and obtain the cross-covariance between the states and the outputs. Propagating distributions through nonlinear functions is quite challenging, so we have to resort to approximations, losing many guarantees.

The *extended Kalman filter* [Goodwin and Sin, 1984a] (EKF) works by linearizing around the state estimate at each point in time and applying the time-varying linear Kalman filter to the obtained dynmics. In essence, we propagate the mean and covariances through the linearized dynamics.

The *unscented Kalman filter* (UKF) [Julier and Uhlmann, 2004] works by representing the underlying probability distribution of the states by a finite set of deterministically chosen points called sigma points. These points are then propagated through the nonlinear dynamics and can be used to approximate the moments of the posterior distribution. The UKF is obtained by computing the minimum variance estimate based on these approximated moments and differs from the particle filter in that the sigma points are deterministically chosen. In contrast, the particles of a particle filter are randomly generated.

Joint- and dual Kalman filtering. A popular way to do simultaneous estimation of states and parameters is to augment the state vector with the uncertain elements of (F, G, H) and use nonlinear estimation techniques for this (bilinear) estimation problem. [Ljung, 1979] studies the convergence properties of join estimation using EKFs and shows that the algorithm can lead to biased estimates and diverge. Global convergence can be ensured by adding a correction term to the updates. We are unaware of any convergence studies of joint estimation using the UKF, although there seem to be many successful applications.

Dual Kalman filtering (DKF) was introduced in [Nelson and Stear, 1976]. The method works by running two interconnected filters, one for the uncertain parameters and one for the states. The state estimates are fed into the parameter estimation filter, whose parameter estimates are fed into the state estimation filter. DKF seems to be an approach of great practical value, and there are many extensions to nonlinear dynamics using the EKF [Wan and Nelson, 2000] and the UKF [Wan and Van Der Merwe, 2000].

Multiple-model estimation

Multiple-model adaptive estimation has been around since the '60s [Magill, 1965; Lainiotis, 1976] and has been an active research field since. 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. 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 the expected value of z_N conditioned on past measurements. The 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 nonlinear 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., 2021a; Deng et al., 2020] and showed improved accuracy in simulations.

1.2 Control

Robust control

Robust control emerged in the '70s as a response to the poor utilization and performance of the modern "optimal" multivariable control methods of the '60s. We consider a system robust if it is unlikely to fail. We quantify robustness as the ability to keep stability and performance invariant over a family of systems, usually quantified by similarity to a nominal one. Robust control has roots in early work on input/output descriptions by Zames and Sandberg [Zames, 1966], 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 in the textbooks [Francis, 1987; Zhou and Doyle, 1998; Dullerud and Paganini, 2013].

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 submultiplicative 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, 2013]. The Riccati equation approach strongly connects to the theory of dynamic differential games [Basar and Bernhard, 1995; Tadmor, 1993]. Game theoretic and passivity-based approaches can be extended to the nonlinear setting [James, 1995]. In the nonlinear setting, the Riccati equations are replaced by partial differential 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 if the block structure has a size no larger than three. 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.

The ℓ_1 optimal control problem was formulated in [Vidyasagar, 1986], where the authors solved the case for unstructured perturbations for SISO minimum-phase systems with at most one RHP zero. The solution is based on Youla-Kucera parameterization to transform the problem into an interpolation problem and considers both the discrete-time and the continuous-time cases. MIMO systems were considered in [Dahleh and Pearson, 1987] and sampled-data systems in [Bamieh et al., 1993; Dullerud and Francis, 1992]. Robust stability and performance with respect to structured uncertainty in the ℓ_1 was introduced in [Khammash and Pearson, 1991] and is described nicely in the tutorial paper [Dahleh and Khammash, 1993].

Adaptive control

Adaptive control dates back to the control of high-speed aircraft in the '50s. Due to the wide range of operating conditions, the performance of conventional controllers was unsatisfactory. Therefore, researchers and engineers started to look for controllers that could *adapt* to a changing environment online. Nowadays, many adaptive controller schemes and textbooks exist. We refer the reader to [Åström and Wittenmark, 1994] and [Goodwin and Sin, 1984b] for an introduction to the subject and its history. The topics of adaptive control that relates most closely to this thesis are those of *dual control* and *multiple-model adaptive control*.

Dual Control. In dual control [Feldbâum, 1963], one is tasked with minimizing a cumulative performance quantity given incomplete information about the dynamical equations governing the plant and exogenous signals affecting the system. Feldbaum showed that the optimal control policy will have two purposes and that the optimal controller is characterized by the solution to a functional equation (often called the Bellman equation). The first purpose is to excite the dynamics of the system (explore) in order to increase the knowledge about the system. The second purpose is to ensure satisfactory behavior of the system, that is, ensure that the performance quantity is small. In learning theory, this is sometimes called the exploration/exploitation trade-off.

Although the Bellman equation for the dual control problem is straightforward to derive, it is generally tough to compute due to the exponential growth of the state space (curse of dimensionality). Analytical or tractable numerical solutions have only been found in a few exceptional cases. We refer the reader to the review [Wittenmark, 1995] and [Åström and Wittenmark, 1994] for a more in-depth discussion on dual control and its approximations. The example in Paper III concerns an integrator with unknown gain. A similar problem was studied from the perspective of stochastic dual control in [Astrom and Helmersson, 1986].

Supervisory– and multiple model adaptive control. Multiple model adaptive control (MMAC) originated in the '70s as a way to control uncertain linear stochastic systems with parameter uncertainty, where the parameter uncertainty belonged 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 M was associated with a corresponding control law $u_M(\hat{x}_M)$ and a Kalman filter that estimates the states $\hat{x}_M(t)$. [Magill, 1965; Lainiotis, 1976] showed that the conditional probability, given measured signals, that a specific model is active can be expressed recursively using the Kalman filter residuals associated with each model. At each time instant, the control signal $u(t)$ was computed as the average of the ones for each model $u_M(t)$ weighted by the conditional probability of that model being active $\rho_{M|y(t),\ldots,y(0),u(t-1),\ldots,u(0)}$, that is

$$u(t) = \sum_M \rho_{M|y(t),\ldots,y(0),u(t-1),\ldots,u(0)} u_M(t).$$

It is not a dual control method, the chosen control law coincides with the last stage in a finite-time dual control problem; it is a myopic controller [Åström and Wittenmark, 1994]. In [Athans et al., 1977], the authors remark that they could not find rigorous proof of convergence of the claim that the probability associated with the actual model will asymptotically converge to unity. The claim is false, as seen by the following counterexample, which also shows that the MMAC is not necessarily stabilizing. The claim is, however, true if certain *distinguishability* conditions are fulfilled [Silvestre et al., 2020].

EXAMPLE 1—COUNTEREXAMPLE TO CONVERGENCE IN PROBABILITY
Consider the uncertain first-order linear stochastic discrete-time system

$$\begin{aligned} x(t+1) &= ax(t) \pm u(t) + w(t), \\ y(t) &= x(t) + v(t). \end{aligned} \tag{1.2}$$

In (1.2) $x(t), u(t), y(t) \in \mathbb{R}$ are the state, input and measured output. The process disturbance $w(t)$ and $v(t)$ 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 $u(0) = 0$. But 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". □

That careful tuning of the Kalman filters to prevent erroneous estimates and identifications is required to make MMAC behave reasonably well, is well known [Athans et al., 1977; Maybeck and Pogoda, 1989], but even then—the closed-loop system can exhibit unwanted oscillative behavior [Greene and Willsky, 1980].

The uncertain system (1.2) 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 (1.2) 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 1.1 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 relates the most 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.

Online learning. The learning community became interested in adaptive control problems not long ago. In particular, learning the linear-quadratic regulator from a single trajectory received much attention. It seems to have started with [Fiechter, 1997], who derived a probably-approximately-correct (PAC) bound with an explicit expression for the number of samples required for the *discounted* LQR problem.

REMARK 1
The discounted LQR problem entails minimizing the cost functional

$$J_\mu(x_0) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^k \left(|x_t|_Q^2 + |u_t|_R^2\right)\right],$$

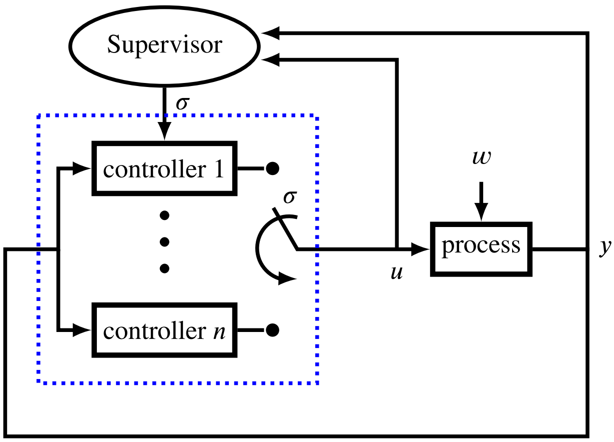


Figure 1.1 Illustration of a supervisory control architecture, recreated from [Hespanha, 2001]

over policies μ, subject to the dynamics

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t \\ u_t &= \mu(x_t). \end{aligned}$$

Here x_t is the state, w_t are i.i.d zero-mean noise sequences with $\mathbb{E}w_t w_t^\top = \sigma^2 I$ and $\gamma \in [0, 1)$ is an *discount factor*.

The discount factor is common in the Reinforcement learning literature. It enters the cost functional similarly to the *forgetting factor* (λ) prevalent in online identification algorithms like recursive least squares [Ljung, 1998]. However, the difference is that while the discount factor is smaller than 1, the forgetting factor is greater than one; conceptually, $\lambda \sim 1/\gamma$. The main implication for control is that $J_\mu(x_0) < \infty$ does not imply stability of the closed-loop for a discount factor $0 \leq \gamma < 1$. □

Regret (the difference between the total cost incurred and the cost of the optimal policy in hindsight) bounds for this "online LQR"-problem were first considered in [Abbasi-Yadkori and Szepesvári, 2011], based on the *Optimism in the Face of Uncertainty* (OFU) principle. The OFU principle dates back to and is a heuristic method to balance exploration and exploitation in learning systems. It entails continuously characterizing a high-probability set of model parameters given data and basing decisions on the parameter values that would give the lowest cost.

The model-free case was considered in [Abbasi-Yadkori et al., 2019]. [Cohen et al., 2018] and [Agarwal et al., 2019b] studied the case where the dynamics are known, but the cost matrices Q_t, R_t are time-varying and revealed only after each u_t has been selected. Regret in the setting where disturbances and time-varying convex loss-functions are chosen adversarially from bounded sets were introduced in [Agarwal et al., 2019a], and extended to unknown dynamics in [Chen and Hazan, 2021].

Large-scale systems

As systems grow in complexity, new problems arise. Standard methods for controller synthesis may become *computationally intractable*; we require synthesis methods that can handle a large number of states. For standard control , the speed of the electronics is usually significantly faster (almost the speed of light) than the controlled system. However, for large systems, the control signals and measurement signals may have to travel large distances and experience *significant delay*. Global communication may be infeasible due to computational tractability, limited transmission capacity but also for legal- and integrity reasons. Control may be constrained to use only *local information*.

These issues motivate the search for methods to analyze and control systems using information from local measurements and information exchange with the closest neighbors. Furthermore, we insist that such controllers should be synthesized locally—exploiting the computational resources in each node.

Team-theory and quadratic invariance. Team decision problems originate from economics, where optimal decentralized decisions in organizations were studied in the papers by [Marschak, 1955] and [Radner, 1962] under stochastic settings. A major contribution of Radner was that the case the linear-quadratic case admits a unique optimal decision policy that is linear in the measurements. In these studies, the agents in the team know the problem parameters. The agents use the information of the problem parameters to find the optimal decentralized decision. Decentralized decisions only depend on local measurements of the state of nature, where the measurements of the agents are typically different. [Gattami et al., 2012] studied linear quadratic robust team decision problems and showed that optimal decisions are linear and can be found by solving a convex (in fact, semi-definite) optimization problem. Team-decision theory has been helpful in understanding distributed control [Mahajan et al., 2012]. Witsenhausens famous counterexample [Witsenhausen, 1968] established that linear decisions are not always optimal for *distributed* LQG problems and sparked an interest in research of team problems in the control community. [Ho and Chu, 1972] showed how linear-quadratic problems with partially nested information can be rewritten as static team-decision problems and Witsenhausen showed that a general class of dynamic team decision problems can be reduced to static ones via a change of measures [Witsenhausen, 1988]. Static reductions for more exotic information structures is still an active research field [Gupta et al., 2014; Sanjari et al., 2021].

The reduction to static team problems by Ho and Chu is equivalent to Q-parameterization, a simplification of Youla parameterization [Youla et al., 1976], and was recently rediscovered in the online learning literature [Simchowitz et al., 2020], under the name “nature’s y’s”. The value in the reduction is that the optimal solution, expressed in the Q-parameter, is linear and unique. Intuitively, the reparametrization subtracts the influence of the other players on each player’s decision, removing the necessity for the player to “guess” the decisions of the other

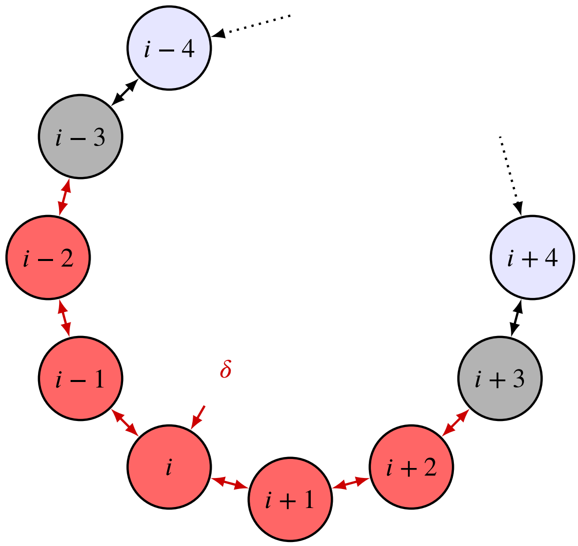


Figure 1.2 Illustration of disturbance localization. A shock δ entering in the ith node spreads to the neighbors at most three hops away, where it is completely canceled.

players. It is the estimation of other players' actions that destroys linearity in distributed control problems. It is sometimes referenced to as the signaling incentive. Q-parameterization is an affine parameterization of all achievable stable closed-loop systems, which sometimes greatly simplifies controller synthesis. In distributed control, one aims to impose constraints on the controller architecture, that is, restrict the set of feasible controllers to those that satisfy different types of communication constraints. A reasonable attempt to simplify controller synthesis is to translate the communication constraints into constraints on the Q-parameter. Convexity of the feasible set of Q-parameters was introduced in [Rotkowitz and Lall, 2006] and shown to be necessary and sufficient for the constraint set of the Q-parameter to be *equal* to that of the controller. [Rotkowitz et al., 2010] showed that communication constraints in distributed control are quadratically invariant as long as information between controllers propagates at least as fast as the dynamics propagetes between links. [Lessard and Lall, 2011] proved that quadratic invariance is both necessary and sufficient for convexity of the constraint set on the Q-parameter, but noted that there are cases when the constraint fails to be quadratically invariant. Yet the closed-loop maps are an affine set. These situations can be handled by *System-Level Synthesis*.

System-level synthesis. The main idea of system-level synthesis is to optimize directly over the closed-loop transfer functions from disturbances to states and control signals. Sparsity constraints on the closed-loop maps can then be translate into a sparse realization of the controller. The set of closed-loop responses that can be achieved with causal linear feedback is described as the kernel of an affine map. In

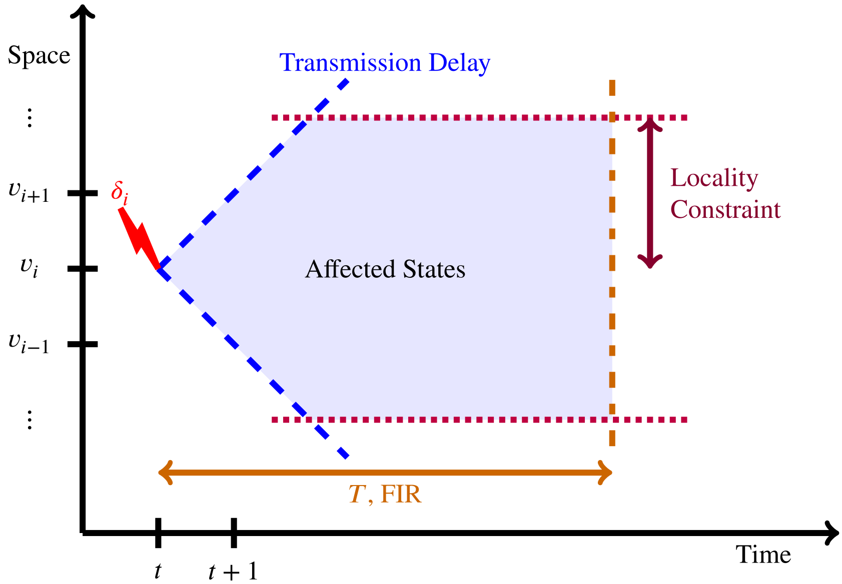


Figure 1.3 Localization in system-level synthesis

contrast, the celebrated Youla-parametrization characterizes the achievable closed-loop responses as the range of an affine map. The kernel representation has a number of advantages over the image representation; we list a few. Firstly, it seems simpler for unstable systems. Secondly, the kernel description can be decomposed into two equality constraints, one which is *column separable* and one that is *row separable*. If the control objective and other constraints can be similarly decomposed, one can use operator splitting methods to synthesize controllers in a scalable manner. Finally, there exists useful bounds on the degradation when constraints can only be approximately fulfilled.

Communication constraints among subsystems in the controller are encoded directly in the closed-loop responses. For example, the disturbance localization requirement in Figure 1.2 implies that no communication has to be sent from the ith any node further than three hops away. The reason is that whatever happens in the ith node is constrained not to affect nodes that far away, implying that one can remove the *signaling incentive* by letting subsystem i communicate their decisions to systems $i \pm 1$, $i \pm 2$ and $i \pm 3$. Similarly, one can encode delayed communication as requirements on the closed-loop response from disturbances to control action as a requirement that a shock entering in the ith node at time t should spread "slowly" to other subsystems, see Figure 1.3.

System-level synthesis started with a series of conference contributions by Wang, Matni, You, and Doyle between 2014–2017, culminating in the PhD thesis [Wang, 2017]. The research up until 2019 is nicely presented in the review article [Anderson et al., 2019]. Since then, system-level synthesis has been combined with Willem's fundamental lemma and incorporated into the data-driven framework [Xue and Matni, 2021],used to study the internal feedbacks prevalent in human vision [Lisa Li, 2022],model-predictive control [Chen et al., 2020; Alonso and Matni, 2020; Alonso et al., 2022], safe statistical learning of controllers [Dean et al., 2019] and many other control-related questions [Li et al., 2021b].

All the contributions above employ a finite impulse-response restriction of the closed-loop dynamics to turn the controller synthesis problem into a *finite-dimensional* convex optimization problem. Such optimization problems can be efficiently solved using off-the-shelf software. However, controller synthesis that avoids the finite impulse-response restriction remains largely unexplored. A notable exception is [Yu et al., 2021], where the authors solve the $\mathcal{H}_2$ optimal control problem under communication constraints but with instantaneous communication using dynamic programming over the impulse responses.

Positivity and symmetry. A tangential question is: "For what systems do the optimal centralized solutions have distributed realizations?". In other words, are there systems and problem formulations that, without explicit sparsity constraints, have sparse solutions? The answer is yes [Bergeling, 2019]; positive systems [Rantzer and Valcher, 2018] that also fulfill some structural conditions have distributed solutions that are centrally optimal.

2

Contributions

Papers I, II, and III consider uncertain linear systems with uncertain parameters belonging to a (known) finite set. These papers are part of an effort to extend minimax adaptive control [Rantzer, 2020; Rantzer, 2021] to the output feedback case.

Paper I. O. Kjellqvist and A. Rantzer (2022c). "Minimax adaptive estimation for finite sets of linear systems". In: *2022 American Control Conference (ACC)*, pp. 260–265. DOI: 10.23919/ACC53348.2022.9867474,
This paper concerns 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 the gain from unmeasured disturbances to the output prediction error is bounded by a constant, γ, provided that γ is a feasible gain bound. We also show how to evaluate online whether γ_N is a feasible gain bound.

The problem formulation extends [Başar, 1991] from the case with known dynamics to the case where the dynamics are uncertain.

Paper II. O. Kjellqvist (2023). *Fundamental worst-case performance limits for multiple-model estimation*. Submitted to IFAC World Congress 2023,
This paper extends Paper I in two ways. Firstly, we consider strictly causal state estimation. Secondly, we provide upper and lower bounds on the *achievable* disturbance gains γ_N (from disturbances to Nth time-step estimation error). The upper and lower bounds are characterized by one forward Riccati recursion per model and one backward Riccati recursion per *pair* of feasible models.

The achievable attenuation level is informative as to whether the set of feasible models is suitable for state estimation. The conclusion is independent of the estimation procedure. In essence, γ_N will be large if highly different state trajectories well explain the same output trajectory for the models. By well explained, we mean that the noise and disturbance trajectories required to explain the data are small (in a least-squares sense). The gain will be small if the same output trajectory requires large disturbance trajectories to be compatible with different state trajectories. We want to point out that the gain-bound γ_N 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., 2020] alone is neither sufficient nor necessary to guarantee that the estimation error will be small.

In practical multiple-model settings, one is often interested in detecting which model has generated the data. For instance, in fault detection, one is interested to know whether one of the sensors has gone bad. The work in papers I and II is limited to the state estimate only and does not consider the ability to identify the underlying model.

Paper III. O. Kjellqvist and A. Rantzer (2022a). "Learning-enabled robust control with noisy measurements". In: *Learning for Dynamics and Control Conference*. PMLR, pp. 86–96,
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 (2022b). *Learning-enabled robust control with noisy measurements*. DOI: 10.48550/ARXIV.2202.08363. URL: https://arxiv.org/abs/2202.08363.

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 memory-less function of a $\mathcal{H}_\infty$ *multi observer*, so 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 has unknown sign. The controllers are of the certainty-equivalence type, which in this case coincides with a multiple-model adaptive (supervisory) control architecture.

The approach differs from supervisory control and multiple-model adaptive control in that we start with the desired property (finite ℓ_2-gain) and characterize controllers that achieve this property. That is, we do not impose a supervisory structure. Instead, the resulting controller architecture is a suboptimal solution to the equivalent problem formulation (2). The literature on supervisory control and multiple-model adaptive control starts with an imposed architecture and proceeds to prove that the resulting controller achieves a bounded ℓ_2-gain.

Paper IV. O. Kjellqvist and J. C. Doyle (2022). *ν-analysis: a new notion of robustness for large systems with structured uncertainties*. To appear in IEEE CDC 2022,
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 were aiming for a convex and separable quantity, so one can synthesize controllers locally. This work was performed with Lisa Li for a course at the California Institute of Technology given by John Doyle. In the end, it resulted in two papers. The first paper (this paper) proposes and analyzes the robustness measure ν. The second paper [Li and Doyle, 2022] uses the results of the first paper to synthesize controllers in an iterative manner, similar to D-K synthesis.

Paper V. O. Kjellqvist and J. Yu (2022). *On infinite-horizon system level synthesis problems*. To appear in IEEE CDC 2022,
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 the controllers 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.

Paper VI. O. Kjellqvist and A. Gattami (2022). *Learning optimal team decisions*. To appear in IEEE CDC 2022,
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 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 the sampling does not require any coordination between players.

A limitation of this work is that static reduction of [Ho and Chu, 1972] requires knowledge of the problem parameters. Therefore, it needs to be clarified how to apply these results to sequential team-decision problems, like the control of an unknown linear system.

Other work not included in this thesis. O. Kjellqvist and O. Troeng (2020). "Numerical pitfalls in q-design". *IFAC-PapersOnLine* **53**:2. 21st IFAC World Congress, pp. 4404–4408. ISSN: 2405-8963. DOI: `https://doi.org/10.1016/j.ifacol.2020.12.368`. URL: `https://www.sciencedirect.com/science/article/pii/S2405896320306522` and
V. Renganathan et al. (2022). "Distributed implementation of minimax adaptive controller for finite set of linear systems". *arXiv preprint arXiv:2210.00081*

Bibliography

Abbasi-Yadkori, Y., N. Lazic, and C. Szepesvari (2019). "Model-free linear quadratic control via reduction to expert prediction". In: Chaudhuri, K. et al. (Eds.). *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*. Vol. 89. Proceedings of Machine Learning Research. PMLR, pp. 3108–3117. URL: `https://proceedings.mlr.press/v89/abbasi-yadkori19a.html`.

Abbasi-Yadkori, Y. and C. Szepesvári (2011). "Regret bounds for the adaptive control of linear quadratic systems". In: Kakade, S. M. et al. (Eds.). *Proceedings of the 24th Annual Conference on Learning Theory*. Vol. 19. Proceedings of Machine Learning Research. 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 Transactions on Automatic Control* **15**:1, pp. 10–17. DOI: `10.1109/TAC.1970.1099359`.

Agarwal, N., B. Bullins, E. Hazan, S. Kakade, and K. Singh (2019a). "Online control with adversarial disturbances". In: Chaudhuri, K. et al. (Eds.). *Proceedings of the 36th International Conference on Machine Learning*. Vol. 97. Proceedings of Machine Learning Research. PMLR, pp. 111–119. URL: `https://proceedings.mlr.press/v97/agarwal19c.html`.

Agarwal, N., E. Hazan, and K. Singh (2019b). "Logarithmic regret for online control". In: Wallach, H. et al. (Eds.). *Advances in Neural Information Processing Systems*. Vol. 32. Curran Associates, Inc. URL: `https://proceedings.neurips.cc/paper/2019/file/78719f11fa2df9917de3110133506521-Paper.pdf`.

Akca, A. and M. Ö. Efe (2019). "Multiple model kalman and particle filters and applications: a survey". *IFAC-PapersOnLine* **52**:3. 15th IFAC Symposium on Large Scale Complex Systems LSS 2019, pp. 73–78. ISSN: 2405-8963. DOI: `https://doi.org/10.1016/j.ifacol.2019.06.013`. URL: `https://www.sciencedirect.com/science/article/pii/S2405896319300977`.

Alonso, C. A. and N. Matni (2020). "Distributed and localized closed loop model predictive control via system level synthesis". In: *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 5598–5605. DOI: 10.1109/CDC42340.2020.9303936.

Alonso, C. A., F. Yang, and N. Matni (2022). "Data-driven distributed and localized model predictive control". *IEEE Open Journal of Control Systems* **1**, pp. 29–40. DOI: 10.1109/OJCSYS.2022.3171787.

Anderson, J., J. C. Doyle, S. H. Low, and N. Matni (2019). "System level synthesis". *Annual Reviews in Control* **47**, pp. 364–393. ISSN: 1367-5788. DOI: https://doi.org/10.1016/j.arcontrol.2019.03.006. URL: https://www.sciencedirect.com/science/article/pii/S1367578819300215.

Astrom, K. and A. Helmersson (1986). "Dual control of an integrator with unknown gain". *Computers & Mathematics with Applications* **12**:6, Part A, pp. 653–662. ISSN: 0898-1221. DOI: https://doi.org/10.1016/0898-1221(86)90052-0. URL: https://www.sciencedirect.com/science/article/pii/0898122186900520.

Åström, K. J. and B. Wittenmark (1994). *Adaptive Control*. 2nd. Addison-Wesley Longman Publishing Co., Inc., USA. ISBN: 0201558661.

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 Transactions on Automatic Control* **22**:5, pp. 768–780. DOI: 10.1109/TAC.1977.1101599.

Athans, M. and L. P. Falb (2007). *Optimal control an introduction to the theory and its applications*. Originally published: New York, McGraw-Hill, 1966. Dover Publications, New York.

Bamieh, B., M. Dahleh, and J. Pearson (1993). "Minimization of the $\mathcal{L}_\infty$-induced norm for sampled-data systems". *IEEE Transactions on Automatic Control* **38**:5, pp. 23–. DOI: 10.1109/9.277236.

Basar, T. and P. Bernhard (1995). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach*. Birkhauser.

Başar, T. (1991). "Optimum performance levels for minimax filters, predictors and smoothers". *Systems & Control Letters* **16**:5, pp. 309–317. ISSN: 0167-6911. DOI: https://doi.org/10.1016/0167-6911(91)90052-G. URL: https://www.sciencedirect.com/science/article/pii/016769119190052G.

Bergeling, C. (2019). *On H-infinity Control and Large-Scale Systems*. English. PhD thesis. Department of Automatic Control. ISBN: 978-91-7895-095-9.

Blom, H. A. P. and Y. Bar-Shalom (1988). "The interacting multiple model algorithm for systems with markovian switching coefficients". *IEEE Transactions on Automatic 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 Transactions on Automatic 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 Transactions on Automatic Control* **61**:3, pp. 693–708. DOI: `10.1109/TAC.2015.2492503`.

Chang, C. B. and M. Athans (1978). "State estimation for discrete systems with switching parameters". *IEEE Transactions on Aerospace and Electronic Systems* **AES-14**:3, pp. 418–425. DOI: `10.1109/TAES.1978.308603`.

Chen, S., H. Wang, M. Morari, V. M. Preciado, and N. Matni (2020). "Robust closed-loop model predictive control via system level synthesis". In: *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 2152–2159. DOI: `10.1109/CDC42340.2020.9304200`.

Chen, X. and E. Hazan (2021). "Black-box control for linear dynamical systems". In: Belkin, M. et al. (Eds.). *Proceedings of Thirty Fourth Conference on Learning Theory*. Vol. 134. Proceedings of Machine Learning Research. PMLR, pp. 1114–1143. URL: `https://proceedings.mlr.press/v134/chen21c.html`.

Cohen, A., A. Hassidim, T. Koren, N. Lazic, Y. Mansour, and K. Talwar (2018). "Online linear quadratic control". In: *ICML*.

Crassidis, J. L. and J. L. Junkins (2011). *Optimal Estimation of Dynamic Systems, Second Edition (Chapman & Hall/CRC Applied Mathematics & Nonlinear Science)*. 2nd. Chapman & Hall/CRC. ISBN: 1439839859.

Dahleh, M. and J. Pearson (1987). "ℓ^1-optimal feedback controllers for mimo discrete-time systems". *IEEE Transactions on Automatic Control* **32**:4, pp. 314–322. DOI: `10.1109/TAC.1987.1104603`.

Dahleh, M. A. and M. H. Khammash (1993). "Controller design for plants with structured uncertainty". *Autom.* **29**, pp. 37–56.

Dean, S., S. Tu, N. Matni, and B. Recht (2019). "Safely learning to control the constrained linear quadratic regulator". In: *2019 American Control Conference (ACC)*, pp. 5582–5588. DOI: `10.23919/ACC.2019.8814865`.

Deng, L., D. Li, and R. Li (2020). "Improved IMM algorithm based on RNNs". *Journal of Physics: Conference Series* **1518**, p. 012055. DOI: `10.1088/1742-6596/1518/1/012055`. URL: `https://doi.org/10.1088/1742-6596/1518/1/012055`.

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

Doyle, J. C. (1982). "Analysis of feedback systems with structured uncertainty".

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

Dullerud, G. and F. Paganini (2013). *A Course in Robust Control Theory: A Convex Approach*. Texts in Applied Mathematics. Springer New York. ISBN: 9781475732900. URL: `https://books.google.se/books?id=KGbTBwAAQBAJ`.

Dullerud, G. and B. Francis (1992). "$\mathcal{L}_1$ analysis and design of sampled-data systems". English (US). *IRE Transactions on Automatic Control* **37**:4, pp. 436–446. ISSN: 0018-9286. DOI: `10.1109/9.126577`.

Fan, M. and A. Tits (1986). "Characterization and efficient computation of the structured singular value". *IEEE Transactions on Automatic 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 Transactions on Automatic 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: *1988 American Control Conference*, pp. 1195–1200. DOI: `10.23919/ACC.1988.4789902`.

Feintuch, A. (1998). *Robust control theory in Hilbert space*. Springer.

Feldbâum, A. (1963). "Dual control theory problems". *IFAC Proceedings Volumes* **1**:2. 2nd International IFAC Congress on Automatic and Remote Control: Theory, Basle, Switzerland, 1963, pp. 541–550. ISSN: 1474-6670. DOI: `https://doi.org/10.1016/S1474-6670(17)69687-3`. URL: `https://www.sciencedirect.com/science/article/pii/S1474667017696873`.

Fiechter, C.-N. (1997). "Pac adaptive control of linear systems". In: *COLT '97*.

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

Gattami, A., B. M. Bernhardsson, and A. Rantzer (2012). "Robust team decision theory". *IEEE Transactions on Automatic Control* **57**:3, pp. 794–798. DOI: `10.1109/TAC.2011.2168071`.

Goodwin, G. C. and K. S. Sin (1984a). *Adaptive filtering prediction and control / Graham C. Goodwin and Kwai Sang Sin.* eng. Prentice-Hall information and system sciences series. Prentice-Hall, Englewood Cliffs, N.J. ISBN: 013004069X.

Goodwin, G. C. and K. S. Sin (1984b). *Adaptive filtering prediction and control / Graham C. Goodwin and Kwai Sang Sin.* eng. Prentice-Hall information and system sciences series. Prentice-Hall, Englewood Cliffs, N.J. ISBN: 013004069X.

Greene, C. S. and A. S. Willsky (1980). "An analysis of the multiple model adaptive control algorithm". In: *1980 19th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, pp. 1142–1145. DOI: `10.1109/CDC.1980.271982`.

Gupta, A., S. Yuksel, T. Başar, and C. Langbort (2014). "On the existence of optimal policies for a class of static and sequential dynamic teams". *SIAM Journal on Control and Optimization* **53**. DOI: `10.1137/14096534X`.

Hespanha, J. P. (2001). “Tutorial on supervisory control”. In: *40th Conf. on Decision and Control*. Lecture notes for the workshop *Control using Logic and Switching*.

Ho, Y.-C. and K. Chu (1972). “Team decision theory and information structures in optimal control problems–part i”. *IEEE Transactions on Automatic Control* **17**:1, pp. 15–22. DOI: 10.1109/TAC.1972.1099850.

James, M. R. (1995). “Recent developments in nonlinear $\mathcal{H}_\infty$ control”. *IFAC Proceedings Volumes* **28**:14. 3rd IFAC Symposium on Nonlinear Control Systems Design 1995, Tahoe City, CA, USA, 25-28 June 1995, pp. 541–552. ISSN: 1474-6670. DOI: https://doi.org/10.1016/S1474-6670(17)46885-6. URL: https://www.sciencedirect.com/science/article/pii/S1474667017468856.

Julier, S. and J. Uhlmann (2004). “Unscented filtering and nonlinear estimation”. *Proceedings of the IEEE* **92**:3, pp. 401–422. DOI: 10.1109/JPROC.2003.823141.

Khammash, M. H. and J. B. Pearson (1991). “Performance robustness of discrete-time systems with structured uncertainty”. *IEEE Transactions on Automatic Control* **36**, pp. 398–412.

Khargonekar, P., K. Poolla, and A. Tannenbaum (1985). “Robust control of linear time-invariant plants using periodic compensation”. *IEEE Transactions on Automatic 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 Transactions on Automatic Control* **33**:12, pp. 1158–1161. DOI: 10.1109/9.14439.

Kjellqvist, O. (2023). *Fundamental worst-case performance limits for multiple-model estimation*. Submitted to IFAC World Congress 2023.

Kjellqvist, O. and J. C. Doyle (2022). *ν-analysis: a new notion of robustness for large systems with structured uncertainties*. To appear in IEEE CDC 2022.

Kjellqvist, O. and A. Gattami (2022). *Learning optimal team decisions*. To appear in IEEE CDC 2022.

Kjellqvist, O. and A. Rantzer (2022a). “Learning-enabled robust control with noisy measurements”. In: *Learning for Dynamics and Control Conference*. PMLR, pp. 86–96.

Kjellqvist, O. and A. Rantzer (2022b). *Learning-enabled robust control with noisy measurements*. DOI: 10.48550/ARXIV.2202.08363. URL: https://arxiv.org/abs/2202.08363.

Kjellqvist, O. and A. Rantzer (2022c). “Minimax adaptive estimation for finite sets of linear systems”. In: *2022 American Control Conference (ACC)*, pp. 260–265. DOI: 10.23919/ACC53348.2022.9867474.

Kjellqvist, O. and O. Troeng (2020). “Numerical pitfalls in q-design”. *IFAC-PapersOnLine* **53**:2. 21st IFAC World Congress, pp. 4404–4408. ISSN: 2405-8963. DOI: https://doi.org/10.1016/j.ifacol.2020.12.368. URL: https://www.sciencedirect.com/science/article/pii/S2405896320306522.

Kjellqvist, O. and J. Yu (2022). *On infinite-horizon system level synthesis problems*. To appear in IEEE CDC 2022.

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

Lessard, L. and S. Lall (2011). “Quadratic invariance is necessary and sufficient for convexity”. In: *Proceedings of the 2011 American Control Conference*, pp. 5360–5362. DOI: 10.1109/ACC.2011.5990928.

Li, D., P. Zhang, and R. Li (2021a). “Improved IMM algorithm based on XGBoost”. *Journal of Physics: Conference Series* **1748**, p. 032017. DOI: 10.1088/1742-6596/1748/3/032017. URL: https://doi.org/10.1088/1742-6596/1748/3/032017.

Li, J. S., C. A. Alonso, and J. C. Doyle (2021b). “Frontiers in scalable distributed control: sls, mpc, and beyond”. In: *2021 American Control Conference (ACC)*, pp. 2720–2725. DOI: 10.23919/ACC50511.2021.9483130.

Li, J. S. and J. C. Doyle (2022). “Distributed robust control for systems with structured uncertainties”. In: *61th IEEE Conference on Decision and Control (CDC)*. To appear.

Liberzon, M. R. (2001). “Lur’e problem of absolute stability - a historical essay”. *IFAC Proceedings Volumes* **34**:6. 5th IFAC Symposium on Nonlinear Control Systems 2001, St Petersburg, Russia, 4-6 July 2001, pp. 25–28. ISSN: 1474-6670. DOI: https://doi.org/10.1016/S1474-6670(17)35141-8. URL: https://www.sciencedirect.com/science/article/pii/S1474667017351418.

Lisa Li, J. S. (2022). “Internal feedback in biological control: locality and system level synthesis”. In: *2022 American Control Conference (ACC)*, pp. 474–479. DOI: 10.23919/ACC53348.2022.9867769.

Ljung, L. (1979). “Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems”. *IEEE Transactions on Automatic Control* **24**:1, pp. 36–50. DOI: 10.1109/TAC.1979.1101943.

Ljung, L. (1998). “System identification”. In: Procházka, A. et al. (Eds.). *Signal Analysis and Prediction*. Birkhäuser Boston, Boston, MA, pp. 163–173. ISBN: 978-1-4612-1768-8. DOI: 10.1007/978-1-4612-1768-8_11. URL: https://doi.org/10.1007/978-1-4612-1768-8_11.

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

Mahajan, A., N. C. Martins, M. C. Rotkowitz, and S. Yüksel (2012). "Information structures in optimal decentralized control". In: *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 1291–1306. DOI: 10.1109/CDC.2012.6425819.

Marschak, J. (1955). "Elements for a theory of teams". *Management Science* **1**:2, pp. 127–137. URL: https://EconPapers.repec.org/RePEc:inm:ormnsc:v:1:y:1955:i:2:p:127-137.

Mårtensson, B. (1985). "The order of any stabilizing regulator is sufficient a priori information for adaptive stabilization". *Systems & Control Letters* **6**:2, pp. 87–91. ISSN: 0167-6911. DOI: https://doi.org/10.1016/0167-6911(85)90002-7. URL: https://www.sciencedirect.com/science/article/pii/0167691185900027.

Maybeck, P. and D. Pogoda (1989). "Multiple model adaptive controller for the stol f-15 with sensor/actuator failures". In: *Proceedings of the 28th IEEE Conference on Decision and Control,* 1566–1572 vol.2. DOI: 10.1109/CDC.1989.70412.

Megretski, A. and A. Rantzer (1997). "System analysis via integral quadratic constraints". *IEEE Transactions on Automatic Control* **42**:6, pp. 819–830. DOI: 10.1109/9.587335.

Morse, A. (1996). "Supervisory control of families of linear set-point controllers - part i. exact matching". *IEEE Transactions on Automatic 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 Transactions on Automatic Control* **42**:11, pp. 1500–1515. DOI: 10.1109/9.649687.

Nelson, L. and E. Stear (1976). "The simultaneous on-line estimation of parameters and states in linear systems". *IEEE Transactions on Automatic Control* **21**:1, pp. 94–98. DOI: 10.1109/TAC.1976.1101148.

Packard, A., M. Fan, and J. Doyle (1988). "A power method for the structured singular value". In: *Proceedings of the 27th IEEE Conference on Decision and Control*, 2132–2137 vol.3. DOI: 10.1109/CDC.1988.194710.

Poola, K. and A. Tikku (1995). "Robust performance against time-varying structured perturbations". *IEEE Transactions on Automatic 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.

Radner, R. (1962). "Team Decision Problems". *The Annals of Mathematical Statistics* **33**:3, pp. 857–881. DOI: 10.1214/aoms/1177704455. URL: https://doi.org/10.1214/aoms/1177704455.

Rantzer, A. (2020). *Minimax adaptive control for state matrix with unknown sign*. arXiv: 1912.03550 [math.OC].

Rantzer, A. (2021). *Minimax adaptive control for a finite set of linear systems*. arXiv: 2011.10814 [math.OC].

Rantzer, A. and M. E. Valcher (2018). "A tutorial on positive systems and large scale control". In: *2018 IEEE Conference on Decision and Control (CDC)*, pp. 3686–3697. DOI: 10.1109/CDC.2018.8618689.

Renganathan, V., A. Rantzer, and O. Kjellqvist (2022). "Distributed implementation of minimax adaptive controller for finite set of linear systems". *arXiv preprint arXiv:2210.00081*.

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

Rotkowitz, M. and S. Lall (2006). "A characterization of convex problems in decentralized control". *IEEE Transactions on Automatic Control* **51**:2, pp. 274–286. DOI: 10.1109/TAC.2005.860365.

Rotkowitz, M., R. Cogill, and S. Lall (2010). "Convexity of optimal control over networks with delays and arbitrary topology". *International Journal of Systems, Control and Communications* **2**:1-3, pp. 30–54. DOI: 10.1504/IJSCC.2010.031157.

Safonov, M. G. (1978). In: *1978 Allerton Conference on Communication, Control and Computing*, pp. 451–460.

Safonov, M. G. (1981). "Stability margins of diagonally perturbed multivariable feedback systems". In: *1981 20th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, pp. 1472–1478. DOI: 10.1109/CDC.1981.269503.

Safonov, M. G. (2012). "Origins of robust control: early history and future speculations". *Annual Reviews in Control* **36**:2, pp. 173–181. ISSN: 1367-5788. DOI: https://doi.org/10.1016/j.arcontrol.2012.09.001. URL: https://www.sciencedirect.com/science/article/pii/S1367578812000363.

Sanjari, S., T. Başar, and S. Yüksel (2021). "Policy-dependent and policy-independent static reduction of stochastic dynamic teams and games and fragility of equivalence properties". In: *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 6231–6236. DOI: 10.1109/CDC45484.2021.9683260.

Shamir, O. (2013). "On the complexity of bandit and derivative-free stochastic convex optimization". In: *Conference on Learning Theory*. PMLR, pp. 3–24.

Shamma, J. (1994). "Robust stability with time-varying structured uncertainty". *IEEE Transactions on Automatic Control* **39**:4, pp. 714–724. DOI: 10.1109/9.286248.

Silvestre, D., P. Rosa, and C. Silvestre (2020). "Distinguishability of discrete-time linear systems". *International Journal of Robust and Nonlinear Control* **31**. DOI: 10.1002/rnc.5367.

Simchowitz, M., K. Singh, and E. Hazan (2020). "Improper learning for non-stochastic control". In: *COLT*.

Simon, D. (2006). *Optimal State Estimation, Kalman, H_∞, and Nonlinear Approaches*. John Wiley & Sons, Inc.

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

Vidyasagar, M. (1986). "Optimal rejection of persistent bounded disturbances". *IEEE Transactions on Automatic Control* **31**:6, pp. 527–534. DOI: 10.1109/TAC.1986.1104315.

Vinnicombe, G. (2004). "Examples and counterexamples in finite l2-gain adaptive control".

Wan, E. and R. Van Der Merwe (2000). "The unscented kalman filter for nonlinear estimation". In: *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, pp. 153–158. DOI: 10.1109/ASSPCC.2000.882463.

Wan, E. and A. Nelson (2000). "Dual kalman filtering methods for nonlinear prediction, smoothing, and estimation". *Advances in Neural Information Processing Systems*.

Wang, Y.-S. (2017). *A System Level Approach to Optimal Controller Design for Large-Scale Distributed Systems*. PhD thesis. Caltech.

Wang, Y.-S., N. Matni, and J. C. Doyle (2018). "Separable and localized system-level synthesis for large-scale systems". *IEEE Transactions on Automatic Control* **63**:12, pp. 4234–4249. DOI: 10.1109/TAC.2018.2819246.

Witsenhausen, H. S. (1968). "A counterexample in stochastic optimum control". *SIAM Journal on Control* **6**:1, pp. 131–147. DOI: 10.1137/0306011.

Witsenhausen, H. S. (1988). "Equivalent stochastic control problems". *Mathematics of Control, Signals and Systems* **1**, pp. 3–11.

Wittenmark, B. (1995). "Adaptive dual control methods: an overview". *IFAC Proceedings Volumes* **28**:13. 5th IFAC Symposium on Adaptive Systems in Control and Signal Processing 1995, Budapest, Hungary, 14-16 June, 1995, pp. 67–72. ISSN: 1474-6670. DOI: https://doi.org/10.1016/S1474-6670(17)45327-4. URL: https://www.sciencedirect.com/science/article/pii/S1474667017453274.

Xiong, K., C. Wei, and L. Liu (2015). "Robust multiple model adaptive estimation for spacecraft autonomous navigation". *Aerospace Science and Technology* **42**, pp. 249–258. ISSN: 1270-9638. DOI: https://doi.org/10.1016/j.ast.2015.01.021. URL: https://www.sciencedirect.com/science/article/pii/S1270963815000371.

Xue, A. and N. Matni (2021). "Data-driven system level synthesis". In: *L4DC*.

Youla, D., J. Bongiorno, and H. Jabr (1976). "Modern wiener–hopf design of optimal controllers parts i and 2". *IEEE Transactions on Automatic Control* **21**:1, pp. 3–13. DOI: 10.1109/TAC.1976.1101139.

Young, P. and J. Doyle (1990). "Computation of mu with real and complex uncertainties". In: *29th IEEE Conference on Decision and 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: *1992 American Control Conference*, pp. 2190–2194. DOI: 10.23919/ACC.1992.4792521.

Yu, J., Y.-S. Wang, and J. Anderson (2021). "Localized and distributed $\mathcal{H}_2$ state feedback control". In: *2021 American Control Conference (ACC)*, 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 Transactions on Automatic 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 Transactions on Automatic Control* **26**:2, pp. 301–320. DOI: 10.1109/TAC.1981.1102603.

Zhou, K. and J. Doyle (1998). *Essentials of Robust Control*. Essentials of Robust Control. Prentice Hall. ISBN: 9780137908745. URL: https://books.google.se/books?id=GdHEAAAACAAJ.

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

$$\begin{aligned} x_{t+1} &= Fx_t + Gu_t + w_t \\ y_t &= Hx_t + v_t, \qquad 0 \le t \le N-1, \end{aligned} \tag{1}$$

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, $\hat{x}_0$,

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

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, 1995]; 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 $A \in \mathbb{R}^{n\times n}$,

$|x|_A^2 := x^\top Ax$. For a vector $x_t \in \mathbb{R}^n$ we denote the sequence of such vectors up to time t by $\mathbf{x}^t := (x_k)_{k=0}^t$.

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, 1995], we do not make explicit assumptions on the distribution of the noise trajectories $\mathbf{w}^t$ and $\mathbf{v}^t$. 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 $(\mathbf{y}^{N-1}, \mathbf{u}^{N-1})$. Since the disturbances are unknown, we cannot evaluate (2) directly. However, define

$$\begin{aligned} J_N(\mathbf{y}^{N-1}, \mathbf{u}^{N-1}, \hat{y}_N) := \sup_{x_0, \mathbf{w}^{N-1}, \mathbf{v}^{N-1}, (F,G,H)} \Bigg\{ & |\hat{y}_N - Hx_N|^2 \\ & - \gamma_N^2 \left(|x_0 - \hat{x}_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1} \left(|w_t|^2_{Q^{-1}} + |v_t|^2_{R^{-1}} \right) \right) \Bigg\}, \end{aligned} \tag{3}$$

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

$$J_N(y^{N-1}, u^{N-1}, \hat{y}_N) \leq 0.$$

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

$$\begin{aligned} J_N(\mathbf{y}^{N-1}, \mathbf{u}^{N-1}, \hat{y}_N) = \sup_{\mathbf{x}^N, (F,G,H)} \Bigg\{ & |\hat{y}_N - Hx_N|^2 - \gamma_N^2 |x_0 - \hat{x}_0|^2_{P_0^{-1}} \\ & - \gamma_N^2 \sum_{t=0}^{N-1} \left(|x_{t+1} - Fx_t - Gu_t|^2_{Q^{-1}} + |y_t - Hx_t|^2_{R^{-1}} \right) \Bigg\}. \end{aligned} \tag{4}$$

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

$$\inf_{\hat{y}_N} J_N(y^{N-1}, u^{N-1}, \hat{y}_N) \;=\; J_N(y^{N-1}, u^{N-1}, \hat{y}_N^\star) \;=:\; J_N^\star(y^{N-1}, u^{N-1}), \tag{5}$$

holds, where $\hat{y}_N$ are functions of past data $\mathbf{y}^{N-1}$ and u^{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

$$J_N(\mathbf{y}^{N-1}, u^{N-1}, \hat{y}_N) = \max_i \left\{ |\hat{y}_N - H_i\check{x}_{N,i}|^2_{(I-\gamma_N^{-2}H_iP_{N,i}H_i^\top)^{-1}} - \gamma_N^2 c_{N,i} \right\}. \quad (6)$$

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

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

□

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

$$\sup_{x_N, i} \left\{ |\hat{y}_N - H_ix_N|^2 - \gamma_N^2 \inf_{\mathbf{x}^{N-1}} \left\{ |x_0 - \hat{x}_0|^2_{P_0^{-1}} + \sum_{t=0}^{N-1} \left(|x_{t+1} - F_ix_t - G_iu_t|^2_{Q^{-1}} + |y_t - H_ix_t|^2_{R^{-1}} \right) \right\} \right\},$$

[1] $\max_{\mathbf{x}^N}\{\dots\} = \max_{x_N}\left\{\max_{\mathbf{x}^{N-1}}\{\dots\}\right\}$.

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

$$J_N(y^{N-1}, u^{N-1}, \hat{y}_N) = \sup_{x_N, i} \left\{ |\hat{y}_N - H_i x_N|^2 - \gamma_N^2 V_{N,i}((x_N, y^{N-1}) \right\}$$
$$= \sup_{i, x_N} \left\{ |\hat{y}_N - H_i x_N|^2 - \gamma_N^2 \left(|x_N - \check{x}_N|^2_{P_{N,i}^{-1}} + c_{N,i} \right) \right\}.$$

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],

$$J_N(y^{N-1}, u^{N-1}, \hat{y}_N) = \max_i |\hat{y}_N - H_i \check{x}_{N,i}|^2_{(I - \gamma_N^{-2} H_i P_{N,i} H_i^\top)^{-1}} - \gamma_N^2 c_{N,i}. \qquad \square$$

Remark 2
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. $\square$

Remark 3
Equation (6) is monotonically increasing in γ_N and the smallest $\gamma_N^\star$ such that $J_N(y^{N-1}, u^{N-1}, \hat{y}_N) \le 0$ can be found efficiently through bisection. $\square$

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

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

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

Remark 4
If the model set is a singleton, then $\hat{y}_N^\star = H x_N^\star = H \check{x}_N$ is the estimate generated by the Kalman filter, which is a well known result [Basar and Bernhard, 1995]. $\square$

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

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 $\mathbf{y}^N$ has been generated by the model (F_i, G_i, H_i) and the input $\mathbf{u}^N$ can be expressed as

$$p(i|\mathbf{y}^N, \mathbf{u}^N) = \frac{\alpha_N e^{-|y_N - H_i \check{x}_{N,i}|^2_{\tilde{R}_{N,i}}}}{\det(2\pi \tilde{R}_{N,i})^{1/2}} p(i|\mathbf{y}^{N-1}, \mathbf{u}^{N-1}).$$

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 $e^{-c_{N+1,i}}$,

$$p(i|\mathbf{y}^{N-1}, \mathbf{u}^{N-1}) \propto e^{-c_{N+1,i}} \prod_{t=1}^{N} \det(2\pi \tilde{R}_{t,i})^{-1/2}.$$

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 $F_1, \dots, F_K \in \mathbb{R}^{n \times n}$, $H_1, \dots, H_K \in \mathbb{R}^{m \times n}$ 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

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

and (1), is the minimizing argument of

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

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

$$\begin{aligned}
\check{x}_{0,i} &= x_0 \\
\check{x}_{t+1,i} &= F_i\check{x}_{t,i} + K_i(y_t - H_i\check{x}_{t,i}) + G_iu_t \\
K_i &= F_iP_iH_i^\top(R + H_iP_iH_i^\top)^{-1} \\
c_{0,i} &= 0 \\
c_{t+1,i} &= |H_i\check{x}_{t,i} - y_t|^2_{(R+H_iP_iH_i^\top)^{-1}} + c_{t,i}.
\end{aligned} \qquad \square$$

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

5. Example

In this example, we compare a minimax estimator synthesized using Corollary 1, bisecting over γ_N, to find the estimator $\hat{y}^\star_N$ 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

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

The weights in (2) are chosen to be $Q = R = P_0 = 1$. We generate data $\mathbf{y}^{N-1}$ 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

$$P_{5,1} = P_{5,-1} = 1.62,$$

$$\check{x}_{5,1} = -2.34, \quad \check{x}_{5,-1} = 1.50,$$

$$c_{5,1} = 3.56, \quad c_{5,-1} = 8.11.$$

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 $N = 1 \dots 20$.

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

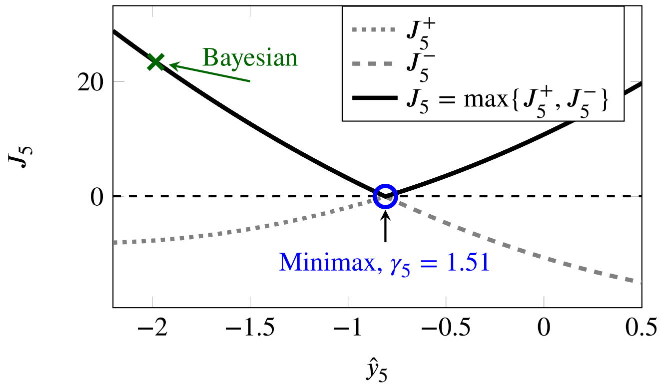


Figure 1. Illustration of the optimization problem (6) for $N = 5$, together with the minimax solution and the one given by a Bayesian multiple model estimator for $\gamma_N = 1.51$. The minimax estimate has a guaranteed worst-case gain bound from disturbances to observer error lower than 1.51, whereas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

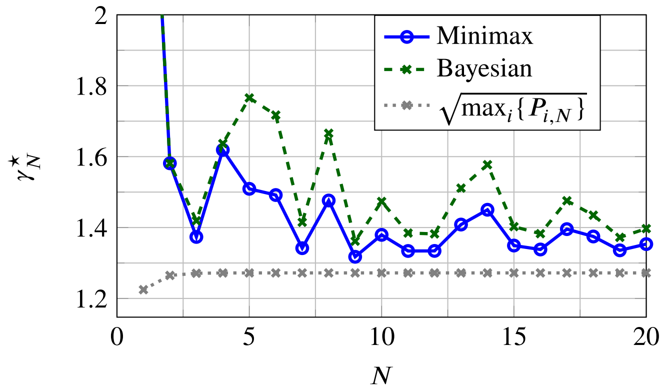


Figure 2. The smallest γ_N such that $J_N(\mathbf{y}^{N-1}, 0, \hat{y}_N) \leq 0$ for the minimax estimator (blue) compared to the Bayesian multiple-model adaptive estimator (green) for one realization.

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

$$\sup_{\mathbf{y}^{N-1}} J_N^\star(\mathbf{y}^{N-1}, \mathbf{u}^{N-1}) \leq 0.$$

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.

Appendix — Supporting Lemmata

LEMMA 1
The cost function

$$\begin{aligned} V_{N,i}(x_N, \mathbf{y}^{N-1}) = \min_{\mathbf{x}^{N-1}} \Bigg\{ & |x_0 - \hat{x}_0|^2_{P_0^{-1}} \\ & + \sum_{k=1}^{N-1} (|x_{t+1} - F_i x_t - G_i u_t|^2_{Q^{-1}} + |y_t - H_i x_t|^2_{R^{-1}}) \Bigg\} \end{aligned} \tag{7}$$

under the dynamics (1), is of the form

$$V_{t,i}(x, \mathbf{y}^{t-1}) = |x - \check{x}_{t,i}|^2_{P_{t,i}} + c_{t,i},$$

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

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

□

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

$$V_1(x, \mathbf{y}^0) = |x - x_0|^2_{P_0^{-1}} \tag{8}$$

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

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

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

[3] We relax the index i in this proof

Since this should hold for arbitrary x and

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

we get

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

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

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

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

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

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

$$\check{x}_{t+1} = F \check{x}_t + K_t (y_t - H \check{x}) + G u_t,$$

where

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

□

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

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

□

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

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

□

References

Ackerson, G. and K. Fu (1970). "On state estimation in switching environments". *IEEE Transactions on Automatic 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". *IFAC-PapersOnLine* **52**:3. 15th IFAC Symposium on Large Scale Complex Systems LSS 2019, pp. 73–78. ISSN: 2405-8963. DOI: https://doi.org/10.1016/j.ifacol.2019.06.013. URL: https://www.sciencedirect.com/science/article/pii/S2405896319300977.

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

Basar, T. and P. Bernhard (1995). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach*. Birkhauser.

Blom, H. A. P. and Y. Bar-Shalom (1988). "The interacting multiple model algorithm for systems with markovian switching coefficients". *IEEE Transactions on Automatic 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 Transactions on Aerospace and Electronic Systems* **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 Transactions on Automatic Control* **9**:1, pp. 5–12.

Crassidis, J. L. and J. L. Junkins (2011). *Optimal Estimation of Dynamic Systems, Second Edition (Chapman & Hall/CRC Applied Mathematics & Nonlinear Science)*. 2nd. Chapman & Hall/CRC. ISBN: 1439839859.

Deng, L., D. Li, and R. Li (2020). "Improved IMM algorithm based on RNNs". *Journal of Physics: Conference Series* **1518**, p. 012055. DOI: 10.1088/1742-6596/1518/1/012055. URL: https://doi.org/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.

Goodwin, G. C. and K. S. Sin (1984). *Adaptive filtering prediction and control / Graham C. Goodwin and Kwai Sang Sin.* eng. Prentice-Hall information and system sciences series. Prentice-Hall, Englewood Cliffs, N.J. ISBN: 013004069X.

Lainiotis, D. G. (1976). "Partitioning: a unifying framework for adaptive systems, i: estimation". *Proceedings of the 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". *Journal of Physics: Conference Series* **1748**, p. 032017. DOI: 10.1088/1742-6596/1748/3/032017. URL: https://doi.org/10.1088/1742-6596/1748/3/032017.

Magill, D. (1965). "Optimal adaptive estimation of sampled stochastic processes". *IEEE Transactions on Automatic 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*. arXiv: 2011.10814 [math.OC].

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

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

Xiong, K., C. Wei, and L. Liu (2015). “Robust multiple model adaptive estimation for spacecraft autonomous navigation”. *Aerospace Science and Technology* **42**, pp. 249–258. ISSN: 1270-9638. DOI: https://doi.org/10.1016/j.ast.2015.01.021. URL: https://www.sciencedirect.com/science/article/pii/S1270963815000371.

Paper II

Fundamental Worst-Case Performance Limits for Multiple-Model Estimation

Olle Kjellqvist

Abstract

This article provides upper- and lower bounds for the worst-case performance limit of strictly causal state estimation for uncertain linear systems with uncertainty belonging to a finite set. We quantify performance as the gain from noise and disturbances to point-wise estimation error. The bounds rely on forward Riccati recursions, one per feasible model, and backward Riccati recursions, one for each pair of models in the uncertain set.

1. Introduction

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

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

where $x_t \in \mathbb{R}^n$, and $y_t \in \mathbb{R}^m$ are the states and the measured 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, $(F, H) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{m\times n}$ is unknown but fixed, belonging to a (known) finite set

$$\mathcal{M} = \{(F_1, H_1), \dots, (F_K, H_K)\}.$$

The state estimate, $\hat{x}_N$, is generated by a strictly causal map from measured outputs to $\mathbb{R}^n$ that is agnostic of the model realization.

$$\hat{x}_N = \mu(y_{N-1}, \dots, y_0). \tag{2}$$

We make no assumptions about the distributions of w and v. Instead, we employ a deterministic framework and characterize the achievable worst-case performance. The worst-case performance is quantified as the smallest gain from disturbance trajectories $(w_t, v_t)_{t=0}^{N-1}$ to pointwise estimation error $(x_N - \hat{x}_N)$, denoted $\gamma_N^\star$, that can be achieved with a strictly causal estimator μ. The disturbance norm can be different for every model, i. The norms are defined by the positive definite matrices $P_{0,i} \in \mathbb{R}^{n\times n}$, $R_i \in \mathbb{R}^{m\times m}$ and $Q_i \in \mathbb{R}^{n\times n}$ and nominal values of the initial state, $\hat{x}_{0,i}$. More precisely, $\gamma_N^\star$ is defined as the smallest γ_N for which there exists an estimator, μ, such that for all possible realizations of (1) and (2),

$$\frac{|\hat{x}_N - x_N|^2}{|x_0 - \hat{x}_{0,i}|^2_{P_{0,i}^{-1}} + \sum_{t=0}^{N-1}\left(|w_t|^2_{Q_i^{-1}} + |v_t|^2_{R_i^{-1}}\right)} \leq \gamma_N^2. \tag{3}$$

We provide a lower and an upper bound for $\gamma_N^\star$. Both bounds depend on the pairwise interaction between candidate models and are computed by validating the positive definiteness of certain matrices appearing in forward and backward Riccati recursions. The Riccati recursions describe the worst-case trajectory given a minimax-optimal state estimate. This estimator was described in [Kjellqvist and Rantzer, 2022b].

1.1 Background

[Bar-Shalom, 1972] studied simultaneous estimation of parameters and states in linear systems and showed that the maximum-likelihood approach leads to estimates that cannot be put on recursive form. One suboptimal approach is to parametrize the dynamics, augment the state vector with the parameters and apply nonlinear estimation techniques like the extended Kalman Filter. The state and parameter estimates may diverge [Ljung, 1979], but the algorithm can be tweaked to guarantee asymptotic convergence. The resulting algorithm is equivalent to a recursive prediction error method [Goodwin and Sin, 1984]. Unfortunately, optimality guarantees for these methods are difficult to obtain. However, if we restrict the parameter uncertainty to belong to a finite set, and the driving noise is Gaussian. Then the conditional probability of each model having generated the observed trajectory can be expressed recursively. The conditional probability can then be used to extract the maximum likelihood estimate, the expected value of the state and other meaningful statistical quantities.

This *multiple-model approach* was introduced in the '60s by [Magill, 1965] and [Lainiotis, 1976] and has then been adapted to switching systems [Ackerson and Fu, 1970; Chang and Athans, 1978; Blom and Bar-Shalom, 1988]. The extensions for switching systems approximate the solution by combining the estimates and model pruning. The approximation is necessary to avoid the exponential growth of the feasible parameter trajectories due to switching. [Ronghua et al., 2008] shows that multiple-model estimators are sensitive to deviations from the assumed 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. [Li et al., 2021; Deng et al., 2020] combined multiple-model estimators with machine-learning approaches to classification and showed improved accuracy in simulations.

Recently, [Silvestre et al., 2020] introduced the notion of absolute distinguishability to evaluate the suitability of sets in multiple-model systems. Distinguishability means that there exists some input sequence such that the different systems generate different outputs. If the outputs are different for *all* system inputs, then the systems are absolutely distinguishable.

The work in this article has been inspired by recent progress in *minimax adaptive control* [Rantzer, 2020; Rantzer, 2021; Kjellqvist and Rantzer, 2022a; Renganathan et al., 2022] and builds on the optimal minimax adaptive estimator for finite sets of linear systems introduced in [Kjellqvist and Rantzer, 2022b].

1.2 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 contained in Section 4. We give conclusions and final remarks in Section 5. The proof of the main results are contained in Section 6

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 $A \in \mathbb{R}^{n\times n}$, $|x|^2_A := x^\top A x$. 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 performance limits

This section studies the minimax performance index

$$J_N^\star(\hat{x}_0) := \inf_{\mu} \sup_{x_0, w_{[0:T-1]}, v_{[0:T-1]}, i} \Bigg\{ |x_N - \hat{x}_N|^2 - \gamma^2 \Bigg(|x_0 - \hat{x}_{0,i}|^2_{P_{0,i}} + \sum_{t=0}^{N-1} \Big[|w_t|^2_{Q_i^{-1}} + |v_t|^2_{R_i^{-1}} \Big] \Bigg) \Bigg\}, \quad (4)$$

where $x_{[0:T]}$ is generated according to (1) with $(F_i, H_i) \in \mathcal{M}$. This is a two-player game where the adversary picks the disturbance sequences $w_{[0:T-1]}$ and $v_{[0:T-1]}$

and the active model $i = 1, \dots, K$ and the minimizing player picks the estimation policy μ. In particular, we are interested in the infimum of γ_N, denoted $\gamma_N^\star$, such that the performance index (4) has a finite value.

The case when the model is known, equivalently when $\mathcal{M} = (F_1, H_1)$, was studied in [Basar and Bernhard, 1995]. In that case, the smallest achievable attenuation level $\gamma_N^\star$ can be characterized using two Riccati recursions, one forward recursion and one backward recursion. The Riccati recursions are functions of γ_N, and one can evaluate whether γ_N is greater or smaller than $\gamma_N^\star$ by checking whether certain matrix expressions occuring in the backward recursion are positive definite. These criteria are both necessary and sufficient, when the model is fully known.

The remainder of this article is devoted upper- and lower bounds of $\gamma_N^\star$ for uncertain linear systems, where the uncertainty is described by a finite set. These bounds are characterized using forward and backward Riccati recursions. There is one forward recursions per model, $i = 1, \dots, K$, and one backward recursions for each *pair* of models (i, j).

The forward recursions are those of a Kalman filter with zero-mean independent white noise sequences w_t and v_t with covariance matrices Q_i and R_i respectively,

$$\begin{aligned} K_{t,i} &= F_i P_{t,i} H_i^\top (R_i + H_i P_{t,i} H_i^\top)^{-1}, \\ P_{t+1,i} &= Q_i + F_i P_{t,i} F_i + K_{t,i}(R_i + H_i P_{t,i} H_i^\top) K_{t,i}^\top. \end{aligned} \tag{5}$$

$P_{0,i}$ can be interpreted as the covariance of the initial estimate $\hat{x}_{0,i}$. The output-error covariance matrix, $\tilde{R}_{t,i}$, will play an important role in the backward recursion and is given by

$$\tilde{R}_{t,i} = R_i + H_i P_{t,i} H_i^\top. \tag{6}$$

We now state our first condition on γ_N.

PROPOSITION 1
$\gamma_N \geq \gamma_N^\star$ only if $P_{N,i} \preceq \gamma_N^2 I$ for all $i = 1, \dots, K$. □

Let

$$F_t^{ij} = \begin{bmatrix} F_i - K_{t,i} H_i & \\ & F_j - K_{t,j} H_j \end{bmatrix}, \quad K_t^{ij} = \begin{bmatrix} K_{t,i} \\ K_{t,j} \end{bmatrix}. \tag{7}$$

F_t^{ij} corresponds to the closed-loop of a pair (i, j) of Kalman filters with filter gains $K_{t,i}$ and $K_{t,j}$. We will express the necessary and sufficient conditions using the following Riccati recursions. Given some symmetric matrix $T_N \in \mathbb{R}^{2n \times 2n}$,

$$X_t^{ij} = \left((K_t^{ij})^\top T_{t+1}^{ij} K_t^{ij} - \gamma^2 (\tilde{R}_{t,i}^{-1} + \tilde{R}_{t,j}^{-1}) \right), \tag{8a}$$

$$L_t^{ij} = (X_t^{ij})^{-1} \left((K_t^{ij})^\top T_{t+1}^{ij} F_t^{ij} + \gamma^2 \begin{bmatrix} \tilde{R}_{t,i}^{-1} H_i & \tilde{R}_{t,j}^{-1} H_j \end{bmatrix} \right), \tag{8b}$$

$$T_t^{ij} = (F_t^{ij})^\top T_{t+1}^{ij} F_t^{ij} - \gamma^2 \begin{bmatrix} H_i^\top \tilde{R}_{t,i}^{-1} H_i & \\ & H_j^\top \tilde{R}_{t,j}^{-1} H_j \end{bmatrix} - (L_t^{ij})^\top X_t^{ij} L_t^{ij}. \tag{8c}$$

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 stated in Theorems 3 and 4 below.

THEOREM 3—UPPER BOUND
Given γ_N such that $P_N \prec \gamma_N^2 I$. Let $S \in \mathbb{R}^{n\times n}$ be a positive definite matrix such that $S \preceq I - \gamma^{-2} P_{N,i}$ for all $i = 1, \dots, K$ and

$$T_N^{ij} = \begin{bmatrix} S^{-1} & -S^{-1} \\ -S^{-1} & S^{-1} \end{bmatrix}.$$

Assume that for all i, j that X_t^{ij} in (8a) is negative definite. Then $\gamma_N^\star \le \gamma_N$ and

$$J_N^\star(\hat{x}_0) \le \frac{1}{2} \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix}^\top T_0^{ij} \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix}.$$

□

THEOREM 4—LOWER BOUND
Given γ_N such that $P_N \prec \gamma_N^2 I$, if $X_t^{ij} \npreceq 0$ for some pair i, j and $0 \le t \le N-1$, where X_t^{ij} is generated according to the Riccati recursion (8a)–(8c), with

$$T_N^{ij} = \begin{bmatrix} T^{ij} & -T^{ij} \\ -T^{ij} & T^{ij} \end{bmatrix}, \quad T^{ij} = (2I - \gamma^{-2}(P_{N,i} + P_{N,j}))^{-1}.$$

then $\gamma_N^\star > \gamma_N$. If $X_t^{ij} \succ 0$, for all $t = 0, \dots, N-1$ then

$$J_N^\star(\hat{x}_0) \ge \frac{1}{2} \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix}^\top T_0^{ij} \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix}.$$

□

We conclude this section with a causal estimator that achieves (4).

PROPOSITION 2—[KJELLQVIST AND RANTZER, 2022B]
Given $\gamma_N \ge \gamma_N^\star$, then the infimum in (4) is achieved by

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

With $P_{t,i}$ and $K_{t,i}$ as in (5), $\check{x}_{N,i}$ and $c_{N,i}$ are generated according to

$$\check{x}_{0,i} = x_0, \quad c_{0,i} = 0, \tag{9a}$$

$$\check{x}_{t+1,i} = F_i \check{x}_{t,i} + K_{t,i}(y_t - H_i \check{x}_{t,i}), \tag{9b}$$

$$c_{t+1,i} = |H_i \check{x}_{t,i} - y_t|^2_{\tilde{R}_{t,i}^{-1}} + c_{t,i}. \tag{9c}$$

□

4. Examples

Figures 1–4 show upper and lower bounds for $\gamma_N^\star$ for four different scalar systems computed using Theorem 3 and 4. In all cases $Q_1 = Q_2 = R_1 = R_2 = 1$ and $P_{0,i}$ were taken as the solution to the *Riccati equation* resulting from removing the time index in (5). The systems in Fig. 1 are unstable and *indistinguishable* and the resulting attenuation level $\gamma_N^\star$ grows exponentially in N. This is because very different state trajectories result in the same output trajectory. Fig. 2 is also *indistinguishable*, but here both systems are stable. The minimax attenuation level $\gamma_N^\star$ is bounded and lies somewhere between 2 to 2.5 times that of the case when the model is known. This is because the systems are BIBO stable, so picking $\hat{x}_N = 0$ results in an estimation error that is bounded by the disturbance norm. Fig. 3 contains two stable systems that are *indistinguishable*, but here the minimax attenuation level is not much worse than in the case where the model is known. This is because even though the systems are different, and cannot be distinguished from data, state trajectories resulting in the same output trajectory cannot differ very much. Fig. 4 contains two unstable *distinguishable* systems. We see that the minimax attenuation level is a bounded function of N. Comparing Fig. 4 to Figs. 2 and 3, we see that distinguishability alone is not enough to guarantee good performance of multiple-model estimators.

5. Conclusions

This article provided an upper and a lower bound for the minimax attenuation level for uncertain linear systems, where the uncertainty belongs to a finite set. The attenuation level provides an alternative to the notion of *distinguishability* in a priori analysis of the problem set-up for multiple-model estimation. In the case when similar output trajectories can come from highly different state trajectories, the gain seems to be large. If similar output trajectories must come from similar state trajectories, the gain seems to be small. Interesting directions for future research are:

- H_∞-(sub)optimal multiple-model filtering
- Tighter bounds on the minimax gain from disturbances to estimation error.
- Switched linear systems.
- Uncertain linear systems, with the uncertainty belonging to a compact (but infinite) set.
- Experimental studies.

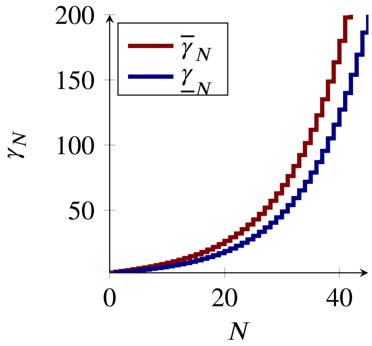


Figure 1. Bounds on the minimax gain from disturbances to estimation error for two systems where $F1 = F2 = 1.1$, $H_1 = -H_2 = 1$, $P_1 = P_2 = 1.77$.

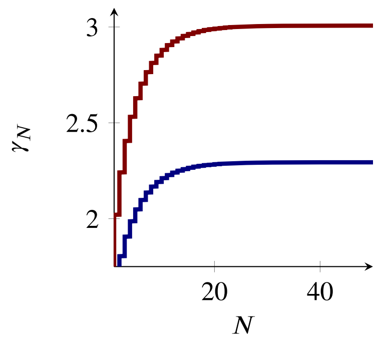


Figure 2. Bounds on the minimax gain from disturbances to estimation error for two systems where $F_1 = F_2 = 0.9$, $H_1 = -H_2 = 1$, $P_1 = P_2 = 1.48$.

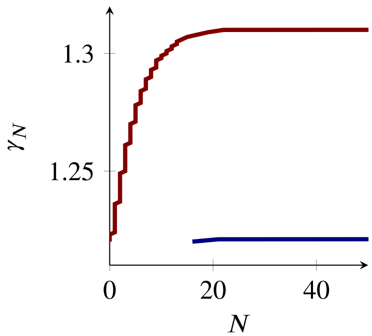


Figure 3. Bounds on the minimax gain from disturbances to estimation error for two systems where $F_1 = F_2 = 0.9$, $H_1 = 1$, $H_2 = 1.5$, $P_1 = 1.48$, $P_2 = 1.27$.

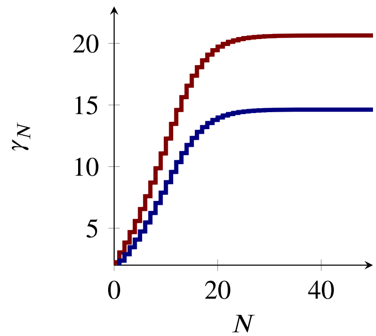


Figure 4. Bounds on the minimax gain from disturbances to estimation error for two systems where $F_1 = 1.1$, $F_2 = 1.2$, $H_1 = -H_2 = 1$, $P_1 = 1.77$, $P_2 = 1.95$.

6. Proofs

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. Thus, we can substitute $w_t = x_{t+1} - F_i x_t$ and $v_t = y_t - H_i x_t$ into (4),

$$J_N^\star(\hat{x}_0) = \inf_{\mu} \sup_{x_{[0:N]}, y_{[0:N-1]}, i} \Bigg\{ |x_N - \hat{x}_N|^2 - \gamma^2 |x_0 - \hat{x}_{0,i}|^2_{P_{0,i}} - \gamma^2 \sum_{t=0}^{N-1} \left[|x_{t+1} - F_i x_t|^2_{Q_i^{-1}} + |y_t - H_i x_t|^2_{R_i^{-1}} \right] \Bigg\}. \tag{10}$$

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 vector $\hat{x}_0 \in \mathbb{R}^n$. Let,

$$J_N^{\text{inner}}(y_{[0,N-1]}, \hat{x}_N, \hat{x}_0) = \sup_{x_{[0:N]}, i} \Bigg\{ |x_N - \hat{x}_N|^2 - \gamma_N^2 |x_0 - \hat{x}_{0,i}|^2_{P_{0,i}} - \gamma^2 \sum_{t=0}^{N-1} \left[|x_{t+1} - F_i x_t|^2_{Q_i^{-1}} + |y_t - H_i x_t|^2_{R_i^{-1}} \right] \Bigg\}. \tag{11}$$

Then (10) can be written as

$$J_N^\star(\hat{x}_0) = \sup_{y_{[0:N-1]}} \inf_{\hat{x}_N} J_N^{\text{inner}}(y_{[0,N-1]}, \hat{x}_N, \hat{x}_0). \tag{12}$$

The following proposition shows how to compute the inner optimization problem (11) using K parallel Kalman filters.

PROPOSITION 3
Consider matrices $F_i \in \mathbb{R}^{n\times n}$ and $H_i \in \mathbb{R}^{n\times n}$, and positive definite $Q_i, P_{0,i} \in \mathbb{R}^{n\times n}$ and $R_i \in \mathbb{R}^{m\times m}$ for $i = 1, \dots, K$ and a positive number γ_N. Let $P_{t,i}$ be the solution to the Riccati recursion (5), If $P_{N,i} \prec \gamma_N^2 I$, then the inner optimization problem (11) is equivalent to

$$J_N^{\text{inner}}(y_{[0,N-1]}, \hat{x}_N, \hat{x}_0) = \max_i \left\{ |\hat{x}_N - \check{x}_{N,i}|^2_{(I-\gamma^{-2}P_{N,i})^{-1}} - \gamma^2 c_{N,i} \right\}, \tag{13}$$

where $\check{x}_{N,i}$ and $c_{N,i}$ are generated according to (9). If $P_{N,i} \nprec \gamma_N^2 I$ for some i, then the value (11) is unbounded. □

Proof. See [Kjellqvist and Rantzer, 2022b]. □

The last statement of Proposition 3, is precisely Proposition 1.

6.1 Upper- and Lower bounds of $J_N^\star$

In this section we develop upper and lower bounds on the objective, (4), that we translate into bounds on $\gamma_N^\star$ via bisection. Substitute (13) into (12) to get

$$J_N^\star(\hat{x}_0) = \sup_{y_{[0:N-1]}} \min_{\hat{x}_N} \max_i \left\{ |\hat{x}_N - \check{x}_{N,i}|^2_{(I-\gamma^{-2}P_{N,i})^{-1}} - \gamma^2 c_{N,i} \right\}. \tag{14}$$

As the maximum is greater than the average of any two points, we have that

$$\begin{aligned} J_N^\star(\hat{x}_0) \geq & \sup_{i,j,y_{[0:N-1]}} \min_{\hat{x}_N} \frac{1}{2}\Big\{ |\hat{x}_N - \check{x}_{N,i}|^2_{(I-\gamma^{-2}P_{N,i})^{-1}} \\ & - \gamma^2 c_{N,i} + |\hat{x}_N - \check{x}_{N,j}|^2_{(I-\gamma^{-2}P_{N,j})^{-1}} - \gamma^2 c_{N,j} \Big\} \\ = & \sup_{i,j,y_{[0:N-1]}} \frac{1}{2}\Big\{ |\check{x}_{N,i} - \check{x}_{N,j}|^2_{(2I-\gamma^{-2}P_{N,i}-\gamma^{-2}P_{N,j})^{-1}} \\ & - \gamma^2 c_{N,i} - \gamma^2 c_{N,j} \Big\} =: \max_{i,j} \underline{J}_N^{ij}(\hat{x}_0). \end{aligned} \tag{15}$$

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 $S \preceq I - \gamma^{-2}P_{t,i}$ for all $i = 1, \ldots, K$. Then

$$\begin{aligned} J_N^\star(\hat{x}_0) &\leq \max_y \min_{\hat{x}_N} \max_i \{ |\hat{x}_N - \check{x}_{N,i}|^2_{S^{-1}} - \gamma^2 c_{N,i} \} \\ &= \max_y \max_\theta \left\{ \sum_{i,j}^{M} \theta_i\theta_j |\check{x}_{N,i} - \check{x}_{N,j}|^2_{S^{-1}}/2 - \gamma^2 \sum_i \theta_i c_{N,i} \right\} \\ &\leq \max_{i,j} \max_y \frac{1}{2} \left\{ |\check{x}_{N,i} - \check{x}_{N,j}|^2_{S^{-1}} - \gamma^2 (c_{N,i} + c_{N,j}) \right\} \\ &=: \max_{i,j} \overline{J}_N^{ij}(\hat{x}_0). \end{aligned} \tag{16}$$

Similarly, 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 weighting matrix of the term $|\check{x}_{N,i} - \check{x}_{N,j}|$. Computing the upper and the lower bound amounts to solving LQR problems with an indefinite terminal penalty on the state.

6.2 Indefinite LQR

Define the cost function

$$J_N^{ij}(T_N) = \sup_{y_{[0:N-1]}} \left\{ \begin{bmatrix} \check{x}_{N,i} \\ \check{x}_{N,j} \end{bmatrix}^\top T_N \begin{bmatrix} \check{x}_{N,i} \\ \check{x}_{N,j} \end{bmatrix} - \gamma^2 \sum_{t=0}^{N-1} |H_i \check{x}_{t,i} - y_t|^2_{\tilde{R}_{t,i}^{-1}} - \gamma^2 \sum_{t=0}^{N-1} |H_j \check{x}_{t,j} - y_t|^2_{\tilde{R}_{t,j}^{-1}} \right\} \quad (17)$$

The cost (17) is a standard finite-horizon time-varying linear-quadratic regulator problem and readily obtained by dynamic programming. We summarize the solution in the following lemma.

LEMMA 3
Fix $i, j \in \{1, \dots, K\}$. Let $T_N^{ij} \in \mathbb{R}^{2n \times 2n}$ be a symmetric matrix. Define T_t^{ij}, X_t^{ij} and L_t^{ij} for $t = 0, \dots, N-1$ by the Riccati recursions (8). Then $J_N^{ij}(T_N, \hat{x}_0) = +\infty$ if $X_t \npreceq 0$ for some t. If $X_t \prec 0$ for all $t = 0, \dots, N-1$, then

$$J_N^{ij}(T_N, \hat{x}_0) = \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix}^\top T_0^{ij} \begin{bmatrix} \hat{x}_{0,i} \\ \hat{x}_{0,j} \end{bmatrix}. \qquad \square$$

REMARK 5
The output sequence $y_t = -L_t^{ij} \begin{bmatrix} \check{x}_{t,i} \\ \check{x}_{t,j} \end{bmatrix}$ maximizes (17). $\square$

Theorems 3 and 4 follow from applying Lemma 3 to the upper bound $\overline{J}_N^{ij}(\hat{x}_0)$ in (16) and the lower bound $\underline{J}_N^{ij}(\hat{x}_0)$ in (15).

Acknowledgements

This author is thankful to Venkatraman Renganathan and Anders Rantzer for helpful discussions on previous versions of the manuscript.

References

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

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

Basar, T. and P. Bernhard (1995). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach*. Birkhauser.

Blom, H. A. P. and Y. Bar-Shalom (1988). "The interacting multiple model algorithm for systems with markovian switching coefficients". *IEEE Transactions on Automatic 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 Transactions on Aerospace and Electronic Systems* **AES-14**:3, pp. 418–425. DOI: 10.1109/TAES.1978.308603.

Deng, L., D. Li, and R. Li (2020). "Improved IMM algorithm based on RNNs". *Journal of Physics: Conference Series* **1518**, p. 012055. DOI: 10.1088/1742-6596/1518/1/012055. URL: https://doi.org/10.1088/1742-6596/1518/1/012055.

Goodwin, G. C. and K. S. Sin (1984). *Adaptive filtering prediction and control / Graham C. Goodwin and Kwai Sang Sin.* eng. Prentice-Hall information and system sciences series. Prentice-Hall, Englewood Cliffs, N.J. ISBN: 013004069X.

Kjellqvist, O. and A. Rantzer (2022a). "Learning-enabled robust control with noisy measurements". In: *Learning for Dynamics and Control Conference*. PMLR, pp. 86–96.

Kjellqvist, O. and A. Rantzer (2022b). "Minimax adaptive estimation for finite sets of linear systems". In: *2022 American Control Conference (ACC)*, pp. 260–265. DOI: 10.23919/ACC53348.2022.9867474.

Lainiotis, D. G. (1976). "Partitioning: a unifying framework for adaptive systems, i: estimation". *Proceedings of the 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". *Journal of Physics: Conference Series* **1748**, p. 032017. DOI: 10.1088/1742-6596/1748/3/032017. URL: https://doi.org/10.1088/1742-6596/1748/3/032017.

Ljung, L. (1979). "Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems". *IEEE Transactions on Automatic Control* **24**:1, pp. 36–50. DOI: 10.1109/TAC.1979.1101943.

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

Rantzer, A. (2020). *Minimax adaptive control for state matrix with unknown sign*. arXiv: 1912.03550 [math.OC].

Rantzer, A. (2021). *Minimax adaptive control for a finite set of linear systems*. arXiv: 2011.10814 [math.OC].

Renganathan, V., A. Rantzer, and O. Kjellqvist (2022). "Distributed implementation of minimax adaptive controller for finite set of linear systems". *arXiv preprint arXiv:2210.00081*.

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

Silvestre, D., P. Rosa, and C. Silvestre (2020). "Distinguishability of discrete-time linear systems". *International Journal of Robust and Nonlinear Control* **31**. DOI: 10.1002/rnc.5367.

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 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

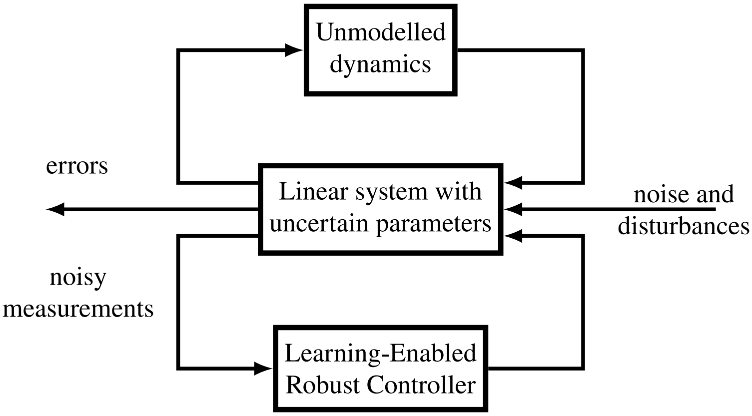


Figure 1. For a finite set of linear time-invariant models, the Learning-Enabled Robust Controller minimizes the ℓ_2-gain from noise and disturbances to errors for any realization of the unknown model parameters. This gain bound guarantees robustness to unmodelled dynamics.

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 5 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 6. 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,

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

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

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

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 6
The cases $b = -1$ and $b = 1$ cannot be simultaneously stabilized by a static feedback controller when $a \geq 1$ □

REMARK 7
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, 1995] to bound (3) in a way which leads itself to recursive computation. We need the following lemma:

LEMMA 4—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

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

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

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

REMARK 8
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, 1995]. 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 4. The system is equivalent to (6.1) and (6.2) in [Basar and Bernhard, 1995, 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

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

and conclude $l_M(t+1) = -m_k$. □

Lemma 4 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

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

Then $\max_M \alpha_M(t)$ is the largest possible value of (3) at time t. In the following theorem, we use Lemma 4 to express α_M recursively and construct equivalent conditions using computable quantities.

THEOREM 5—INFORMATION-STATE CONDITION
Given a causal output-feedback control policyμ, a positive quantity γ, and an un-

certainty set $\mathcal{M}$. Assume that for all $(a,b,c) = M \in \mathcal{M}$ the Riccati equation

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

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

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

$$l_M(t+1) = l_M(t) - P_M \hat{x}_M(t)^2 - \gamma^2 y(t)^2 + \frac{(P_M \hat{x}_M(t) + \gamma^2 c y(t))^2}{P_M + \gamma^2 c^2 - 1}, \quad l_M(0) = 0. \tag{11}$$

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 4 to express $\alpha_M(t)$ in the known quantities $\hat{x}_M(t)$, P_M and $l_M(t)$[1]:

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

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. □

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

From Theorem 5 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 6—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), (11) and an observer-based control policy $\eta^\star$

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

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):

$$\mu_t^\star\left(y(0), y(1) \ldots, y(t)\right) = \eta^\star \left\{ \left(\hat{x}_M(t), l_M(t+1), y(t) \right) \,:\, m \in \mathcal{M} \right\} \tag{12}$$

□

REMARK 9

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 6 *(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, $l_M(T) \le 0$ for all T, M and $y \in \ell_2[0,T]$ for the controller (12), which we know implies that the system is finite gain by Theorem 5.

(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 5 are fulfilled and there exist observers $\hat{x}_M$ and l_M generated by (10) and

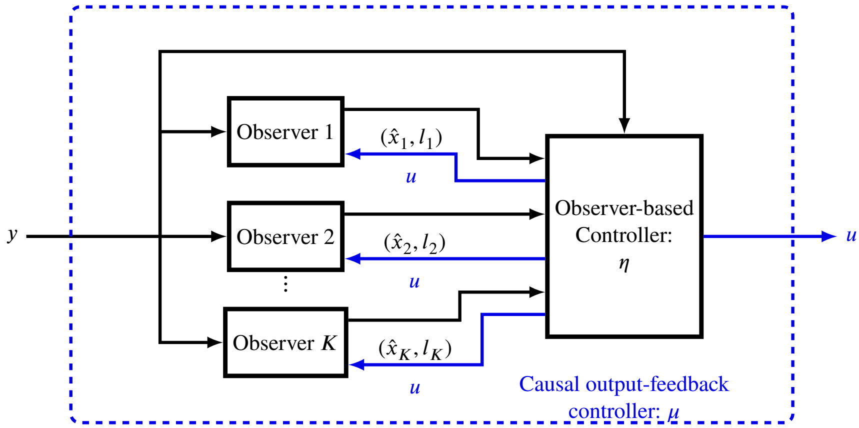


Figure 2. Illustration of the controller architecture in Theorem 6 for uncertainty sets consisting of K linear models. The controller η only considers the current state of the observers.

(11). Define the set of feasible generating trajectories given observer states $\hat{x}_M(t)$, $l(t)$ and current measurement $y(t)$:

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

Then $\mathcal{T}\left\{(\hat{x}(0), l_M(1), y(0)) \,:\, M \in \mathcal{M}\right\}$ is nonempty since it is compatible with any trajectory of length 1 such that $\check{y}(0) = y(0)$. Fix $t \geq 0$ and observer states $\hat{x}_M(t), l_M(t+1)$ and measurement $y(t)$. Assume that $\mathcal{T}\left\{\hat{x}_M(t), l_M(t), y(t)) \,:\, M \in \mathcal{M}\right\}$ is non empty. Then there exists a sequence $\check{y}$, and final time T so that $l_M(t+1) = \alpha_M(T)$ with $\alpha_M(t)$ as in (8) generated by $\check{y}$ and the controller $u(\tau) = \mu^\star(\check{y}(0), \ldots, \check{y}(\tau))$. By assumption, $l_M(t+1) = \alpha_M(T) \leq 0$. Taking

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

for some $\check{y}, \in \mathcal{T}\left\{(\hat{x}_M(t) l_M(t+1), y(t)) \,:\, M \in \mathcal{M}\right\}$ 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 6 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

$$P_i = P = \frac{1}{2}(1 - \gamma^2 a^2) + \sqrt{\gamma^2(-1 + \gamma^2) + (\gamma^2 a^2 - 1)^2/4}. \tag{13}$$

Construct the observers $\hat{x}_i$ and cumulative objectives l_i using (10) and (11) 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

$$u(t) = \begin{cases} -(\hat{a}\hat{x}_1(t) + \hat{g}y(t)) & \text{if } l_1(t+1) \geq l_{-1}(t+1) \\ \hat{a}\hat{x}_{-1}(t) + \hat{g}y(t) & \text{if } l_1(t+1) < l_{-1}(t+1). \end{cases} \tag{14}$$

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.

Proposition 4
With $\hat{a}$, $\hat{g}$, P as above, $\hat{x}_i$ and l_i as in (10) and (11), and the control signal given by (14), let

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

[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.

Then the following is true:

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

$$2: \begin{cases} l_1(t+1) & = \begin{cases} l_1(t) - \gamma^2 y(t)^2 + \frac{(\gamma^2 y(t))^2}{P+\gamma^2-1} & \text{if } l_1(t) \ge l_{-1}(t) \\ l_1(t) - P\hat{x}(t)^2 - \gamma^2 y(t)^2 + \frac{(P\hat{x}(t)+\gamma^2 y(t))^2}{P+\gamma^2-1}, & \text{if } l_1(t) < l_{-1}(t) \end{cases} \\ l_{-1}(t+1) & = \begin{cases} l_{-1}(t) - P\hat{x}(t)^2 - \gamma^2 y(t)^2 + \frac{(P\hat{x}(t)+\gamma^2 y(t))^2}{P+\gamma^2-1}, & \text{if } l_1(t) \ge l_{-1}(t) \\ l_{-1}(t) - \gamma^2 y(t)^2 + \frac{(\gamma^2 y(t))^2}{P+\gamma^2-1}, & \text{if } l_1(t) < l_{-1}(t) \end{cases} \end{cases} \tag{15}$$

□

Proof. We start by proving the first claim. Consider the case when $l_1(t+1) \ge l_{-1}(t+1)$. 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 $l_1(t) \ge l_{-1}(t)$ then $\hat{x}_1(t) = 0$ and $\hat{x}_{-1}(t) = \hat{x}(t)$. The claim follows by substitution into (11). □

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 5

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

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

Then with $l_i(t+1)$ as in (15), it holds that $l_i(t+1) \le 0$ for $i \in \{1, -1\}$. □

Proof Lemma 5, full. We will give the proof for the case $0 \ge l_1(t) \ge l_{-1}(t)$. The case $0 \ge l_{-1}(t) \ge l_1(t)$ is similar. Note that $l_1(t+1)$ and $l_{-1}(t+1)$ are concave in $y(t)$ if and only if

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

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

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

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 5 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 7—Certainty equivalence, upper bound

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

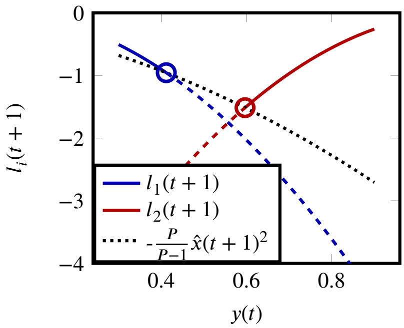
$$P = \frac{1}{2}(1 - \gamma^2 a^2) + \sqrt{\gamma^2(-1 + \gamma^2) + (\gamma^2 a^2 - 1)^2/4} > 1.$$

If P and γ fulfill the *curvature condition* (16) and *strong negativity condition* (17) below, then the closed-loop system (1) controlled with the certainty-equivalence deadbeat controller (14) has gain from $(w, v) \to x$ bounded above by γ.

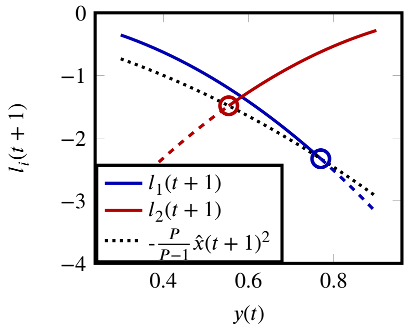
$$P > 2\gamma - 1 \tag{16}$$

$$(P + 2\gamma^2 - 1)\left(P - 1 - 2\sqrt{\gamma^2 - P})^2\right) \geq (P-1)\left((P+1)^2 - 4\gamma^2\right) \tag{17}$$

□



(a) $a = 1, \gamma = 4$



(b) $a = 1, \gamma = 3.4$

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 5 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.

REMARK 10
We can solve (17) 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 11
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 7, full. By assumption $P > 1$ is positive so Theorem 5 applies. We will show that if the *curvature condition* and the *strong negativity condition* are fulfilled, then the assumptions in Lemma 5 will hold for all t. Then, by Theorem 6 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 5 $\max_i\{l_i(t+1)\} \leq 0$. It remains to show that

$$\min_i\{l_i(t+1)\} \leq -\frac{P}{P-1}\hat{x}(t+1)^2. \tag{18}$$

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 4, 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 (18) 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

$$\gamma^4(1/X - 1/\gamma^2) \le -4\frac{P}{P-1}\hat{g}^2 \qquad (19)$$
$$\iff \quad \gamma^4 \ge -4\frac{P}{P-1}\frac{1}{1/X - 1/\gamma^2}\hat{g}^2$$

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

$$-4\frac{P}{P-1}\frac{1}{1/X - 1/\gamma^2}\hat{g}^2 = 4\frac{\gamma^2 - P}{(P-1)^2}\gamma^4. \qquad (20)$$

The concavity condition (19) simplifies to the *curvature condition* (16),

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

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 (18) 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 (18) with equality. We will start with characterizing $\bar{\mathcal{I}}_1$ by looking for the solutions to $\bar{l}_1(t+1) = -4\frac{P}{P-1}\hat{g}^2 z(t)^2$:

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

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

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

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

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

Using (20) we get

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

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 $\hat{x}(t)$,

$$
\begin{aligned}
\bar{\mathcal{I}}_{-1}=\Bigg[&\frac{1}{2\gamma^2}(P+2\gamma^2-1)\frac{P-1-2\sqrt{\gamma^2-P}}{(P+1)^2-4\gamma^2}P\hat{x}(t), \\
&\qquad\qquad \frac{1}{2\gamma^2}(P+2\gamma^2-1)\frac{P-1+2\sqrt{\gamma^2-P}}{(P+1)^2-4\gamma^2}P\hat{x}(t)\Bigg]
\end{aligned}
$$

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 $\#\left(\bar{\mathcal{I}}_1\cap\bar{\mathcal{I}}_2\right)\le 1$ is equivalent to

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

which simplifies to (17). The case when $\hat{x}(t)$ is negative is similar. □

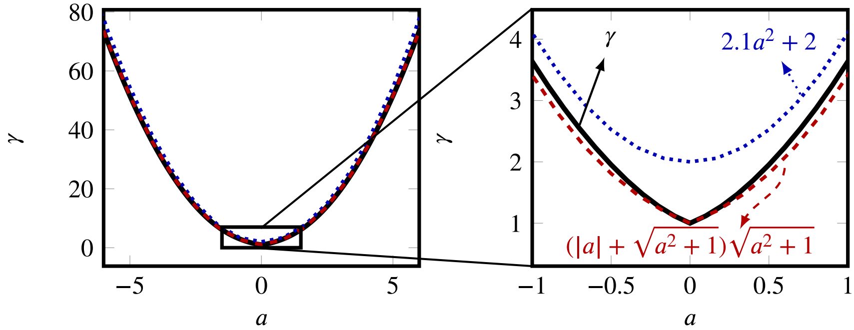


Figure 4. Guaranteed bound on the ℓ_2-gain from disturbances to error under feedback with the certainty equivalence controller with respect to a. We note that experimentally γ is lower bounded by $(|a| + \sqrt{a^2+1})\sqrt{a^2+1}$ and upper bounded by $\leq 2.1a^2 + 2$. The lower bound becomes tighter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

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 6 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 5 and 6 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.). *Proceedings* *of the 36th International Conference on Machine Learning*. Vol. 97. Proceedings of Machine Learning Research. PMLR, pp. 111–119. URL: `https://proceedings.mlr.press/v97/agarwal19c.html`.

Basar, T. and P. Bernhard (1995). *H_∞-Optimal Control and Related Minimax Design Problems — A dynamic Game Approach*. Birkhauser.

Dean, S., S. Tu, N. Matni, and B. Recht (2019). "Safely learning to control the constrained linear quadratic regulator". In: *2019 American Control Conference (ACC)*, pp. 5582–5588. DOI: `10.23919/ACC.2019.8814865`.

Didinsky, G. and T. Basar (1994). "Minimax adaptive control of uncertain plants". In: *Proceedings of 1994 33rd IEEE Conference on Decision and Control*. Vol. 3, 2839–2844 vol.3. 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: pp. 3724–3740. DOI: `10.1109/CDC40024.2019.9029916`.

Rantzer, A. (2021). *Minimax adaptive control for a finite set of linear systems*. arXiv: `2011.10814 [math.OC]`.

Simchowitz, M. (2020). "Making non-stochastic control (almost) as easy as stochastic". In: Larochelle, H. et al. (Eds.). *Advances in Neural Information Processing Systems*. Vol. 33. Curran Associates, Inc., pp. 18318–18329.

Vinnicombe, G. (2004). "Examples and counterexamples in finite l2-gain adaptive control".

Zhou, K. and J. C. Doyle (1998). *Essentials of Robust Control*. Prentice-Hall.

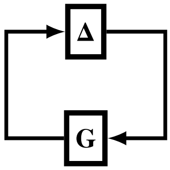


Figure 1. Interconnection of the stable system **G** and uncertainty **Δ** considered for robust stability.

Paper IV

ν-Analysis: A New Notion of Robustness for Large Systems with Structured Uncertainties

Olle Kjellqvist **John C. Doyle**

Abstract

We present a new, scalable alternative to the structured singular value, which we call ν, provide a convex upper bound, study their properties and compare them to ℓ_1 robust control. The analysis relies on a novel result on the relationship between robust control of dynamical systems and non-negative constant matrices.

1. Introduction

We consider a system to be robust if it is unlikely to fail. The usual setting to analyze the robustness of a system is to study how it interacts with uncertainty. Standard approaches impose structure on the uncertainty and certify robustness against its size. However, the way we currently measure the size of uncertainty is unsuitable for large-scale networks.

To see this, consider the standard robust control set-up in Fig. 1. **G** is a stable causal linear system with n inputs and outputs. **Δ** is unknown but belongs to the set

$\mathscr{D}$ consisting of diagonal linear time-varying (LTV) systems that are strictly causal, stable, and have n inputs and outputs. We want to determine which of the following two systems are most likely to fail:

$$\mathbf{P}_1 : \begin{cases} x_1(t+1) = \delta_1 x_1(t) \\ x_2(t+1) = \delta_2 x_2(t) \\ \quad \vdots \\ x_n(t+1) = \delta_n x_n(t) \end{cases}, \quad \mathbf{P}_2 : \begin{cases} x_1(t+1) = \delta_1 x_2(t) \\ x_2(t+1) = \delta_2 x_3(t) \\ \quad \vdots \\ x_n(t+1) = \delta_n x_1(t). \end{cases}$$

$\mathbf{P}_1$ is a set of decoupled first-order systems with uncertain time constants, and $\mathbf{P}_2$ is a delayed ring with uncertain weights. Robustness measures based on structured singular values [Zhou and Doyle, 1998; Dullerud and Paganini, 2010] or ℓ_1 robust control methods [Dahleh and Khammash, 1993] agree that both systems are robust against diagonal uncertainties whose *largest*[1] diagonal element is bounded by one. It is tempting to conclude that $\mathbf{P}_1$ and $\mathbf{P}_2$ are equally likely to fail. A more careful study reveals that destabilizing P_1 is easy; a constant gain $|\delta_k| > 1$ for any k will render the closed-loop unstable. However, all of the uncertainties must simultaneously be large ($\|\delta_1\|\|\delta_2\| \cdots \|\delta_n\| \geq 1$ to destabilize P_2. In plain words, destabilizing $\mathbf{P}_2$ requires large globally coordinated perturbations directly affecting every node.

This article proposes a new robustness measure ν[2] that captures sparsity in the uncertainty. ν is large for systems that are easily destabilized by sparse perturbations and small for systems that can withstand sparse perturbations. For example, $\nu(\mathbf{P}_1) = 1$ and $\nu(\mathbf{P}_2) = 1/n$. We focus on diagonal linear time-varying and nonlinear uncertainty in discrete time.

This work is primarily motivated by recent progress to distributed and localized controller design for large-scale networks [Anderson et al., 2019], modeling and analysis of the feedback in neuroanatomy [Stenberg et al., 2022; Li, 2022; Sarma et al., 2022] and the need for better control methods for emerging large-scale systems such as smart-grids and intelligent transportation systems. It is similar in spirit to [You and Matni, 2015] where the authors considered a sparse $\mathcal{H}_\infty$ analysis, but differ in that we consider systems in input/output form. Another approach to reduce conservativeness is to consider stochastic formulations for multiplicative uncertainty as in [Bamieh and Filo, 2020].

1.1 Outline

Section 2 introduces notation and gives some background on robust stability for static and dynamic matrices. We introduce and analyze the new robustness measure in Section 3 and provide a convex upper bound. Section 4 describes the properties of

[1] In $\mathcal{H}_\infty$– and ℓ_1–norm respectively

[2] The robustness measure ν is unrelated to Vinnicombe's ν-gap metric. We apologize for the confusion caused by overloading ν and highlight the need for further research into new Greek letters.

the upper bound and in Section 5 we show how to compute it and characterize the optimal solution. Concluding remarks and directions for future research are contained in Section 6.

2. Preliminaries and notation

This section contains a brief mathematical background, the reader is referred to the excellent textbook [Desoer and Vidyasagar, 1975]. Latin letters denote real-valued vectors and matrices like $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n\times m}$. For a matrix $A \in \mathbb{R}^{n\times m}$, A_{ij} means the element on the ith row and jth column, and we refer to the ith element of a vector $x \in \mathbb{R}^n$ by x_i. The p-norm of a vector $x \in \mathbb{R}^n$ is defined by

$$|x|_p := \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} & \text{if } p \in [0,\infty), \\ \max_i |x_i| & \text{if } p = \infty. \end{cases}$$

For a matrix, $A \in \mathbb{R}^{n\times m}$, the induced norm from q to p is defined by

$$|A|_{q,p} := \max_x \frac{|Ax|_p}{|x|_q}.$$

For an infinite sequence $\mathbf{x} = \{x(0), x(1), \ldots\}$, $x(k) \in \mathbb{R}^n$, $m_x = (|x_1|_\infty, \cdots, |x_n|_\infty)$ is called the magnitude vector of $\mathbf{x}$ and ℓ_∞^n denotes the set of all such sequences that satisfy

$$\|\mathbf{x}\|_\infty := |m_x|_\infty < \infty.$$

We define the truncation operator P_T on ℓ_∞^n by

$$P_T\mathbf{x} = (x(0), \ldots, x(T), 0, \ldots).$$

By $\ell_{\infty,e}^n$ we mean the extended ℓ_∞^n-space: $\{\mathbf{x} \in \ell_\infty^n : (P_T\mathbf{x}) \in \ell_\infty^n : T \geq 0\}$. An operator $\mathbf{H} : \ell_\infty^n \to \ell_\infty^n$ is causal if $P_k\mathbf{H} = P_k\mathbf{H}P_k$ and time-invariant if it commutes with the delay operator z^{-1}. We say that $\mathbf{H}$ is ℓ_∞ stable if

$$\|\mathbf{H}\|_1 := \|\mathbf{H}\|_{\infty,\infty} = \sup_t \sup_{0\neq\mathbf{x}\in\ell_{\infty,e}} \frac{\|P_t\mathbf{H}\mathbf{x}\|_\infty}{\|P_t x\|_\infty} < \infty,$$

where $\|\mathbf{H}\|_{\infty,\infty}$ is called the induced norm on ℓ_∞^n. A linear time-varying operator $\mathbf{G}$ is fully characterized by its impulse response (convolution kernel) $G(t,\tau)$ and operates on signals $\mathbf{x} \in \ell_{\infty,e}$ by convolution, $(\mathbf{G}\mathbf{x})(t) = \sum_{\tau=0}^t G(t,\tau)x(\tau)$. Expressed in the elements of its convolution kernel, the induced norm becomes $\|\mathbf{G}\|_1 = \max_i \sum_{j=1}^n \sup_t \sum_{\tau=0}^t |G_{ij}(t,\tau)|$. It will be convenient to express the norm in terms of the magnitude matrix of $\mathbf{G}$

$$M_G := \begin{bmatrix} \|G_{11}\|_1 & \cdots & \|G_{1n}\|_1 \\ \vdots & \ddots & \vdots \\ \|G_{n1}\|_1 & \cdots & \|G_{nn}\|_1 \end{bmatrix}, \tag{1}$$

and $\|\mathbf{G}\|_1 = \max_i \sum_j M_{ij}$.

Table 1. Summary of matrix induced norms, adapted from [Tropp, 2004]. The norm on the domain (D) is determined by the column, and the codomain (CD) by the row.

CD\D	$\lvert\cdot\rvert_1$	$\lvert\cdot\rvert_2$	$\lvert\cdot\rvert_\infty$
$\lvert\cdot\rvert_1$	$\max_j \sum_{i=1}^n \lvert A_{ij}\rvert$	NP-HARD	NP-HARD
$\lvert\cdot\rvert_2$	$\sqrt{\max_j \sum_{i=1}^n \lvert A_{ij}\rvert^2}$	$\overline{\sigma}(A)$	NP-HARD
$\lvert\cdot\rvert_\infty$	$\max_{ij} \lvert A_{ij}\rvert$	$\sqrt{\max_i \sum_{j=1}^n \lvert A_{ij}\rvert^2}$	$\max_i \sum_{j=1}^n \lvert A_{ij}\rvert$

2.1 Matrix induced norms and stability of static systems

Before diving into induced norms for dynamical systems, we explore norms on constant n-dimensional vectors and square matrices. For a constant matrix $M \in \mathbb{R}^{n\times n}$, robust stability with respect to bounded *unstructured* uncertainty means that $\det(I - \Delta M)$ is invertible for all $|\Delta| \leq \gamma$ in some norm. Let $\frac{1}{p} + \frac{1}{q} = 1$, then $\det(I - \Delta M)$ is invertible for all $|\Delta|_{p,q} \leq 1$ if and only if $|M|_{q,p} < 1$. See Table 1 for a table of the most common compatible p-norms.

2.2 Robust stability with diagonal uncertainty

Let $\mathscr{D}$ be the set of ℓ_∞-stable causal linear time-varying operators whose off-diagonal elements are zero, and $D \subset \mathbb{R}^{n\times n}$ be the set of diagonal matrices with positive diagonal entries and define

$$\mu_{\mathscr{D}}(\mathbf{G}) = \frac{1}{\inf\{\|\mathbf{\Delta}\|_{\infty,\infty} : \mathbf{\Delta} \in \mathscr{D},\ (I - \mathbf{G}\mathbf{\Delta})^{-1} \text{ unstable}\}}. \quad (2)$$

The following Theorem characterizes robust stability of Fig. 1 as conditions on M_G.

THEOREM 8—THEOREM 2 IN [DAHLEH AND KHAMMASH, 1993]
For $\mathbf{\Delta} \in \mathscr{D}$ with $\|\mathbf{\Delta}\|_{\infty,\infty} \leq 1$, the following are logically equivalent :

1. The system in Fig. 1 is robustly stable.
2. $\rho(M_G) < 1$, where $\rho(\cdot)$ denotes the spectral radius.
3. $x \leq M_G x$ and $x \geq 0$ imply that $x = 0$.
4. $\inf_{D\in D} |DM_G D^{-1}|_{\infty,\infty} < 1$
5. $\mu_{\mathscr{D}}(\mathbf{G}) < 1$. □

3. ν: The new μ

Inspired by the role of LASSO [Tibshirani, 1996] in favoring sparse solutions to regression problems, we propose using the sum of ℓ_1 norms, $\sum_{i=1}^n \|\boldsymbol{\delta}_i\|_1$. One could go one step further and study robustness in the ℓ_0 (number of nonzero $\boldsymbol{\delta}_i$s) setting. However, any system that has a nonzero diagonal element can be destabilized by a local (possibly very large) perturbation and the corresponding robustness measure will be 1 for almost all systems and not really informative. The new robustness metric, ν, hits the sweet spot and is defined as follows:

DEFINITION 1—ν
Let $\mathscr{D}$ be the set of ℓ_∞-stable causal linear time-varying operators with n inputs and outputs, whose off-diagonal elements are zero. Given a causal linear system $\mathbf{G}$ with n inputs and outputs

$$\nu_{\mathscr{D}}(\mathbf{G}) := \frac{1}{\inf\{\sum_{i=1}^n \|\boldsymbol{\delta}_i\|_1 \,:\, \boldsymbol{\Delta} \in \mathscr{D},\ (I - \mathbf{G}\boldsymbol{\Delta})^{-1} \text{ unstable}\}}. \qquad \square$$

To study the properties of the new robustness measure and its relationship to μ, we require insight into the relationship between that of destabilizing the dynamical system $\mathbf{G}$ and its magnitude matrix M_G. From Theorem 8 we know that if there exists a $\boldsymbol{\Delta}$ that destabilizes Fig. 1, then there exists a constant matrix M_Δ with the same ℓ_1 norm so that $I - M_\Delta M_G$ is singular, and vice versa. Surprisingly, it turns out that the bounds on *each diagonal entry* of $\boldsymbol{\Delta}$ are equal to that of M_Δ.

THEOREM 9
Let $\mathscr{D}$ be the set of ℓ_∞-stable causal linear time-varying operators with n inputs and outputs, whose off-diagonal elements are zero. Further, let $\mathscr{C} \subset \mathbb{R}^{n\times n}$ be the set of non-negative diagonal matrices. Given upper bounds $\overline{\delta}_{ii}$ for $i = 1, \dots, n$ and a stable, causal $n \times n$-dimensional system $\mathbf{G}$, the following are logically equivalent:

1. There exists a $\boldsymbol{\Delta} \in \mathscr{D}$, where each diagonal element is bounded from above, $\|\boldsymbol{\delta}_{ii}\|_{\infty,\infty} \le \overline{\delta}_{ii}$, such that the system in Fig. 1 is unstable.
2. There exists a matrix $M_\Delta \in \mathscr{C}$, where each diagonal element is bounded from above, $\delta_{ii} \le \overline{\delta}_{ii}$, such that $I - \Delta M_G$ is singular. $\square$

Proof. We start by showing that the first claim implies the second. Let $R_\Delta = \mathrm{Diag}(\bar{\delta}_{11}, \dots, \overline{\delta}_{nn})$, then $\boldsymbol{\Delta} = \widehat{\boldsymbol{\Delta}} R_\Delta$ for some $\widehat{\boldsymbol{\Delta}} \in \mathscr{D}, \|\widehat{\boldsymbol{\Delta}}\|_1 \le 1$. As Fig. 1 with $\boldsymbol{\Delta}\mathbf{G} = \widehat{\boldsymbol{\Delta}}(R_\Delta \mathbf{G})$ is unstable, we conclude the existence of a diagonal non-negative matrix $\hat{M}_\Delta$ with $|\hat{M}_\Delta|_{\infty,\infty} \le 1$ so that $I - \widehat{\Delta} R_\Delta M_G$ is singular. Taking $M_\Delta = \widehat{\Delta} R_\Delta$ completes the first part of the proof.

The proof of the converse is identical but starts with M_G. $\square$

Theorem 9 implies that we can replace the ℓ_1 norm in (2) with any norm on the magnitude matrix of $\mathbf{\Delta}$ and get $\mu_{\mathscr{D}}(\mathbf{G}) = \mu_{\mathscr{C}}(M_G)$ for free.

Although we do not yet know how to compute ν, from Fig. 1 and Theorem 9 we know that ν must be absolutely homogeneous and invariant to similarity transforms with matrices that commute with $\mathscr{D}$. Furthermore, we can translate the equivalence relationship between $|\cdot|_1$ and $|\cdot|_\infty$ into a corresponding relationship between ν and μ. We summarize the above discussion with the following proposition:

PROPOSITION 5
With $\mathbf{G}$, $\mathscr{D}$, $\mathscr{C}$ as in Theorem. 9, let $\mathcal{D} \subset \mathbb{R}^{n\times n}$ be the set of non-negative diagonal matrices, then the following statements are true:

1. $\nu_{\mathscr{D}}(\mathbf{G}) = \nu_{\mathscr{C}}(M_G)$
2. $\nu_{\mathscr{D}}(a\mathbf{G}) = |a|\nu_{\mathscr{D}}(\mathbf{G})$ for $a \in \mathbb{R}$.
3. $\nu_{\mathscr{D}}(D\mathbf{G}D^{-1}) = \nu_{\mathscr{D}}(\mathbf{G})$ for $D \in \mathcal{D}$.
4. $\mu_{\mathscr{D}}(\mathbf{G})/n \le \nu_{\mathscr{D}}(\mathbf{G}) \le \mu_{\mathscr{D}}(\mathbf{G})$. □

The following theorem tightens the lower bound in 4) by zeroing out different diagonal elements. This result agrees with intuition because we can study how a system interacts with sparse uncertainty by testing the different sparsity patterns separately.

THEOREM 10
Given $\mathbf{G}$ and $\mathscr{D}$ as in Definition 1. Let $I = (i_1, i_2, \dots, i_m)$ with $m \le N$ and $i_k \ne i_l$ for $k \ne l$ be an index tuple, and consider the sub-matrix of M_G:

$$M_I = \begin{bmatrix} M_{i_1 i_1} & \cdots & M_{i_1 i_m} \\ \vdots & \ddots & \vdots \\ M_{i_m i_1} & \cdots & M_{i_m i_m} \end{bmatrix}. \tag{3}$$

Then $\nu_{\mathscr{D}}(\mathbf{G}) \ge \frac{\rho(M_I)}{m}$. □

Proof. Assume without loss of generality that $i_k = k$ for $k = 1, \dots, m$. This assumption can always be enforced by renaming the signals. Restrict $\mathbf{\Delta}$ by setting $\boldsymbol{\delta}_{kk} = 0$ for $k > m$. By Proposition 5 $\nu_{\mathscr{D}}(\mathbf{G}) = \nu_{\mathscr{D}}(M_G)$, so it is sufficient to give the proof in the constant matrix case. Let $\Delta_1 = \mathrm{Diag}(\delta_{11}, \dots, \delta_{mm})$ be the submatrix of Δ that is nonzero and partition M_G into

$$M_G = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix},$$

where $M_{11} \in \mathbb{R}^{m\times m}$. Thus $I - \Delta M_G$ is invertible if and only if $(I - \Delta_1 M_{11})$ is invertible, which is equivalent to $|\Delta_1|_{\infty,\infty} \le 1/\rho(M_I)$. From the fourth property of Proposition 5 we conclude that $\nu_{\mathscr{D}}(\mathbf{G}) \ge \rho(M_I)/(m)$. □

3.1 An upper bound of ν

If the norm on the magnitude matrix of Δ is one in the upper triangle of Table 1, then we can use the corresponding dual norm in the lower triangle to construct an upper bound.

Although the induced norm from ∞ to 1, in general, is NP-hard to compute, it coincides with the absolute sum for diagonal matrices. To see this, consider

$$|\Delta|_{\infty,1} = \sup_{|x|_\infty=1} \sum_{i=1}^{n} |\delta_{ii}x_i| = \sum_{i=1}^{n} |\delta_{ii}|. \tag{4}$$

Thus, if $|M_G|_{1,\infty} < 1/|\Delta|_{\infty,1}$ then $I - \Delta M_G$ is non-singular. As $\nu_{\mathscr{D}}$ is invariant under similarity transformations with $D \in \mathcal{D}$, we suggest the following upper bound:

$$\overline{\nu}_{\mathscr{D}}(\mathbf{G}) := \inf_{D\in\mathcal{D}} |DM_G D^{-1}|_{1,\infty} \tag{5}$$

The 1 to ∞ norm is the maximum absolute element of a matrix, see Table 1, and can be computed for large-scale connected systems by local evaluation and communication with the closest neighbors.

We end this section by noting that for positive systems, the $\mathcal{H}_\infty$-norm is achieved by a stationary input [Rantzer, 2015; Colombino and Smith, 2016], so robustness analysis can be done entirely on positive matrices in that case too. We suspect one can derive similar results for positive systems as those in this article.

CONJECTURE 1

For positive systems, there exists a similar convex upper bound for a robustness measure against a causal, diagonal, linear time-varying uncertainty $\mathbf{\Delta}$ bounded in the following norm $\|\mathbf{\Delta}\| = \|\boldsymbol{\delta}_1\|_\infty + \|\boldsymbol{\delta}_2\|_\infty + \cdots + \|\boldsymbol{\delta}_n\|_\infty$. □

4. Properties of $\overline{\nu}$.

The lower bound in Theorem 10 shows that if the maximum absolute value is achieved on the diagonal of M_G, then the upper bound coincides with the lower bound and is exact. These types of systems are called *diagonally maximal* and merit a formal definition.

DEFINITION 2—DIAGONALLY MAXIMAL

A Matrix $A \in \mathbb{R}^{n\times n}$ is *diagonally maximal* if the maximum absolute element of A appears on the diagonal. A dynamical system $\mathbf{G}$ is diagonally maximal if its magnitude matrix M_G is diagonally maximal. □

The following important corollary follows from applying Theorem 10 to each diagonal element.

COROLLARY 2
If the matrix DM_GD^{-1} is *diagonally maximal* for some $D \in \mathcal{D}$, then $\overline{\nu}_{\mathscr{D}}(\mathbf{G}) = \nu_{\mathscr{D}}(\mathbf{G})$. □

Going back to the systems $\mathbf{P}_1$ and $\mathbf{P}_2$ in the introduction, we see that $\mathbf{P}_1$ is diagonal and hence diagonally maximal and $\nu_{\mathscr{D}}(\mathbf{P}_1) = \overline{\nu}_{\mathscr{D}}(\mathbf{P}_1) = 1$. However, for $\mathbf{P}_2$ the upper bound is conservative. Indeed, $1/n = \nu_{\mathscr{D}}(\mathbf{P}_2) \le \overline{\nu}_{\mathscr{D}}(\mathbf{P}_2) = 1$. The following theorem describes the gap between ν and $\overline{\nu}$.

THEOREM 11
With ν, $\mathbf{G}$ and $\mathscr{D}$ as in Definition 1 and $\overline{\nu}$ as in (5), it is true that $1 \le \overline{\nu}_{\mathscr{D}}(\mathbf{G})/\nu_{\mathscr{D}}(\mathbf{G}) \le n$. Furthermore, the lower bound is achieved by systems $\mathbf{G}$ that are *diagonally maximal* under some similarity transform D that commutes with $\mathscr{D}$. Pure rings achieve the upper bound. □

Proof. By construction $\overline{\nu}_{\mathscr{D}}(\mathbf{G}) \ge \nu_{\mathscr{D}}(\mathbf{G})$, and by Corollary 2 the upper bound is exact for systems that are diagonally maximal under some similarity transform that commutes with $\mathscr{D}$.

By $|M_G|_{1,\infty} \le |M_G|_{\infty,\infty}$ and Proposition 5 we have that $\overline{\nu}_{\mathscr{D}}(\mathbf{G}) \le \mu_{\mathscr{D}}(\mathbf{G}) \le n\nu_{\mathscr{D}}(\mathbf{G})$. It remains to show that the upper bound is achieved for pure ring systems. After scaling, balancing, and relabeling the signals, a pure ring system is of the form

$$x_1(t+1) = \delta_{11}x_2(t), \quad \cdots, \quad x_n(t+1) = \delta_{nn}x_1(t).$$

By Proposition 5, $\nu_{\mathscr{D}}(G) = \nu_{\mathscr{C}}(M_G)$, so we will study the null space of $I - M_G\Delta$. $I - M_G\Delta$ has a nontrivial null space if for some non-zero $w \in \mathbb{R}^n$,

$$(I - M_G\Delta)w = 0 \iff \begin{bmatrix} w_1 - \delta_{22}w_2 \\ w_2 - \delta_{33}w_3 \\ \vdots \\ w_n - \delta_{11}w_1 \end{bmatrix} = 0.$$

If $w_1 = 0$, then by substitution we must have $w = 0$. So assume without loss of generality that $w_1 = 1$. Then we have that $I - M_G\Delta$ has a nontrivial null space if and only if

$$\delta_{11} \cdots \delta_{nn} = 1. \tag{6}$$

We proceed to lower bound $\sum_{i=1}^n \delta_{ii}$ by minimizing it subject to (6). Substitute $\delta_{nn} = 1/\prod_{i=1}^{n-1} \delta_{ii}$ into the sum to transform the constrained optimization problem into a convex optimization problem over $\delta_{ii} > 0$ with the solution $\min_{\delta_{ii}} \sum_{i=1}^n \delta_{ii} = n$. Substitute the lower bound on a destabilizing Δ into Definition 1 to get $\overline{\nu}_{\mathscr{D}}(\mathbf{G}) \ge n\nu_{\mathscr{D}}(\mathbf{G})$ as $\overline{\nu}_{\mathscr{D}}(\mathbf{G}) = 1$. Since the upper bound is equal lower bound, we conclude that the bound is achieved. □

By the discussion in this section it is clear that even though $\overline{\nu}$ bounds ν, the gap can be pretty significant. It stands to reason that $\overline{\nu}$ is exact for some class of disturbances.

CONJECTURE 2
$\overline{\nu}$ is exact for some class of norm-bounded disturbances. □

We conclude this section by studying 2×2 matrices.

4.1 A closed-form formula for 2×2 matrices

Consider without loss of generality, matrices $M \in \mathbb{R}^{2\times 2}$ of the form

$$M = \begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix}.$$

If $x > 1$ or $y > 1$ we know that $\nu_{\mathscr{C}}(M) = \max\{x, y\}$ so only the case $0 < x, y < 1$ remains. We begin by parameterizing all destabilizing Δ in δ_{22}. Setting the determinant to zero we get

$$\frac{1}{\det(M)}\left(y + \frac{-1}{x - \det(M)\delta_{22}}\right) = \delta_{11}.$$

Thus $\nu_{\mathscr{C}}(M) = \delta_{11}(\delta_{22}) + \delta_{22}$ is convex on the domain $[0, 1]$ and the minimum is achieved either on the boundary or at a stationary point. For $0 < x, y < 1$ we have that

$$\delta_{11} = \frac{y-1}{\det(M)}, \quad \delta_{22} = \frac{x-1}{\det(M)}, \nu_{\mathscr{C}}(M) = \frac{\det(M)}{x+y-2}. \tag{7}$$

In Fig. 2 we compare the new robustness metric ν, the upper bound $\overline{\nu}$ and μ for 2×2-matrices. We see that $\overline{\nu}$ is exact for and only for matrices that are diagonally maximal under some $D \in \mathcal{D}$ and conclude that even for diagonally maximal systems, ν and μ can be very different. As the closed-loop maps generated by system-level synthesis often seem to be diagonally maximal, we conclude that for a large class of relevant systems, computing both $\overline{\nu}$ and μ gives additional information into the nature of destabilizing disturbances even for this class of systems. Based on this observation we state the following conjecture.

CONJECTURE 3
$\overline{\nu}_{\mathscr{C}}(M) = \nu_{\mathscr{C}}(M)$ only if DMD^{-1} is diagonally maximal for some $D \in \mathscr{C}$. □

5. Computing $\overline{\nu}$

5.1 The convex approach

This section explains how to formulate $\overline{\nu}$ as a linear program. Let $M \in \mathbb{R}^{n\times n}$ be a positive matrix. We want to compute

$$\inf_{D\in \mathcal{D}} \max_{ij} \left\{ M_{ij} \frac{d_i}{d_j} \right\}. \tag{8}$$

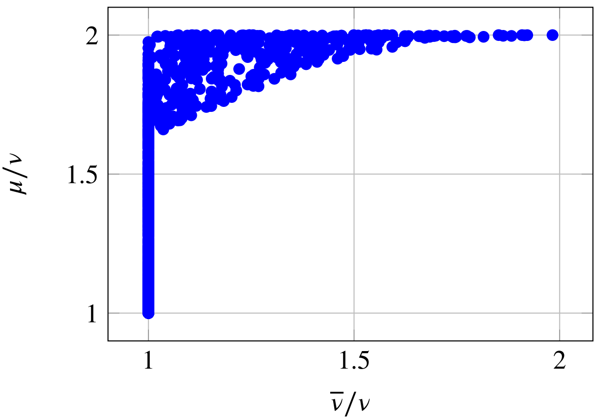


Figure 2. Comparing the ℓ_1-robustness metric μ, the new metric ν and the upper bound $\overline{\nu}$ for matrices of the form $M = \begin{bmatrix} x & w \\ w & y \end{bmatrix}$ for $x, w, y \in [0, 1]$. The matrices along the $\overline{\nu}/\nu = 1$ line are the diagonally maximal matrices. In the bottom left corner we have the identity matrix, in the top left corner we have the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and in the top right corner we have $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

As the logarithm is strictly increasing, (8) is equivalent to

$$\min_{D \in \mathcal{D}} \max_{ij} \{\log(M_{ij}) + \log(d_i) - \log(d_j)\},$$

where we use the convention that $\log(0) = -\infty$. Let $\beta_i = \log(d_i)$, then (8) is equivalent to the following linear program that can be solved efficiently using simplex or interior-point methods [Todd, 2002]:

$$\begin{aligned} &\underset{\beta \in \mathbb{R}^n, \gamma \in \mathbb{R}}{\text{minimize}} && \gamma \\ &\text{subject to:} && \log(M_{ij}) + \beta_i - \beta_j \leq \gamma. \end{aligned} \tag{9}$$

5.2 Characterizing the solutions of the upper bound

We will relax the positivity assumption of $d_1, \dots, d_n$ (8) to allow d_is to be zero. Consider the function

$$\phi_d(M, i, j) = \begin{cases} M_{ij} \frac{d_i}{d_j} & \text{if } M_{ij} > 0 \\ 0 & \text{if } M_{ij} = 0. \end{cases} \tag{10}$$

Then (8) is equivalent to

$$\inf_{d_1, \dots, d_n \geq 0} \max_{ij} \phi_d(M, i, j). \tag{11}$$

The following theorem shows that if for some $D \in \mathcal{D}$, the maximizing indices of DMD^{-1} only consists of loops, then D minimizes (11).

THEOREM 12—SUFFICIENT CONDITION FOR OPTIMALITY
Given a non-negative, non-zero matrix $M \in \mathbb{R}^{n\times n}$ and non-negative constants $d_1, \dots, d_n$. With ϕ as in (10), let $\mathcal{I}$ be the set of maximizing indices of (8), i.e.

$$\mathcal{I} = \left\{ (k,l) : \phi_d(M,k,l) = \max_{ij} \phi_d(M,i,j) \right\}.$$

If for all $(k,l) \in \mathcal{I}$ it holds that

$$\phi_d(M,k,l) = \max_j \phi_d(M,l,j). \tag{12}$$

Then $d_1, \dots, d_n$ is an optimal solution to (8). □

Proof. First, we show that $\mathcal{I}$ must contain at least one loop. Let $(j_0, j_1) \in \mathcal{I}$, and let j_{k+1} be the smallest integer such that $\phi_d(M, j_k, j_{k+1}) = \max_j \phi_d(M, j_k, j)$. By induction $(j_k, j_{k+1}) \in \mathcal{I}$. Furthermore, as n is finite, and the selection rule for j_{k+1} is unique given j_k, there is a $K \geq 0$ and a $T \geq 1$ so that $j_{k+T} = j_k$ for all $k \geq K$. We denote the limit set containing such points by $I_* = \{j_k : k \geq K\}$.

Assume towards a contradiction that there are $d'_1, \dots, d'_n$ so that $\max_{ij} \phi_{d'}(M,i,j) < \max_{ij} \phi_d(M,i,j)$, and let $(j_0, j_1) \in \mathcal{I}_*$. Assume without loss of generality that $d'_{j_1} > d_{j_1}$, otherwise multiply every d'_i by a positive constant so that the assumption holds true. Let $j_2 = \arg\max_j \phi_d(M, j_1, j)$. By assumption, it must hold that $d'_{j_2} > d_{j_2} d'_{j_1}/d_{j_1}$. Continuing, we have that

$$d'_{j_{k+T}} > d_{j_{k+T}} \frac{d'_{j_{k+T-1}}}{d_{j_{k+T-1}}} > d_{j_{k+T}} \frac{d'_{j_k}}{d_{j_k}}.$$

However, since $j_{k+T} = j_k$ we have that $d'_k > d'_k$ which is a contradiction. □

By the above theorem, we know that if the maximum is achieved on a loop, then the solution is optimal. It turns out that an optimal solution must contain a loop. This is because if the maximum is achieved on a chain, we can perturb the scales at the end of the chain to make that value smaller, making the chain shorter. Repeating this process reduces all the elements in the maximal chain. We formalize this statement in the following Lemma:

LEMMA 6
Let $d^\star_1, \dots, d^\star_n$ be an optimal solution to (11), and let $\mathcal{I}$ be the set of maximizing indices as in Theorem 12. Then $\mathcal{I}$ contains at least one loop. □

Proof. If the optimal value is zero, all diagonal elements must be zero, and $(i,i) \in \mathcal{I}$ implies that $\mathcal{I}$ contains a loop. Assume towards a contradiction that $\mathcal{I}$ does not contain a loop and that the optimal value is greater than zero. Let $(j_0, j_1) \in \mathcal{I}$, and let j_{k+1} be the smallest integer such that $\phi_d(M, j_k, j_{k+1}) = \max_j \phi_d(M, j_k, j)$. By assumption there is a k such that

$$\phi_d(M, j_k, j_{k+1}) < \max_i \phi_d(M, i, j_k) \tag{13}$$

This means that there is a $d'_k > 0$ that decreases the right hand side of (13) so that the inequality still holds for j_k, but also holds for j_{k-1}. By induction, this must hold for $1, \dots, k$. Repeating for any other chain in $\mathcal{I}$, we conclude that $\max_{ij} \phi_d(M, i, j) > \max_{ij} \phi_{d'}(M, i, j)$, contradicting optimality. □

Theorem 12 and Lemma 6 indicate a relationship between solving (11) and balancing the matrix M in the maximum absolute element norm. The following theorem strengthens that connection and shows that we can always find a solution to (11) by balancing M.

THEOREM 13
For any non-negative matrix $M \in \mathbb{R}^{n \times n}$, there exists a non-negative solution $d_1, \dots, d_n$ to (11) such that

$$\max_{r \neq k} \phi_d(M, r, k) = \max_{c \neq k} \phi_d(M, k, c), \quad \forall\, k = 1, \dots, n. \tag{14}$$

□

Proof. We begin by proving the existence of a solution. Assume there is a sequence $i_1, \dots, i_m$ such that $M_{i_k i_{k+1}}, M_{i_m i_1} \neq 0$ for $k = 1, \dots, m$. Then (8) is bounded below by $\min\{M_{ij}, M_{ji}\}$ and (8) is equivalent to a linear program with a bounded solution and the minimum is achieved by some $d_1, \dots, d_n$. If the assumption is false, we can take $d = 0$ and the optimal value is zero. If M is a diagonal matrix, then the claim holds trivially. Assume M is not diagonal and let $\hat{M}$ be the matrix where $\hat{M}_{ij} = M_{ij}$ for $i \neq j$ and $\hat{M}_{ii} = 0$. Then $d_1, \dots, d_n$ are optimal for M if and only if they are optimal for $\hat{M}$. Note that (14) holds for a maximizing loop of $\hat{M}$. Let $d_1, \dots, d_n$ be an optimal solution to (8) for $\hat{M}$. By Lemma 6, the set of maximizing indices $\mathcal{I}$ contains at least one loop. Remove the rows and columns pertaining the loop from $\hat{M}$ to get the smaller matrix $\hat{M}_1$. By recursion on $\hat{M}_k$ we end up with a new set $d'_1, \dots, d'_n$ so that (14) is true. □

5.3 An algorithm for balancing the magnitude matrix

We end this section with a simple heuristic algorithm for computing (8) that results from enforcing (14) coordinate-wise in Algorithm 1. The algorithm is similar to Osborne's algorithm for balancing matrices in the Frobenius norm [Osborne, 1960], but balances a matrix in the maximum absolute-element norm and can be computed using local information and communication with the closest neighbors. We show some empirical convergence properties in Figs. 3 and 4. We remark that naively

Algorithm 1 Heuristic algorithm for solving (8)

Require: Non-negative $M \in \mathbb{R}^{n\times n}$, $\theta \in (0,1)$, T.

$d_k[1] \leftarrow 1$ **for** each $k = 1, \dots, n$

for $t = 1, \dots, T$ **do**

  for $k = 1, \dots, n$ **do**

  $$d_k[t+1] \leftarrow (1-\theta)d_k[t] + \theta \frac{\sqrt{\max_{r\neq k} M_{rk} d_r[t]}}{\sqrt{\max_{c\neq k} M_{kc}/d_c[t]}}$$

  end for

end for

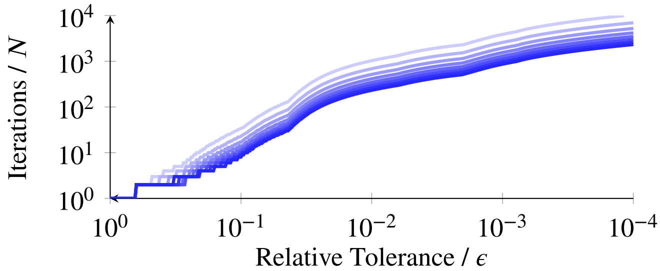


Figure 3. The largest number of iterations N required to reach a relative tolerance level for 500 randomly generated non-negative matrices $M \in \mathbb{R}^{128\times 128}$, with respect to tolerance. θ ranges from 0.2 (light) to 0.9 (dark).

taking $\theta = 1$ may cause the algorithm to fail to converge. Consider the matrix,

$$M = \begin{bmatrix} 0 & 1 \\ x^2 & 0 \end{bmatrix}.$$

Then $d_1(2) = x$ and $d_2(2) = 1/x$, leading to $D(2)MD^{-1}(2) = M^{\mathsf{T}}$ and the iteration will continue to oscillate back and forth. This is because we are updating each coordinate simultaneously, which is desirable for localized computation. Introducing the interpolation $\theta \in (0,1)$ seems to solve this issue. Based on the numerical results[3] we conjecture that our algorithm is guaranteed to converge.

CONJECTURE 4
Algorithm 1 always converges. Moreover the number of iterations required to reach a given tolerance is of $O(n)$ for a fixed ϵ, and $O(\sqrt{\epsilon^{-1}})$ for fixed n. □

[3] A Julia implementation of Algorithm 1 can be found at `https://github.com/kjellqvist/NuSynthesis.jl`.

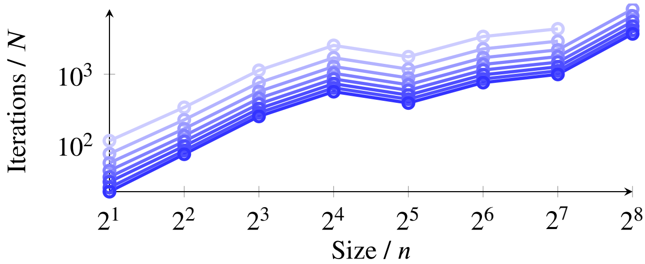


Figure 4. The largest number of iterations N required to reach a relative tolerance level of 10^{-3} for 500 randomly generated non-negative $n \times n$-dimensional matrices with respect to dimension n. θ ranges from 0.2 (light) to 0.9 (dark).

6. Conclusions

This work introduced and analyzed a new robustness measure ν that reasonably handles sparsity. We provided a convex upper bound $\overline{\nu}$, characterized its sub-optimality, and gave simple ways to compute it in a distributed way. The companion paper, [Li and Doyle, 2022] shows how to compute robust controllers for large-scale systems using μ and ν. Throughout this article, we gave four conjectures representing important research topics. We conclude with a final conjecture on the computation of ν.

CONJECTURE 5
There exists a polynomial-time algorithm to compute ν within arbitrary precision. □

References

Anderson, J., J. C. Doyle, S. H. Low, and N. Matni (2019). “System level synthesis”. *Annual Reviews in Control* **47**, pp. 364–393.

Bamieh, B. and M. Filo (2020). “An input–output approach to structured stochastic uncertainty”. *IEEE Transactions on Automatic Control* **65**:12, pp. 5012–5027. DOI: 10.1109/TAC.2020.2970393.

Colombino, M. and R. S. Smith (2016). “A convex characterization of robust stability for positive and positively dominated linear systems”. *IEEE Transactions on Automatic Control* **61**:7, pp. 1965–1971. DOI: 10.1109/TAC.2015.2480549.

Dahleh, M. A. and M. H. Khammash (1993). “Controller design for plants with structured uncertainty”. *Autom.* **29**, pp. 37–56.

Desoer, C. and M. Vidyasagar (1975). *Feedback Systems: Input–Output Properties*. Academic Press. ISBN: 978-0-12-212050-3. DOI: https://doi.org/10.1016/B978-0-12-212050-3.50009-8.

Dullerud, G. E. and F. Paganini (2010). *A Course in Robust Control Theory*. Springer New York.

Li, J. S. (2022). "Internal feedback in biological control: locality and system level synthesis". In: *2022 American Control Conference (ACC)*. To appear.

Li, J. S. and J. C. Doyle (2022). "Distributed robust control for systems with structured uncertainties". In: *61th IEEE Conference on Decision and Control (CDC)*. To appear.

Osborne, E. E. (1960). "On pre-conditioning of matrices". *J. ACM* **7**, pp. 338–345.

Rantzer, A. (2015). "Scalable control of positive systems". *European Journal of Control* **24**. SI: ECC15, pp. 72–80. ISSN: 0947-3580. DOI: `https://doi.org/10.1016/j.ejcon.2015.04.004`.

Sarma, A. A., J. S. Li, J. Stenberg, G. Card, E. S. Heckscher, N. Kasthuri, T. Sejnowski, and J. C. Doyle (2022). "Internal feedback in biological control: architectures and examples". In: *2022 American Control Conference (ACC)*. To appear.

Stenberg, J., J. S. Li, A. A. Sarma, and J. C. Doyle (2022). "Internal feedback in biological control: diversity, delays, and standard theory". In: *2022 American Control Conference (ACC)*. To appear.

Tibshirani, R. (1996). "Regression shrinkage and selection via the lasso". *Journal of the Royal Statistical Society (Series B)* **58**, pp. 267–288.

Todd, M. (2002). "The many facets of linear programming". *Mathematical Programming* **91**. DOI: `10.1007/s101070100261`.

Tropp, J. (2004). "Topics in sparse approximation".

You, S. and N. Matni (2015). "A convex approach to sparse $\mathcal{H}_\infty$ analysis & synthesis". In: *54th IEEE Conference on Decision and Control (CDC)*, pp. 6635–6642. DOI: `10.1109/CDC.2015.7403264`.

Zhou, K. and J. C. Doyle (1998). *Essentials of Robust Control*. Prentice-Hall.

Paper V

On Infinite-horizon System Level Synthesis Problems

Abstract

System level synthesis is a promising approach that formulates structured optimal controller synthesis problems as convex problems. This work solves the distributed linear-quadratic regulator problem under communication constraints directly in infinite-dimensional space, without the finite-impulse response relaxation common in related work. Our method can also be used to construct optimal distributed Kalman filters with limited information exchange. We combine the distributed Kalman filter with state-feedback control to perform localized LQG control with communication constraints. We provide agent-level implementation details for the resulting output-feedback state-space controller.

1. Introcution

In recent years, control design for networked dynamical systems has seen tremendous interest and progress. An important problem is to impose structures on the controllers, such as sparsity for distributed or localized control [Lessard and Lall, 2012; Wang et al., 2014], and communication delay constraints [Wang et al., 2014; Feyzmahdavian et al., 2012]. Such control design problems are challenging due to the non-convex nature of the problem [Rotkowitz and Lall, 2005].

Lately, researchers have focused on novel controller parameterization that admits convex formulation [Wang et al., 2019; Furieri et al., 2019; Sab**u***a*u et al., 2021], with System Level Synthesis (SLS) emerging as a promising and unified framework for structured controller synthesis [Anderson et al., 2019]. A vital feature of the SLS framework is that both the synthesis and the implementation of the structured controller can be done *locally*, thus scaling favorably with the number of subsystems in a network.

All current SLS-based control methods require both the parameterization and implementation to have finite impulse responses (FIR), with the exception of [Yu et al., 2021; Fisher et al., 2022]. This is because optimal controller synthesis, despite the choice of convex reparameterization, is an infinite-dimensional non-convex optimization problem over dynamical systems. The current method of choice to relax the problem into a tractable finite-dimensional optimization problem is to restrict the optimization variable to having a finite impulse response. Such relaxation technique is required for many parameterizations other than SLS [Zheng et al., 2022]. Although previous work almost exclusively uses FIR approximations, we emphasize that FIR is not a requirement for SLS, but rather a convenient way to use off-the-shelf optimization software. Therefore, lifting the FIR constraint will render SLS applicable to detectable and stabilizable systems.

An exception is [Yu et al., 2021], where the authors showed that a class of the infinite-horizon state-feedback SLS problem has natural connections to the Riccati solution for linear quadratic regulator (LQR) problems. The resulting SLS controller has a state-space form that significantly reduces the required memory compared to the FIR SLS controllers.

It is a natural next step to investigate the correspondence between infinite-horizon output-feedback SLS and the linear quadratic Gaussian (LQG) control [Kalman et al., 1960; Wonham, 1968].

Contribution. This paper generalizes a previous result on infinite-horizon state-feedback SLS problem, and investigates a class of infinite-horizon output-feedback SLS problems. In particular, we study an output-feedback SLS problem that corresponds to a class of LQG problems with structural constraints expressable as convex sparsity constraints under the SLS parameterization. Our contribution is three-fold. (1) We generalize the result on the infinite-horizon state-feedback SLS solution [Yu et al., 2021] to scenarios where communication delay constraints among subsystems in a network can be incorporated into controller synthesis and implementation. (2) We provide a suboptimal solution to a class of infinite-horizon output-feedback SLS problems. Our solution leverages an analogous separation principle for SLS parameterization, where the proposed generalized infinite-horizon state-feedback SLS solution is used. A key advantage of our approach is the ability to compute the parameters of each subsystem locally in one swoop using *local* information without iterations or communications among subsystems, which were required by previous methods. (3) We demonstrate an internally stabilizing output-feedback controller that is distributed and localized based on the proposed suboptimal solution. The proposed state-space controller has a fixed, low memory requirement, unlike existing FIR-based SLS controllers where the length of the memory grows linearly with the FIR horizon.

Paper Structure. We introduce the infinite-horizon SLS problems and related concepts in Section 2. Section 3 proposes a generalized approach that solves the infinite-horizon state-feedback SLS problems with communication delay and localization constraints. In Section 4, we construct a suboptimal solution to the infinite-horizon output-feedback SLS problem using the proposed optimal state-feedback solution. In Section 5, we show the implementation of the state-space controller associated with the constructed suboptimal solution. Section 6 demonstrates numerical simulation that corroborates with the theoretical results.

Notation. Bold font $\mathbf{x}$ and $\mathbf{G}$ denote signals $\mathbf{x} = \{x(t)\}_{t=0}^{\infty}$ with $x(t) \in \mathbb{R}^n$, and proper transfer matrices $\mathbf{G}(z) = \sum_{i=0}^{\infty} z^{-i} G[i]$ with convolution kernels $G[i] \in \mathbb{R}^{m \times n}$. The j^{th} standard basis vector is denoted as e_j. For a matrix A, $A(i,j)$ refers to the $(i,j)^{\text{th}}$ element, $A(i,:)$ to the i^{th} row and $A(:,j)$ to the j^{th} column. Non-negative integers are denoted as $\mathbb{N}_+$. We write $A \succ B$ ($A \succeq B$) to mean that $A - B$ is a positive (semi)definite matrix. We use $\mathbb{R}\mathcal{H}_\infty$ for the space of all proper and real rational stable transfer matrices and denote $F \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty$ if and only if $zF \in \mathbb{R}\mathcal{H}_\infty$.

2. Preliminaries

We consider the following dynamical system

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + w(t) \\ y(t) &= Cx(t) + v(t), \end{aligned} \tag{1}$$

where $w(t) \sim \mathcal{N}(0, W)$, and $v(t) \sim \mathcal{N}(0, V)$ are respectively state and measurement noise independently and identically drawn at each time from the zero-mean Gaussian distributions. Here W and V are positive definite matrices. The general control design objective is to synthesize a linear controller $\mathbf{K}$ such that control action computed as $\mathbf{u} = \mathbf{K}\mathbf{y}$ stabilizes the closed-loop system while optimizing over a quadratic cost $J(\mathbf{x}, \mathbf{u}) := \lim_{T\to\infty} \mathbb{E} \sum_{t=0}^{T} x(t)^\mathsf{T} R x(t) + u(t)^\mathsf{T} Q u(t)$, with Q, $R \succ 0$. This is known as the linear quadratic Gaussian (LQG) problem, where the optimal controller is given by the combination of a state-feedback controller and a state observer[Kalman et al., 1960].

In this paper, we investigate a *distributed* variant of the LQG problem, where (1) is constructed from a network of heterogeneous subsystems with dynamical coupling. In particular, we aim to design a stabilizing output-feedback controller that respects *communication delays* and *localization constraints*. For detalied discussions of such constraints, we refer interested readers to [Yu et al., 2021; Wang et al., 2018].

2.1 System Level Synthesis

The SLS theory approaches the constrained output-feedback control design problem described above by characterizing all achievable closed-loop mappings (CLMs) from $\mathbf{w}, \mathbf{v}$ to $\mathbf{x}, \mathbf{u}$ under an internally stabilizing controller $\mathbf{K}$. Then, using any achievable CLMs, SLS provides an implementation of the controller $\mathbf{K}$ that realizes the prescribed CLMs. This is made precise in the following result.

Theorem 14—[Anderson et al., 2019]
Strictly proper linear CLMs $\mathbf{\Phi}_{xx}, \mathbf{\Phi}_{xy}, \mathbf{\Phi}_{ux}, \mathbf{\Phi}_{uy} \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty$ can be achieved by a

linear internally stabilizing controller $\mathbf{K}$ if and only if

$$\begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{xx} & \mathbf{\Phi}_{xy} \\ \mathbf{\Phi}_{ux} & \mathbf{\Phi}_{uy} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix} \tag{2a}$$

$$\begin{bmatrix} \mathbf{\Phi}_{xx} & \mathbf{\Phi}_{xy} \\ \mathbf{\Phi}_{ux} & \mathbf{\Phi}_{uy} \end{bmatrix} \begin{bmatrix} zI - A \\ -C \end{bmatrix} = \begin{bmatrix} I \\ 0 \end{bmatrix}, \tag{2b}$$

where $\mathbf{\Phi}_{xx}$, $\mathbf{\Phi}_{xy}$, $\mathbf{\Phi}_{ux}$, $\mathbf{\Phi}_{uy}$ maps $\mathbf{w}, \mathbf{v}$ to $\mathbf{x}$, $\mathbf{u}$ under a controller $\mathbf{K}$, i.e., $\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_{xx} & \mathbf{\Phi}_{xy} \\ \mathbf{\Phi}_{ux} & \mathbf{\Phi}_{uy} \end{bmatrix} \begin{bmatrix} \mathbf{w} \\ \mathbf{v} \end{bmatrix}$. In particular, $\mathbf{K}$ can be implemented as the following, which is illustrated in Figure 1:

$$\begin{aligned} z\boldsymbol{\beta} &= \widehat{\mathbf{\Phi}}_{xx}\boldsymbol{\beta} + \widehat{\mathbf{\Phi}}_{xy}\mathbf{y} \\ \mathbf{u} &= \widehat{\mathbf{\Phi}}_{ux}\boldsymbol{\beta} + \mathbf{\Phi}_{uy}\mathbf{y}, \end{aligned} \tag{3}$$

where $\widehat{\mathbf{\Phi}}_{xx} = z(I - z\mathbf{\Phi}_{xx})$, $\widehat{\mathbf{\Phi}}_{ux} = z\mathbf{\Phi}_{ux}$, $\widehat{\mathbf{\Phi}}_{xy} = -z\mathbf{\Phi}_{xy}$, and β is the controller internal state. □

Further, it was shown in [Wang et al., 2019] that (2) is equivalent to stabilizability and detectability of (1). Therefore, (3) parameterizes all internally stabilizing linear controller $\mathbf{K}$ for (1).

A special case of Theorem 14 is when the controller is *state-feedback*, i.e., $\mathbf{u} = \mathbf{Kx}$. In this scenario, the SLS CLMs reduce to only $\mathbf{\Phi}_{xw} : \mathbf{w} \to \mathbf{x}$ and $\mathbf{\Phi}_{uw} : \mathbf{w} \to \mathbf{u}$ with the following variation of Theorem 14.

THEOREM 15—[WANG ET AL., 2019]
For the dynamics (1) with $C = I$ and $v(t) \equiv 0$, CLMs $\mathbf{\Phi}_{xw}$ and $\mathbf{\Phi}_{uw}$ can be achieved by a linear internally stabilizing controller $\mathbf{K}$ if and only if

$$\begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{xw} \\ \mathbf{\Phi}_{uw} \end{bmatrix} = I, \quad \mathbf{\Phi}_{xw}, \mathbf{\Phi}_{uw} \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty. \tag{4}$$

□

2.2 System-level Constraints (SLCs)

In this paper, we consider a networked system (1) composed of N subsystems, where each subsystem i has the following dynamics

$$x_i(t+1) = \sum_{j \in \mathcal{N}_{\text{in}}^1(i)} \left(A_{ij}x_j(t) + B_{ij}u_j(t) \right) + w_i(t) \tag{5}$$

$$y_i(t) = C_i x_i(t) + v_i(t) \tag{6}$$

where we denote $j \in \mathcal{N}_{\text{in}}^k(i)$ if the states and control actions of subsystem j affect those of subsystem i in k time steps through the open-loop network dynamics.

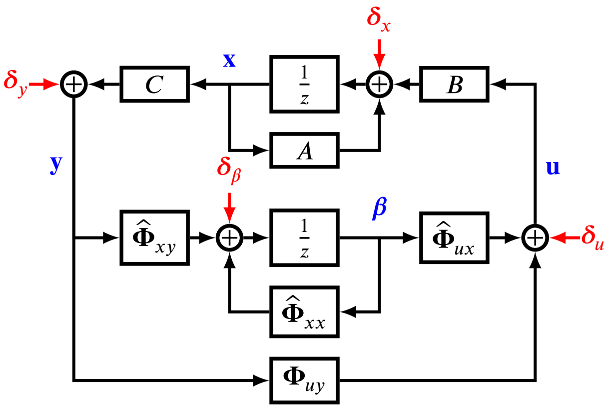


Figure 1. Output-feedback controller architecture. Here, $\widehat{\mathbf{\Phi}}_{xx} = z(I - z\mathbf{\Phi}_{xx})$, $\widehat{\mathbf{\Phi}}_{ux} = z\mathbf{\Phi}_{ux}$ and $\widehat{\mathbf{\Phi}}_{xy} = -z\mathbf{\Phi}_{xy}$. The controller is internally stable; the closed-loop maps from perturbations $(\boldsymbol{\delta}_x, \boldsymbol{\delta}_y, \boldsymbol{\delta}_u, \boldsymbol{\delta}_\beta)$ to internal signals $(\mathbf{x}, \mathbf{y}, \mathbf{u}, \boldsymbol{\beta})$ are stable.

Analogously, we write $i \in \mathcal{N}^k_{\text{out}}(j)$ when the states of subsystem i are affected by subsystem j via dynamics in k time steps.

For such systems, it is common to impose structural constraints such as localization requirement and communication delay for control design. A large class of such structural constraints can be characterized as convex constraints on CLMs. This is called the system-level constraints (SLCs) [Wang et al., 2019]. Examples of SLCs include sparsity constraints on the Youla parameter, QI subspace constraints on controller $\mathbf{K}$, and FIR constraints on the closed-loop responses to noises.

In this paper, we consider the family of SLCs that can be formulated as sparsity constraints on the CLMs. This family includes disturbance localization and communication delay constraints considered in [Yu et al., 2021]. We denote any such sparsity constraints as $\mathcal{S} := \{S[k]\}_{k=1}^{\infty}$ where each $S[k] \in \{0,1\}^{n\times n}$ are binary matrices specifying the sparsity of the kernels of CLMs. Specifically, we consider the *localization SLCs*, which specifies how disturbances at each subsystem must be localized to a neighborhood.

Definition 3—Localization SLCs
An SLC $\mathcal{S} := \{S[k]\}_{k=1}^{\infty}$ is called the localization SLC if $S[k]$ for all k are the same binary matrices. □

In addition to disturbance localization which assumes instantaneous information exchange within a localization neighborhood, one may wish to incorporate *communication delay SLCs*[1]. Let $\mathrm{Sp}(\cdot)$ denote the support of a matrix.

[1] Such delayed localization pattern corresponds to scenarios when communication delays allow disturbances to propagate through dynamics, before subsystems are able to completely attenuate and

DEFINITION 4—D-DELAYED LOCALIZATION SLCS
For a fixed integer $d \geq 1$ and sparsity pattern of the dynamics $\mathbb{A} := \text{Sp}(A)$, a delayed localization SLC is such that $S[k] = \text{Sp}\left(\mathbb{A}^k\right)$ for $k \leq d$ and $S[k] = \text{Sp}\left(\mathbb{A}^d\right)$ for all $k \geq d$. □

This is sometimes called the (A, d)-sparsity [Wang et al., 2014] and generalizes the localization SLCs. We say $\mathbf{\Phi}_{xx} \in \mathcal{S}$ if $\Phi_{xx}[k]$ has the same sparsity pattern as $S[k]$ for all $k \in \mathbb{N}_+$. For the rest of paper, we consider d-delayed localization SLCs for structured controller synthesis. We assume that any given SLCs are feasible for the underlying system.

2.3 The State-feedback and Output-feedback SLS Problem

Given d-delayed localization SLCs $\mathcal{S}_x$ and $\mathcal{S}_u$ specifying the state and input closed-loop sparsity respectively, we now state the output-feedback (OF) SLS problem [Wang et al., 2019].

$$
\begin{aligned}
\min_{\mathbf{\Phi} \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty} \quad & \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{xx} & \mathbf{\Phi}_{xy} \\ \mathbf{\Phi}_{ux} & \mathbf{\Phi}_{uy} \end{bmatrix} \begin{bmatrix} W^{1/2} & 0 \\ 0 & V^{1/2} \end{bmatrix} \right\|_{\mathcal{H}_2} \\
\text{s.t.} \quad & \text{Constraints (2)}, \\
& \Phi_{xx}[k]\,, \Phi_{xy}[k] \in S_x[k] \text{ for } k \in \mathbb{N}_+ \\
& \Phi_{ux}[k]\,, \Phi_{uy}[k] \in S_u[k] \text{ for } k \in \mathbb{N}_+\,.
\end{aligned}
\tag{OF-SLS}
$$

where we used $\mathbf{\Phi}$ to collectively refer to the tuple $\left(\mathbf{\Phi}_{xx}, \mathbf{\Phi}_{ux}, \mathbf{\Phi}_{xy}, \mathbf{\Phi}_{uy}\right)$ to reduce notation. Control problem (OF-SLS) is similar to the classical LQG problem but differs in the additional constraints on the system responses that corresponds to disturbance localization and information delay.

As a special case, the state-feedback (SF) SLS problem is

$$
\begin{aligned}
\min_{\mathbf{\Phi}_{xw}, \mathbf{\Phi}_{uw}} \quad & \left\| \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{xw} \\ \mathbf{\Phi}_{uw} \end{bmatrix} \right\|_{\mathcal{H}_2} \\
\text{s.t.} \quad & \text{Constraints (4)} \\
& \Phi_{xw}[k] \in S_x[k], \;\; \Phi_{uw}[k] \in S_u[k] \text{ for } k \in \mathbb{N}_+\,.
\end{aligned}
\tag{SF-SLS}
$$

In this section, we derive the optimal solution to (SF-SLS). This generalizes [Yu et al., 2021] to account for the *delayed* localization constraints. This solution allows the subsequent output-feedback controller synthesis.

localize them. For ease of exposition, here we consider a communication delay pattern among subsystems that matches the dynamics. The results in this paper can be generalized to broader classes of communication patterns.

Because (SF-SLS) is column-wise separable [Wang et al., 2018], we will synthesize the closed-loop maps one column at a time so that each subsystem can synthesize the columns corresponds to its local states, in a parallel fashion. Such parallel synthesis scales favorably with the number of subsystems in the network. From here on, everything will be seen by the ith subsystem[2]. Let $\boldsymbol{\varphi_x} := \boldsymbol{\Phi}_{xw}(:,i)$ and $\boldsymbol{\varphi_u} := \boldsymbol{\Phi}_{uw}(:,i)$ with kernels $\varphi_x[k]$ and $\varphi_u[k]$ for $k \in \mathbb{N}_+$, respectively corresponding to the ith column of $\boldsymbol{\Phi}_{xw}$ and $\boldsymbol{\Phi}_{uw}$. Furthermore, we use $s_x[k]$ and $s_u[k]$ denote the ith column of $S_x[k]$ and $S_u[k]$ respectively. Each corresponding column problem to be solved locally by subsystem i becomes

$$\begin{aligned} \min_{\boldsymbol{\varphi_x},\boldsymbol{\varphi_u}} \quad & \sum_{k=1}^{\infty} \varphi_x[k]^\mathsf{T} Q \varphi_x[k] + \varphi_u[k]^\mathsf{T} R \varphi_u[k] \\ \text{s.t.} \quad & \varphi_x[k+1] = A\varphi_x[k] + B\varphi_u[k] && \text{(7a)} \\ & \varphi_x[0] = 0, \quad \varphi_x[1] = e_i \\ & \varphi_x[k] \in s_x[k],\ \varphi_u[k] \in s_u[k]. && \text{(7b)} \end{aligned}$$

This new problem is a constrained linear quadratic optimal control problem, and would be a standard infinite-horizon LQR problem if not for the sparsity constraints (7b). We will show how to transform this problem to a finite-horizon LQR problem with time-varying dynamics.

3. Derivation of the Optimal Solution

Let $n_x[k]$ be the number of nonzero elements in the n-dimensional vector $s_x[k]$ and $n_u[k]$ be the number of nonzero elements in $s_u[k]$. Then there exists a surjective matrix $M_x[k] \in \mathbb{R}^{(n-n_x[k])\times n}$ and an injective matrix $M_u[k] \in \mathbb{R}^{m\times n_u[k]}$ such that (7b) is equivalent to

$$M_x[k]\varphi_x[k] = 0, \quad \varphi_u[k] = M_u[k]q[k], \tag{8}$$

where $q[k] \in \mathbb{R}^{n_u[k]}$ becomes the new variable. In particular, one can construct $M_x[k]$ by horizontally stacking standard basis vectors with nonzero positions corresponding to the positions that are zero in $\varphi_x[k]$. On the other hand, $M_u[k]$ can be obtained similarly but with basis vectors corresponding to the nonzero positions in $\varphi_u[k]$. Since $\varphi_x[k+1]$ is uniquely determined by $\varphi_x[k]$ and $\varphi_u[k]$, substitution of (8) into (7a) yields

$$M_x[k+1]A\varphi_x[k] + \underbrace{M_x[k+1]BM_u[k]}_{F[k]}q[k] = 0. \tag{9}$$

[2] To reduce notation, we assume each subsystem has scalar dynamics. One can alleviate this assumption by running the algorithm for multiple columns per subsystem.

The solutions to (9) can be expressed as

$$q[k] = F[k]^{\dagger} M_x[k+1] A \varphi_x[k] + N_F[k] r[k], \tag{10}$$

where $N_F[k] \in \mathbb{R}^{n_u[k] \times n_r[k]}$ is a bijection onto the nullspace of $F[k]$. The vector $r[k] \in \mathbb{R}^{n_r[k]}$ is now our new unconstrained optimization variable. Substituting $\varphi_u[k] = M_u[k] q[k]$ and (10) into the optimal control problem we get the equivalent time-varying LQR problem

$$\begin{aligned} \min_{r[k] \in \mathbb{R}^{n_r[k]}} \quad & \sum_{k=1}^{\infty} \Big(\varphi_x[k]^{\mathsf{T}} \tilde{Q}[k] \varphi_x[k] + \\ & \qquad 2 r[k]^{\mathsf{T}} \tilde{Z} \varphi_x[k] + r[k]^{\mathsf{T}} \tilde{R}[k] r[k] \Big) \\ \text{s.t.} \quad & \varphi_x[k+1] = \tilde{A} \varphi_x[k] + \tilde{B}[k] r[k] \\ & \varphi_x[0] = 0, \quad \varphi_x[1] = e_i, \end{aligned} \tag{11}$$

where

$$\begin{aligned} \kappa[k] &= M_u[k] F[k]^{\dagger} M_x[k+1] A \\ \tilde{Z}[k] &= N_F[k]^{\mathsf{T}} M_u[k]^{\mathsf{T}} R \kappa[k], \quad \tilde{Q}[k] = Q + \kappa[k]^{\mathsf{T}} R \kappa[k] \\ \tilde{R}[k] &= (M_u[k] N_F[k])^{\mathsf{T}} R[k] M_u[k] N_F[k] \\ \tilde{A}[k] &= A - B\kappa[k], \quad \tilde{B}[k] = B M_u[k] N_F[k]. \end{aligned} \tag{12}$$

Finally, we note that for $k \geq d+1$, the localization patterns are constant, implying that the dynamics matrices of the transformed problem are static for $k \geq d+1$. Standard dynamic programming arguments allow us to first solve the Riccati equation for the time-invariant problem for $k \geq d+1$ to get the positive definite solution $\tilde{X}_\star$ and the feedback gain $\tilde{K}_\star$, and then to solve a finite-horizon time-varying problem by replacing the cost function of each column problem (11) with equivalent cost function

$$\begin{aligned} J = & \sum_{k=1}^{d} \Big(\varphi_x[k]^{\mathsf{T}} \tilde{Q}[k] \varphi_x[k] + 2 r[k]^{\mathsf{T}} \tilde{Z} \varphi_x[k] + r[k]^{\mathsf{T}} \tilde{R}[k] r[k] \Big) \\ & + \varphi_x[d+1]^{\mathsf{T}} \tilde{X}_\star \varphi_x[d+1]. \end{aligned} \tag{13}$$

To obtain $\tilde{X}_\star$ and $\tilde{K}_\star$, we invoke the results in [Yu et al., 2021] for the time-invariant problem for with staic localization pattern. Finally, the solution to the time-varying finite-horizon problem (11) with cost (13) is given by the Riccati iteration with $\tilde{X}[d+1] = \tilde{X}_\star$, and for $k = 1, \ldots, d$,

$$\begin{aligned}\tilde{X}[k] &= \tilde{Q}[k] + \tilde{A}[k]^\mathsf{T}\tilde{X}[k+1]\tilde{A}[k] - \\ &\quad \left(\tilde{A}[k]^\mathsf{T}\tilde{X}[k+1]\tilde{B}[k] + \tilde{Z}[k]\right) \\ &\quad \cdot \left(\tilde{R}[k] + \tilde{B}[k]^\mathsf{T}\tilde{X}[k+1]B[k]\right)^{-1}\left(\tilde{B}[k]^\mathsf{T}\tilde{X}[k+1]\tilde{A}[k]\right) \\ \tilde{K}[k] &= \left(\tilde{R}[k] + \tilde{B}[k]\tilde{X}[k+1]\tilde{B}[k]\right)^{-1} \\ &\quad \cdot \left(\tilde{B}[k]^\mathsf{T}\tilde{X}[k+1]\tilde{A}[k] + \tilde{Z}[k]^\mathsf{T}\right). \end{aligned} \tag{14}$$

Substituting $r[k] = \tilde{K}[k]\varphi_x[k]$ into (10) and further into (9), one can obtain the solution to the original problem (7). We formally state the optimality of the proposed solution.

THEOREM 16
Fix delayed localization SLCs $\mathcal{S}_x$ and $\mathcal{S}_u$. The optimal solution to the infinite-horizon state-feedback SLS problem in (SF-SLS) is given, in a column-wise fashion, by

$$\boldsymbol{\Phi}^\star_{xw}(:,i) = \left[\begin{array}{c|c} A^i_{\mathrm{SF}} & B^i_{\mathrm{SF}} \\ \hline C^i_{\mathrm{SF}} & 0 \end{array}\right], \ \boldsymbol{\Phi}^\star_{uw}(:,i) = \left[\begin{array}{c|c} A^i_{\mathrm{SF}} & B^i_{\mathrm{SF}} \\ \hline K^i_{\mathrm{SF}} & 0 \end{array}\right], \tag{15}$$

where

$$\begin{aligned} A^i_{\mathrm{SF}} &= \begin{bmatrix} 0 & \ldots & & & 0 \\ \tilde{A}_{\mathrm{CL},i}[1] & 0 & \ldots & & 0 \\ 0 & \tilde{A}_{\mathrm{CL},i}[2] & 0 & \ldots & 0 \\ 0 & 0 & \ddots & & \vdots \\ & & & \tilde{A}_{\mathrm{CL},i}[d] & \tilde{A}_{\mathrm{CL},\star,i} \end{bmatrix} \\ B^i_{\mathrm{SF}} &= \begin{bmatrix} e_i^\mathsf{T} & 0\ldots 0 \end{bmatrix}^\mathsf{T} \quad C^i_{\mathrm{SF}} = \begin{bmatrix} I & I\ldots & I \end{bmatrix} \\ K^i_{\mathrm{SF}} &= \begin{bmatrix} \tilde{K}_i[1] & \tilde{K}_i[2] & \ldots & \tilde{K}_i[d] & \tilde{K}_{\star,i} \end{bmatrix}, \end{aligned} \tag{16}$$

□

with $\tilde{K}_i[k]$ and $\tilde{A}_{\mathrm{CL},i}[k] := \tilde{A}[k] - \tilde{B}[k]\tilde{K}[k]$ computed using (12) and (14) for the ith column problem. Matrix $\tilde{A}_{\mathrm{CL},\star,i}$ and $\tilde{K}_{\star,i}$ are given by the infinite-horizon solution with static localization SLCs $\mathcal{S}_x \equiv S_x[d]$ and $\mathcal{S}_u \equiv S_u[d]$ using the method in [Yu et al., 2021, Section IV.A].

Proof. The optimality follows directly from the column separable property of (SF-SLS), and the equivalent transformations between (7) and (11). The finite-horizon LQR problem with cost (13) is equivalent to (11) by Bellman's optimality principle. It is straightforward to verify that (15) is a state-space realization of the solution to (11) by substituting the optimal solution $r[k]$ via (14) into (10). □

Compared to [Yu et al., 2021], the state-space realization of the optimal CLMs given by our approach has a higher order because of the first d-delay pattern. If we let the first d delay sparsity pattern to be the same as the localization pattern, then our approach subsumes the results in [Yu et al., 2021].

Given the column-wise state-space description of the optimal CLMs $\mathbf{\Phi}^{\star}_{xw}(:, i)$ and $\mathbf{\Phi}^{\star}_{xw}(:, i)$, we can adopt the same state-space agent-level controller proposed in [Yu et al., 2021, Section IV.B] by simply replacing the state-space implementation of the CLMs with the ones presented here.

3.1 Structured Kalman Filter Design

Theorem 16 can be used to solve the dual problem of optimal structured Kalman filter design with delayed localization SLCs for (1) [Wang et al., 2015]. In particular, the optimal structured infinite-horizon CLMs that map **w** and **v** to state estimation error **e** under a linear observer **L** with respect to the mean estimation error is given by the solution to the dual problem of (SF-SLS) as shown below:

$$\min_{\mathbf{\Phi}_{ew}, \mathbf{\Phi}_{ev} \in \frac{1}{z}\mathbb{R}\mathcal{H}_\infty} \quad \left\| \begin{bmatrix} W^{1/2} & 0 \\ 0 & V^{1/2} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{ew}^\top \\ \mathbf{\Phi}_{ev}^\top \end{bmatrix} \right\|_{\mathcal{H}_2} \tag{17}$$

$$\text{subject to} \quad \begin{bmatrix} \mathbf{\Phi}_{ew} & \mathbf{\Phi}_{ev} \end{bmatrix} \begin{bmatrix} zI - A \\ -C \end{bmatrix} = I \tag{18}$$

$$\mathbf{\Phi}_{ew} \in \mathcal{S}_x, \quad \mathbf{\Phi}_{ev} \in \mathcal{S}_u.$$

Readers are referred to [Wang et al., 2015] for detailed derivation. We highlight the resemblance between constraints (4) and (18), and (2). In what follows, we will use the optimal solutions from the state-feedback and Kalman-filter SLS problem to construct a suboptimal solution to the output-feedback SLS problem.

4. A solution Inspired by Separation Principle

It is well known that for a linear system, observer-based feedback is always stabilizing if the observer error dynamics are stable and the feedback gain stabilizes the state-feedback case. In [Wang et al., 2019], the authors pointed out that a similar property holds for CLMs from state-feedback and Kalman-filter SLS problems described in Section 3.

THEOREM 17—[ANDERSON ET AL., 2019]
Assume there exist stable and strictly proper transfer matrices $\mathbf{\Phi}^{\text{SF}} = (\mathbf{\Phi}^{\text{SF}}_{xw}, \mathbf{\Phi}^{\text{SF}}_{uw})$

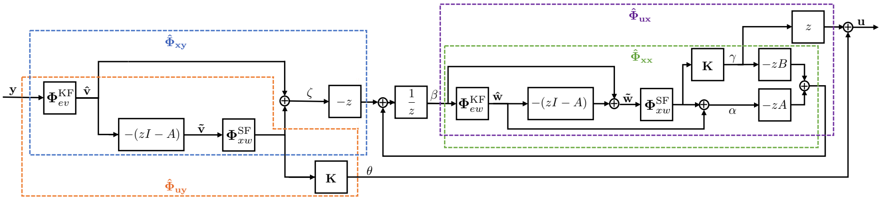


Figure 2. Controller implementation of Fig. 1 after plugging (19) in the controller (3).

and $\mathbf{\Phi}^{\mathrm{KF}} = (\mathbf{\Phi}^{\mathrm{KF}}_{ew}, \mathbf{\Phi}^{\mathrm{KF}}_{ev})$ satisfying

$$\begin{bmatrix} zI - A & -B \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}^{\mathrm{SF}}_{xw} \\ \mathbf{\Phi}^{\mathrm{SF}}_{uw} \end{bmatrix} = I,$$

$$\begin{bmatrix} \mathbf{\Phi}^{\mathrm{KF}}_{ew} & \mathbf{\Phi}^{\mathrm{KF}}_{ev} \end{bmatrix} \begin{bmatrix} zI - A \\ -C \end{bmatrix} = I.$$

The transfer functions

$$\begin{aligned} \mathbf{\Phi}_{xx} &= \mathbf{\Phi}^{\mathrm{SF}}_{xw} + \mathbf{\Phi}^{\mathrm{KF}}_{ew} - \mathbf{\Phi}^{\mathrm{SF}}_{xw}(zI - A)\mathbf{\Phi}^{\mathrm{KF}}_{ew} \\ \mathbf{\Phi}_{ux} &= \mathbf{\Phi}^{\mathrm{SF}}_{uw} - \mathbf{\Phi}^{\mathrm{SF}}_{uw}(zI - A)\mathbf{\Phi}^{\mathrm{KF}}_{ew} \\ \mathbf{\Phi}_{xy} &= \mathbf{\Phi}^{\mathrm{KF}}_{ev} - \mathbf{\Phi}^{\mathrm{SF}}_{xw}(zI - A)\mathbf{\Phi}^{\mathrm{KF}}_{ev} \\ \mathbf{\Phi}_{uy} &= -\mathbf{\Phi}^{\mathrm{SF}}_{uw}(zI - A)\mathbf{\Phi}^{\mathrm{KF}}_{ev} \end{aligned} \tag{19}$$

are strictly proper and satisfy (2). □

Then, output-feedback controller (3) can be constructed using CLMs from (19) to stabilize (1) while respecting the prescribed localization and communication delay constraints.

5. Local Controller implementation

This section describes Algorithm 2, which summarizes the local implementation of the global controller (3) in Fig. 1 using the localized state-feedback controllers and Kalman filters of Section 3 and Theorem 17.

Globally, the controller after plugging (19) in the controller (3) is shown in Fig. 2. Consider the intermediate signals in Fig. 2,

$$
\begin{aligned}
\boldsymbol{\zeta} &= \overbrace{-\boldsymbol{\Phi}^{\mathrm{SF}}_{xw}(zI - A)\boldsymbol{\Phi}^{\mathrm{KF}}_{ev}\mathbf{y}}^{\tilde{\mathbf{v}}} + \overbrace{\boldsymbol{\Phi}^{\mathrm{KF}}_{ev}\mathbf{y}}^{\hat{\mathbf{v}}}\\
\boldsymbol{\theta} &= -\boldsymbol{\Phi}^{\mathrm{SF}}_{uw}(zI - A)\boldsymbol{\Phi}^{\mathrm{KF}}_{ev}\mathbf{y}\\
\boldsymbol{\alpha} &= \overbrace{\boldsymbol{\Phi}^{\mathrm{SF}}_{xw}(\boldsymbol{\beta} - (zI - A)\boldsymbol{\Phi}^{\mathrm{KF}}_{ew}\boldsymbol{\beta})}^{\tilde{\mathbf{w}}} + \overbrace{\boldsymbol{\Phi}^{\mathrm{KF}}_{ew}\boldsymbol{\beta}}^{\hat{\mathbf{w}}},\\
\boldsymbol{\gamma} &= \boldsymbol{\Phi}^{\mathrm{SF}}_{uw}(\boldsymbol{\beta} - (zI - A)\boldsymbol{\Phi}^{\mathrm{KF}}_{ew}\boldsymbol{\beta}).
\end{aligned}
\tag{20}
$$

With these intermediate signals, we can compute the controller internal state $\boldsymbol{\beta}$ and the control signal $\mathbf{u}$ in Fig. 1 from $z\boldsymbol{\beta} = -z(A\boldsymbol{\alpha} + B_2\boldsymbol{\gamma}) - z\boldsymbol{\zeta}$, and $\mathbf{u} = z\boldsymbol{\gamma} + \boldsymbol{\theta}$.

Locally, due to the communication constraints specified in Section 2.2, one can not carry out the computation described above in a centralized way. In particular, the local computation of each signal in (20) involves delayed and locally available information. We now describe the information exchange among subsystems and how they compute (20). Recall that the state-feedback solution $\boldsymbol{\Phi}^{\mathrm{SF}}$ and Kalman-filters $\boldsymbol{\Phi}^{\mathrm{KF}}$ from (15) are synthesized to respect the communication constraints expressed as d-Delayed localization SLCs. Denote

$$
\boldsymbol{\Phi}^{\mathrm{SF}}_{w}(:,i) = \left[\begin{array}{c|c} A^i_{\mathrm{SF}} & B^i_{\mathrm{SF}} \\ \hline I & 0 \end{array}\right]. \tag{21}
$$

Then $\boldsymbol{\Phi}^{\mathrm{SF}}_{xw}(:,i) = C^i_{\mathrm{SF}}\boldsymbol{\Phi}^{\mathrm{SF}}_{w}(:,i)$ and $\boldsymbol{\Phi}^{\mathrm{SF}}_{uw}(:,i) = K^i_{\mathrm{SF}}\boldsymbol{\Phi}^{\mathrm{SF}}_{w}(:,i)$ where C^i_{SF} and K^i_{SF} are from (16). Computing the local components of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ requires only one realization of $\boldsymbol{\Phi}^{\mathrm{SF}}_{w}$ as they can share the same copy of the states within each subsystem. An analogous statement holds true for $\boldsymbol{\zeta}$ and $\boldsymbol{\theta}$. Denote the two realizations of (21) as $\boldsymbol{\Phi}^{\mathrm{SF}}_{w,\alpha}$ and $\boldsymbol{\Phi}^{\mathrm{SF}}_{w,\zeta}$. During each time step t, every node observes its local output $y_i(t)$ and goes through four stages of computation and communication with its neighbors leading to an update to the internal controller states and the application of the actuator signal $u_i(t)$. This is summarized in Algorithm 2 with subroutines 1–4 describing these computations in detail.

Control signal computation at subsystem i begins by receiving the measurements from neighbors j at most d steps away (line 8) and computing the ith element of the internal signals $\hat{v}(t+1)$ and $\hat{w}(t+1)$ via Subroutine 1, which is illustrated in Fig. 1. Here the function step(G, u) means that the internal dynamics of the system G is propagated one time-step with the input u.

In the second stage, the node receives $\hat{v}_j(t)$ and $\hat{w}_j(t)$ from its closest neighbors (line 10) and computes the outgoing components of (20). The computations are outlined in Subroutine 2 and illustrated in Fig. 4.

In the third stage, which is demonstrated in Fig. 5, the node receives the components pertaining to its element of the signals in (19) from other nodes a distance at most d steps away with delayed information (line 12) and sums them to compute the ith element of each signal in (20). This step is described in Subroutine 3.

Algorithm 2 Local computation of controller signals

for Each node $i = 1, \dots, N$ **do**
  Input $\mathbf{\Phi}^{\mathrm{SF}}_{w,\alpha}(:, i)$, $\mathbf{\Phi}^{\mathrm{SF}}_{w,\zeta}(:, i)$, $\mathbf{\Phi}^{\mathrm{KF}}_{ew}(:, i)$, $\mathbf{\Phi}^{\mathrm{KF}}_{ev}(:, i)$
  Initialize $\beta_i(0) \leftarrow 0$, $w_i(t) \leftarrow 0$, $v_i(t) \leftarrow 0$
end for
for $t = 0, 1, \dots$ **do**
  for Each node $i = \{1, \dots, N\}$ **do** ▷ parallel
    Observe $y_i(t)$
    Receive $\beta_j(t)$ and $y_j(t)$ from $j \in \mathcal{N}^d_{\mathrm{in}}(i)$
    subroutine1()
    Receive $\widehat{w}_j(t)$ and $\widehat{v}_j(t)$ from $j \in \mathcal{N}^1_{\mathrm{in}}(i)$
    subroutine2()
    Receive $\widehat{\alpha}^{(\mathcal{N}^d_{\mathrm{out}}(j))}(t+1)$, $\widehat{\zeta}^{(\mathcal{N}^d_{\mathrm{out}}(j))}(t+1)$, $\widehat{\gamma}^{(\mathcal{N}^d_{\mathrm{out}}(j))}(t+1)$ and $\widehat{\theta}^{(\mathcal{N}^d_{\mathrm{out}}(j))}(t)$ from $j \in \mathcal{N}^d_{\mathrm{in}}(i)$.
    subroutine3()
    Receive $\alpha_j(t+1)$ and $\gamma_j(t+1)$ from $j \in \mathcal{N}^1_{\mathrm{in}}(i)$
    subroutine4()
    Apply $u_i(t)$
  end for
end for

Subroutine 1 Compute $\widehat{w}_i(t+1)$ and $\widehat{v}_i(t+1)$

Receive $\beta_j(t)$ and $y_j(t)$ from $j \in \mathcal{N}^d_{\mathrm{in}}(i)$
$\beta_{[\mathcal{N}^d_{\mathrm{in}}(i)]}(t) \leftarrow \mathrm{vec}(\beta_{j_1}(t), \dots, \beta_{j_m}(t))$
$y_{[\mathcal{N}^d_{\mathrm{in}}(i)]}(t) \leftarrow \mathrm{vec}(y_{j_1}(t), \dots, y_{j_m}(t))$
$\widehat{w}_i(t+1) \leftarrow \mathrm{step}(z\mathbf{\Phi}^{\mathrm{KF}}_{ew}(i, :), \beta_{[\mathcal{N}^d_{\mathrm{in}}(i)]}(t))$
$\widehat{v}_i(t+1) \leftarrow \mathrm{step}(z\mathbf{\Phi}^{\mathrm{KF}}_{ev}(i, :), y_{[\mathcal{N}^d_{\mathrm{in}}(i)]}(t))$

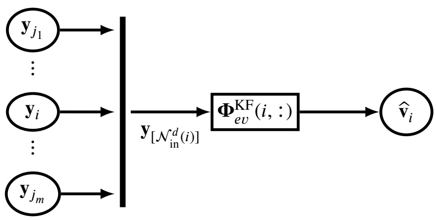


Figure 3. Illustration of subroutine 1. The computation of $\widehat{w}$ is similar.

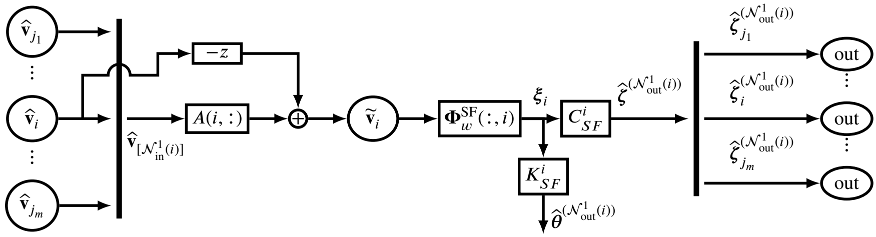


Figure 4. Illustration of subroutine 2.

Subroutine 2 Compute the outgoing components of (20)

$\widehat{w}_{[\mathcal{N}^1_{\text{in}}(i)]}(t) \leftarrow \text{vec}(\widehat{w}_{j_1}(t), \dots, \widehat{w}_{j_m}(t))$
$\widehat{v}_{[\mathcal{N}^1_{\text{in}}(i)]}(t) \leftarrow \text{vec}(\widehat{v}_{j_1}(t), \dots, \widehat{v}_{j_m}(t))$
$\widetilde{w}_i(t) = \beta_i(t) + A\widehat{w}_{[\mathcal{N}^1_{\text{in}}(i)]}(t) - \widehat{w}_i(t+1)$
$\widetilde{v}_i(t) = A\widehat{v}_{[\mathcal{N}^1_{\text{in}}(i)]}(t) - \widehat{v}_i(t+1)$
$\lambda_i(t+1) \leftarrow \text{step}(\mathbf{\Phi}^{\text{SF}}_{w,\alpha}(:,i), e_i\widetilde{w}_i(t))$
$\xi_i(t+1) \leftarrow \text{step}(\mathbf{\Phi}^{\text{SF}}_{w,\zeta}(:,i), e_i\widetilde{v}_i(t))$
$\widehat{\alpha}^{(\mathcal{N}^d_{\text{out}}(j))}(t+1) \leftarrow C^i_{SF}\lambda_i(t+1)$
$\widehat{\zeta}^{(\mathcal{N}^d_{\text{out}}(j))}(t+1) \leftarrow C^i_{SF}\xi_i(t+1)$
$\widehat{\gamma}^{(\mathcal{N}^d_{\text{out}}(j))}(t+1) \leftarrow K^i_{SF}\lambda_i(t+1)$
$\widehat{\theta}^{(\mathcal{N}^d_{\text{out}}(j))}(t) \leftarrow K^i_{SF}\xi_i(t+1)$

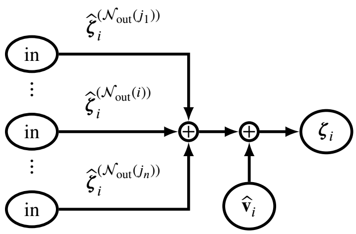


Figure 5. Illustration of subroutine 3.

Subroutine 3 Compute the local elements of (20)

$\alpha_i(t+1) \leftarrow \widehat{w}_i(t+1) + \sum_j \widehat{\alpha}_i^{(\mathcal{N}^1_{\text{out}}(j))}(t+1)$

$\gamma_i(t+1) \leftarrow \sum_j \widehat{\gamma}_i^{(\mathcal{N}^1_{\text{out}}(j))}(t+1)$

$\zeta_i(t+1) \leftarrow \widehat{v}_i(t+1) + \sum_j \widehat{\zeta}_i^{(\mathcal{N}^1_{\text{out}}(j))}(t+1)$

$\theta_i(t) \leftarrow \sum_j \widehat{\theta}_i^{(\mathcal{N}^1_{\text{out}}(j))}(t)$

In the final stage (Subroutine 4) the node receives $\alpha_j(t+1)$ and $\gamma_j(t+1)$ from its closest neighbors (line 14) and computes $\beta_i(t+1)$ and $u_i(t)$. We conclude that the node has now received information from nodes at most $2d+2$ steps away.

Subroutine 4 Compute $u_i(t)$ and $\beta_i(t+1)$

$\alpha_{[\mathcal{N}^1_{\text{in}}(i)]}(t+1) \leftarrow \text{vec}\left(\alpha_{j_1}(t+1), \dots, \widehat{\alpha}_{j_m}(t+1)\right)$

$\gamma_{[\mathcal{N}^1_{\text{in}}(i)]}(t+1) \leftarrow \text{vec}\left(\gamma_{j_1}(t+1), \dots, \widehat{\gamma}_{j_m}(t+1)\right)$

$$\beta_i(t+1) \leftarrow -A(i,:)\alpha_{[\mathcal{N}^1_{\text{in}}(i)]}(t+1) - B(i,:)\gamma_{[\mathcal{N}^1_{\text{in}}(i)]}(t+1) - \zeta_i(t+1)$$

$u_i(t) \leftarrow \gamma_i(t+1) + \theta_i(t)$

We summarize the stability properties of Algorithm 2 in the following theorem:

THEOREM 18
Algorithm 2 with $\mathbf{\Phi}_w^{\text{SF}}$ as in (21), $\mathbf{\Phi}_{ew}^{\text{KF}}$ and $\mathbf{\Phi}_{ev}^{\text{KF}}$ as in Section 3 internally stabilizes system (1). Moreover if $\mathbf{\Phi}_w^{\text{SF}}$, $\mathbf{\Phi}_{ew}^{\text{KF}}$ and $\mathbf{\Phi}_{ev}^{\text{KF}}$ are d-localized, the closed-loop is at most $2d+2$-localized. □

Proof. By Theorem 17 the closed-loop maps satisfy (2a) and (2b). Concatenating $\mathbf{y}_i$, $\boldsymbol{\beta}_i$ and $\mathbf{u}_i$ we get precisely the signals in Fig. 1 which is internally stable [Anderson et al., 2019], we need to show that the closed-loop is internally stable for perturbations entering in the intermediate steps outlined in Subroutines 1–4. Note that a perturbation entering at any of the intermediate signals can be modeled as a disturbance entering as $\boldsymbol{\delta}_x$, $\boldsymbol{\delta}_y$ or $\boldsymbol{\delta}_\beta$ pre-filtered through a stable linear system. Similarly, probing any of the internal signals can be represented as probing $\mathbf{y}$, $\mathbf{u}$ or $\boldsymbol{\beta}$ post-filtered through a stable system. We conclude Algorithm 2 is internally stable in feedback with System 1. Finally, as d-localization is closed under addition, and composition of a d- and a k-localized operator is at most $d+k$-localized, (19) implies that the closed loop is at most $2d+2$-localized. □

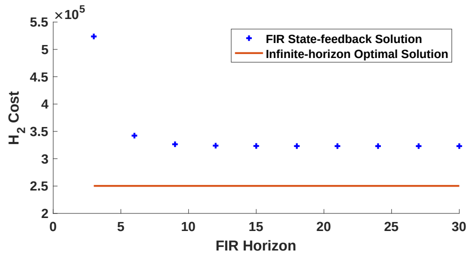


Figure 6. The infinite-horizon SLS solution achieves the optimal cost.

6. Numerical simulations

Consider a bi-directional scalar chain network parameterized by α and ρ:

$$x^i[t+1] = \rho(1-2\alpha)x^i[t] + \rho\alpha \sum_{j\in\{i\pm1\}} x^j[t] + u^i[t] + w^i[t]$$

where α is a coupling constant and ρ is the spectral radius of the global state-transition matrix A, with $\rho \geq 1$ being unstable. We first verify the optimality of the infinite-horizon state-feedback solution given in Section 3. In this simulation, we choose the number of scalar subsystems to be 15, $\alpha = 0.6$ and $\rho = 1$. For the quadratic cost matrices, we let $Q = I$ and $R = 300 \cdot I$. For SLCs, we let the delayed localization parameter be $d = 3$. The result is shown in Figure 6, where the optimality of our approach is clear. Due to the high penalty on the control actions, the performance degradation under FIR approximation can be significant.

Next, we investigate the optimality gap between the suboptimal infinite-horizon output-feedback solution proposed in this work against the FIR output-feedback solution computed numerically with a fixed FIR horizon of 20. We let all other parameters remain the same as before, and change the number of subsystems to 10 in this simulation. First, we study how the d delayed localization parameter influence the optimality gap. This is illustrated in Figure 7. As expected, the more localized the output-feedback problem is, the bigger the optimality gap is between the constructed solution using separation principle and the direct FIR output-feedback solution. As the delayed localization pattern becomes more global, the proposed output-feedback solution becomes more optimal. When the delayed localization SLCs become non-binding (for $d \geq 6$), we see that the proposed infinite-horizon output-feedback solution actually becomes optimal and achieves lower cost than the FIR solution. This is due to the separation principle of centralized LQG.

Next, we investigate how the optimality gap grows with the number of subsystems in the network. Here we have fixed the delayed localization parameter to be $d = 3$. As can be seen in Figure 8, we observe that the optimality gap grows apparently linearly in the number of subsystems. However, we highlight the numerical efficiency and stability of our approach despite the suboptimality. When the number of subsystems exceeds 12 with FIR horizon of 20, the FIR solution solved in MAT-

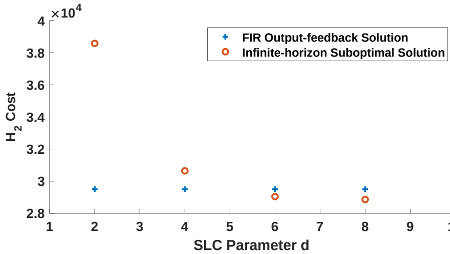


Figure 7. The proposed infinite-horizon suboptimal solution to the output-feedback SLS problem versus the FIR output-feedback solution numerically computed for (OF-SLS) for varying SLC delayed localization parameter d.

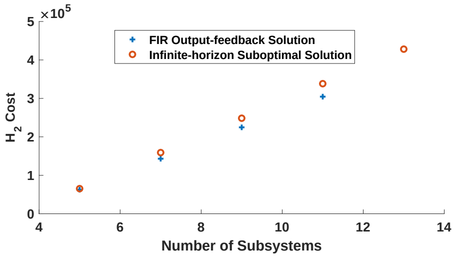


Figure 8. The proposed suboptimal solution to the output-feedback SLS problem versus the FIR output-feedback solution numerically computed for (OF-SLS) for varying number of subsystems in the network.

LAB using CVX renders NaN due to numerical instability (total of 11520 variables).

References

Anderson, J., J. C. Doyle, S. H. Low, and N. Matni (2019). “System level synthesis”. *Annual Reviews in Control* **47**, pp. 364–393.

Feyzmahdavian, H. R., A. Alam, and A. Gattami (2012). “Optimal distributed controller design with communication delays: application to vehicle formations”. In: *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, pp. 2232–2237.

Fisher, M. W., G. Hug, and F. Dörfler (2022). “System level synthesis beyond finite impulse response using approximation by simple poles”. *arXiv preprint arXiv:2203.16765*.

Furieri, L., Y. Zheng, A. Papachristodoulou, and M. Kamgarpour (2019). “An input–output parametrization of stabilizing controllers: amidst youla and system level synthesis”. *IEEE Control Systems Letters* **3**:4, pp. 1014–1019.

Kalman, R. E. et al. (1960). “Contributions to the theory of optimal control”. *Bol. soc. mat. mexicana* **5**:2, pp. 102–119.

Lessard, L. and S. Lall (2012). “Optimal controller synthesis for the decentralized two-player problem with output feedback”. In: *2012 American Control Conference (ACC)*. IEEE, pp. 6314–6321.

Rotkowitz, M. and S. Lall (2005). “A characterization of convex problems in decentralized control”. *IEEE transactions on Automatic Control* **50**:12, pp. 1984–1996.

Sab**u***a*u, Ş., A. Speril**u***a*, C. Oar**u***a*, and A. Jadbabaie (2021). “Network realization functions for optimal distributed control”. *arXiv preprint arXiv:2112.09093*.

Wang, Y.-S., N. Matni, and J. C. Doyle (2014). “Localized LQR optimal control”. In: *53rd IEEE Conference on Decision and Control*. IEEE, pp. 1661–1668.

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

Wang, Y.-S., N. Matni, and J. C. Doyle (2019). “A system-level approach to controller synthesis”. *IEEE Transactions on Automatic Control* **64**:10, pp. 4079–4093.

Wang, Y.-S., S. You, and N. Matni (2015). “Localized distributed Kalman filters for large-scale systems”. *IFAC-PapersOnLine* **48**:22, pp. 52–57.

Wonham, W. M. (1968). “On the separation theorem of stochastic control”. *SIAM Journal on Control* **6**:2, pp. 312–326.

Yu, J., Y.-S. Wang, and J. Anderson (2021). “Localized and distributed $\mathcal{H}_2$ state feedback control”. In: *2021 American Control Conference (ACC)*. IEEE, pp. 2732–2738.

Zheng, Y., L. Furieri, M. Kamgarpour, and N. Li (2022). “System-level, input–output and new parameterizations of stabilizing controllers, and their numerical computation”. *Automatica* **140**, p. 110211.

Paper VI

Learning Optimal Team-Decisions

Abstract

In this paper, we linear quadratic team decision problems, where a team of agents minimizes a convex quadratic cost function over T time steps subject to possibly distinct linear measurements of the state of nature. We assume that the state of nature is a Gaussian random variable and that the agents do not know the cost function nor the linear functions mapping the state of nature to their measurements. We present a gradient-descent based algorithm with an expected regret of $O(\log(T))$ for full information gradient feedback and $O(\sqrt{(}T))$ for bandit feedback. In the case of bandit feedback, the expected regret has an additional multiplicative term $O(d)$ where d reflects the number of learned parameters.

1. Introduction

Team decision problems originate from economics, where optimal decentralized decisions in organizations were studied in the papers by Marschak [Marschak, 1955], and Radner [Radner, 1962] under stochastic settings. In these studies, the agents in the team *know the problem parameters*. The agents use the information of the problem parameters to find the optimal decentralized decision. Decentralized decisions only depend on local measurements of the state of nature, where the measurements of the agents are typically not identical. Gattami [Gattami et al., 2012] studied linear quadratic robust team decision problems and showed that optimal decisions are linear and can be found by solving a convex (in fact, semi-definite) optimization problem. Team-decision theory has been helpful in understanding distributed control research [Mahajan et al., 2012]. Witsenhausens famous counterexample [Witsenhausen, 1968] established that linear decisions are not always optimal for distributed LQG problems and sparked an interest into research of team problems in the control community. Ho and Chu [Ho and Chu, 1972] showed how linear-quadratic problems with partially nested information can be rewritten as static team-decision problems of the type in this paper and Witsenhausen showed that a general class of dynamic team decision problems can be reduced to static ones via a change of measures [Witsenhausen, 1988]. Static reductions for more exotic information structures is still an active research field [Gupta et al., 2014; Sanjari et al., 2021].

In this article, we study *learning* of optimal decentralized decisions with linear information constraints and quadratic cost functions in the stochastic setting, *without the knowledge of the problem parameters*. We consider learning with gradient feedback and bandit feedback. We study expected regret against the optimal policy in hindsight. Our key contributions are:

- We propose a first and a zeroth-order algorithm to learn optimal decentralized decisions with linear information constraints and quadratic cost functions through repeated interactions.
- We extend the regret analysis of online gradient descent to the case with a possibly unbounded gradient oracle that has bounded second-moment.
- We show that our algorithms have expected regret bounded by $O(log(T))$ if the gradient is observed and $O(\sqrt{T})$ if only the incurred loss is observed in each step.

1.1 Outline

We give some background and establish notation in Section 2 and formalize the learning problem in Section 3. Section 4 is devoted to properties of decentralized stochastic team-decision problems. Section 5 contains our extension to the regret analysis of online gradient descent and its application to the stochastic team-decision problem. In Section 7 we summarize our conclusions and give directions for future research.

2. Preliminaries

2.1 Notation

We denote the space of n-dimensional real-valued vectors by $\mathbb{R}^n$ and real-valued matrices with m rows and n columns by $\mathbb{R}^{m\times n}$. For a vector $x \in \mathbb{R}^n$, $\|x\|_2 = \sqrt{x^\top x}$ denotes the Euclidean norm and $A^\top$ denotes the transpose of a matrix A. Tr M denotes the trace of a square matrix M. For matrices $A, B \in \mathbb{R}^{m\times n}$, we denote the operator norm of A as $\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2$, the Frobenius inner product as $\langle A, B\rangle_F = \text{Tr}\, A^\top B$, and the Frobenius norm as $\|A\|_F = \sqrt{\langle A, A\rangle_F}$. We denote the smallest singular value of a matrix $A \in \mathbb{R}^{m\times n}$ by $\sigma_{\min}(A)$. The set of real-valued $n\times n$-dimensional symmetric matrices is denoted as $\mathbb{S}^n$. $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$ refer to the sets of $n\times n$-dimensional of positive semi-definite and positive definite matrices, respectively. For a matrix $A \in \mathbb{R}^{(m_1+\cdots+m_M)\times(n_1+\cdots+n_N)}$, $[A]_i \in \mathbb{R}^{m_i\times(n_1+\cdots+n_N)}$ denotes the ith block row and $[A]_{ij} \in \mathbb{R}^{m_i\times n_j}$ denotes the block element of A in position (i, j). The matrix derivative of a differentiable function $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is denoted $\frac{\partial}{\partial X} f(X)$, where $\left[\frac{\partial}{\partial X} f(X)\right]_{ij} = \partial f(X)/\partial X_{i,j}$. The projection of a variable $y \in \mathcal{Y}$

Algorithm 3 Online Gradient Descent

Input: Convex set $\mathcal{X}$, T, $x_1 \in \mathcal{X}$ step-sizes $\{\eta_t\}$
for $t = 1$ **to** T **do**
  Play x_t, observe f_t and pay $f_t(x_t))$
  Update and project $x_{t+1} = \prod_{\mathcal{X}}(x_t - \nabla f_t(x_t))$
end for

onto a set $\mathcal{X} \subseteq \mathcal{Y}$ is denoted by $\prod_{\mathcal{X}}(y)$. $\mathcal{L}_2$ means the space of square-integrable random variables with the associated inner product $\langle \cdot, \cdot \rangle_{\mathcal{L}_2}$, and semi-norm $\| \cdot \|_{\mathcal{L}_2}$. The set of Gaussian variables with mean m and covariance Σ is denoted $\mathcal{N}(m, \Sigma)$ and $\mathbb{E}[\cdot]$ denotes the expectation operator.

2.2 Online Convex Optimization

The online convex optimization setting is a repeated leader-follower game between a minimizing player and an adversary. At each time-step t, the minimizing player first makes a decision x_t from some compact convex set $\mathcal{X}$. The adversary then observes x_t and selects a convex loss function f_t that is uniformly bounded and has bounded gradients. The minimizing player pays $f_t(x_t)$ and learns the entire function f_t. The goal is to minimize the sum, $\sum_{t=1}^{T} f_t(x_t)$ over an arbitrary sequence of loss functions $f_1, f_2, \dots, f_t$.

Recently, online convex optimization has seen an increasing number of applications across different fields including generator scheduling in smart grids [Narayanaswamy et al., 2012], thermal management of multiprocessors [Zanini et al., 2010], demand steering via real-time electricity pricing [Kim and Giannakis, 2014] and on-policy learning of optimal control policies with linear dynamics [Li et al., 2021; Chen and Hazan, 2021; Hazan et al., 2020; Cohen et al., 2018].

The performance measure is regret against the optimal policy in hindsight,

$$R(T) = \sum_{t=1}^{T} f_t(x_t) - \min_{x \in \mathcal{X}} \sum_{t=1}^{T} f_t(x).$$

Online gradient descent, introduced by Zinkevich [Zinkevich, 2003], is a simple, general yet efficient algorithm that applies to many online convex optimization problems and is given in Algorithm 3. Online gradient descent attains the asymptotic lower bounds $\Omega(DG\sqrt{T})$ and $O(\log T)$ for convex functions and α-strongly convex functions respectively. D bounds the diameter of the feasible set, and G bounds the norm of the gradient.

We will be working with matrix-valued variables and strongly convex functions for the team decision problem, using the below definition of strong convexity.

DEFINITION 5—STRONG CONVEXITY, MATRIX CASE
We say that the differentiable function $f : \mathcal{X} \subseteq \mathbb{R}^{m \times p} \to \mathbb{R}$ is *strongly convex* with *coefficient* α if for all $X, Y \in \mathcal{X}$,

$$f(X) - f(Y) \leq \left\langle \frac{\partial}{\partial X} f(X), (X - Y) \right\rangle_F - \frac{\alpha}{2} \|x - y\|_F^2. \qquad \square$$

An equivalent characterization is to require that the function $X \mapsto f(X) - \frac{\alpha}{2}\|X\|_F^2$ is convex, [Bubeck, 2015]. We refer the reader to [Hazan, 2019] for more details on online convex optimization.

2.3 Bandit and Zeroth-Order Optimization

The minimizing player observes only the incurred cost $f_t(x_t)$ after each round in the bandit setting, rather than the gradient. This necessitates *exploration* to learn properties of the loss functions, such as gradients, to accelerate optimization. Derivative-free methods have a long history in stochastic optimization. Tight convergence rates for strongly convex functions were obtained in [Rakhlin et al., 2012] in the first- and [Shamir, 2013] in the zeroth-order setting. Bandit feedback was introduced to the online convex optimization setting in [Flaxman et al., 2004] where the authors used a one-point gradient estimate. Their method has asymptotic regret upper bounded by $O(T^{3/4})$.

2.4 Stochastic Team Decision theory

The stochastic team-decision problem, is to solve

$$\begin{aligned} \underset{\mu}{\text{minimize}} \quad & \mathbb{E}[\|z\|_2^2] \\ \text{subject to:} \quad & z = Hx + Du \\ & y_i = C_i x + v_i \\ & u_i = \mu_i(y_i), \quad i = 1, \dots, N. \end{aligned} \tag{1}$$

In (1), $x \sim \mathcal{N}(0, V_{xx})$ and $v \sim \mathcal{N}(0, V_{vv})$ are independent Gaussian variables taking values in $\mathbb{R}^n$ and $\mathbb{R}^p$ respectively. $u_i \in \mathbb{R}^{m_i}$ denotes a player, and the players $u_1, \dots, u_N$ make up a team. The function $\mu(\cdot) : \mathbb{R}^p \to \mathbb{R}^m$ represents the decision function of the team, that is, $\mu(Cx) = \begin{bmatrix} \mu_1(y_1)^\top & \cdots & \mu_1(y_N)^\top \end{bmatrix}^\top$. We further assume that $D^\top D \in \mathbb{S}_{++}^m$ where $m = m_1 + \cdots + m_N$. Radner [Radner, 1962] showed that the optimal decision functions $\mu_i^\star$ are unique and linear in y_i. This motivates the search over linear policies in our problem set-up.

3. Problem Formulation

We aim to learn the optimal decision policy through repeated interactions with the environment. At each time-step t, each team-member will decide on a decision policy

K_i^t, receive a noisy partial observation of the system state, y_i^t, play the decision $u_i^t = K_i^t y_i^t$. The team incurs the loss $l_t(K_t) = \|z_t\|_2^2$, generated by

$$\begin{aligned} z_t &= Hx_t + Du_t, & y_i^t &= C_i^t x + v_i^t \\ u_i^t &= K_i^t y_i^t, & i &= 1, \dots, N. \end{aligned} \tag{2}$$

The objective is to minimize the sum of the losses, $J = \sum_{t=1}^{T} l_t(K_t)$, while maintaining $K_t \in \mathcal{K}$, learning good policies locally. $\mathcal{K}$ is the set of real-valued block-diagonal matrices of appropriate dimensions,

$$\mathcal{K} := \{K : K = \mathrm{Diag}(K_1, \dots, K_N), K_i \in \mathbb{R}^{m_i \times p_i}\}. \tag{3}$$

We summarize the interaction in Algorithm 4. Going forward we make the following assumptions.

Assumption 1
x_t and v_t have finite covariance matrices $\mathbb{E}[x_t x_t^\top] = V_{xx}$ and $\mathbb{E}[v_t v_t^\top] = V_{vv}$ and bounded fourth order moments so that $\mathbb{E}[(x_t^\top x_t)^2] \leq \kappa_x$ and $\mathbb{E}[(v_t^\top v_t)^2] \leq \kappa_v$. □

Assumption 2

$$\sigma_{\min}(D^\top D)(\sigma_{\min}(CV_{xx}C^\top) + \sigma_{\min}(V_{vv})) > 0.$$ □

Assumption 1 is motivated by the fact that the variance of an estimator of the derivative $\frac{\partial}{\partial K} J(K)$ will contain fourth-order moments. Assumption 2 is to the losses being strongly convex in expectation, which is summarized in Proposition 7. Finally, we restrict our search to policies with an apriori supplied bound.

Assumption 3
A bound b_K on $\|K\|_2$ is supplied by an oracle. □

Let $K^\star$ be the best policy in hindsight,

$$K^\star = \underset{K \in \mathcal{K}, \|K\|_2 \leq b_K}{\arg\min} \sum_{t=1}^{T} l_t(K). \tag{4}$$

We measure performance as expected regret,

$$\mathbb{E}[R(T)] = \mathbb{E}\left[\sum_{t=1}^{T} l_t(K_t) - \sum_{t=1}^{T} l_t(K^\star)\right]. \tag{5}$$

Algorithm 4 Learning with repeated interactions.

for $t = 1$ **to** T **do**

Sample $x_t \sim \mathcal{N}(0, V_{xx})$ and $v_t \sim \mathcal{N}(0, V_{vv})$

Agents $1, 2, ..., N$ observes $y_1^t, \dots, y_N^t$ as in (2), respectively.

The agents play $K_1^t, \dots, K_N^t$, respectively, and incur a loss $l_t(K_t) := \|z_t\|_2^2$, with z_t as in (2).

Each agent i observes either

- the partial derivative, $\frac{\partial}{\partial K_i}(z_t)^\top z_t$, in the gradient-feedback setting,
- or the incurred loss, $\|z_t\|_2^2$, in the bandit-feedback setting

The agents update their policies K_i^{t+1}.

end for

4. Properties of Stochastic Team Decisions

The losses l_t are differentiable with respect to K_t everywhere. In particular, the derivative with respect to agent i can be viewed as a product of the information available to the agent y_i, and their contribution to the overall cost, $[D^\top]_i z$.

PROPOSITION 6
l_t is differentiable with respect to K_i and the derivative is

$$\frac{\partial}{\partial K_i}\mathbb{E}[l_t(K)] = \mathbb{E}[\left[2[D^\top]_i z_t (y_i^t)^\top\right].$$

□

Proof. By dominated convergence, we can exchange expectation and differentiation.[1]

$$\begin{aligned}\frac{\partial}{\partial K}\mathbb{E}[z^\top z] &= \mathbb{E}\left[\frac{\partial}{\partial K}(Hx + DKy)^\top(Hx + DKy)\right] \\ &= \mathbb{E}\left[2D^\top z y^\top\right].\end{aligned}$$

Identifying the local components $\partial/\partial K_i$ completes the proof. □

The phenomenon that certain large changes to the optimization variable can have (almost) negligible effects on the value can make optimization difficult. The right way to quantify this effect on convergence is through *strong convexity*, a property we can exploit to get better regret bounds in online convex optimization [Hazan, 2019]. In our regret terms, a lower bound on the strong convexity parameter will show up directly as a divisor. The following proposition shows that $\mathbb{E}[l_t]$ is strongly convex as a function of K.

[1] We drop the time-index for readability

PROPOSITION 7
$\mathbb{E}[l_t]$ is α-strongly convex with constant

$$\alpha = 2\sigma_{\min}(D^\top D)(\sigma_{\min}(CV_{xx}C^\top) + \sigma_{\min}(V_{vv})). \qquad \square$$

Proof. We will verify that $\mathbb{E}[l_t](K) - \frac{\alpha}{2}\|K\|_F^2$ is convex. $\mathbb{E}[l_t]$ is a quadratic function of K and

$$\mathbb{E}[l_t](K) = \|Hx\|_{\mathcal{L}_2}^2 + 2\langle Hx, DKy\rangle_{\mathcal{L}_2} + \|DKy\|_{\mathcal{L}_2}^2.$$

Which is convex if and only if $\|DKy\|_{\mathcal{L}_2}^2 \geq \frac{\alpha}{2}\|K\|_F^2$. Consider,

$$\|DKy\|_{\mathcal{L}_2}^2 \geq \sigma_{\min}(D^\top D)\|KCx + Kv\|_{\mathcal{L}_2}^2 \geq \frac{\alpha}{2}\|K\|_F^2 \qquad \square$$

To apply online optimization algorithms to learn the optimal policy through repeated play, we must bound the second and fourth moments of z as we must bound the variance of our derivative estimates. We get the following bounds on the second and fourth order moments of z by Assumptions 1 and 3.

PROPOSITION 8
For $\|K\|_2 \leq b_K$, the loss $l_t(K)$ in Algorithm 4 is bounded from above in expectation, $\mathbb{E}[l_t](K) \leq b_l$, where

$$b_l = (\|H\|_2 + \|D\|_2\|C\|_2 b_K)^2 \operatorname{Tr} V_{xx} + \|D\|_2^2 b_K^2 \operatorname{Tr} V_{vv}.$$

Furthermore, $\mathbb{E}[(z_t^\top z_t)^2 \leq \kappa_z]$ where

$$\kappa_z = \left(\|H\|_2 + \|D\|_2\|C\|_2 b_K + \|D\|_2 b_K\right)^4 \times \left(\kappa_x + \operatorname{Tr} V_{xx} \operatorname{Tr} V_{vv} + \kappa_v\right) \tag{6}$$

$\square$

Proof Proof of Proposition 8. We start with bounding the value function. Let $\|\cdot\|_{\mathcal{L}_2}$ be the $\mathcal{L}_2$ norm, then

$$\begin{aligned}\mathbb{E}[l](K) &= \|(H + DKC)x + DKv\|_{L_2}^2 \\ &= \|(H + DKC)x\|_{L_2}^2 + \|DKv\|_{L_2}^2,\end{aligned}$$

as x and v are independent. By the triangle inequality

$$\|(H + DKC)x\|_{L_2}^2 \leq (\|H\|_2 + \|D\|_2\|K\|_2\|C\|_2)^2\|x\|_{\mathcal{L}_2}^2.$$

Treating the term $\|DKv\|_{\mathcal{L}_2}$ similarly and substituting $\|K\|_2 \leq b_K$ and $\|x\|^2_{\mathcal{L}_2} = \operatorname{Tr} V_{xx}$ completes the proof. To prove the second claim, consider

$$\begin{aligned}
\mathbb{E}[(z^\top z)^2] &= \mathbb{E}\left[\left\|\begin{bmatrix} H + DKC & DK \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}\right\|_2^4\right] \\
&\leq \left\|\begin{bmatrix} H + DKC & DK \end{bmatrix}\right\|_2^4 \mathbb{E}[(x^\top x + v^\top v)^2] \\
&\leq \left(\|H\|_2 + \|D\|_2\|C\|_2 b_K + \|D\|_2 b_K\right)^4 \\
&\quad \times \left(\kappa_x + \operatorname{Tr} V_{xx} \operatorname{Tr} V_{vv} + \kappa_v\right)
\end{aligned}$$

□

5. Learning Optimal Team Decisions

This section describes how to learn the optimal team decision policies using online gradient descent. Due to the stochastic nature of our problem, we cannot hope to bound the objective function or the gradient for an arbitrary realization. We will modify the analysis to give results when these properties hold in expectation. This means our guarantees hold in expectation and are well suited to analyze stochastic problems. We summarize the upper bound for expected regret for strongly convex functions in Theorem 19 The bound is what one would expect; the standard result [Hazan, 2019, Theorem 3.3] for strongly convex functions holds in expectation against an adaptive adversary.

THEOREM 19
Let $l_1, \dots, l_T$ be independent random functions $l_t : R^{m\times n} \to \mathbb{R}$ such that $\mathbb{E}[l_t]$ is α-strongly convex for all $t = 1, \dots, T$. Let $\tilde{\nabla}_t$ be a derivative oracle that is **consistent** $\mathbb{E}[\tilde{\nabla}_t] = \frac{\partial}{\partial K}\mathbb{E}[l_t(K)]$ and has **bounded variance** $\mathbb{E}\left[\|\tilde{\nabla}_t\|_F^2\right] \leq (b_t)^2$ for all $K \in \mathcal{K}$, where $\mathcal{K}$ is convex and compact. Set the step size $\eta_t = \frac{1}{\alpha t}$. Let $K^\star = \arg\min_{K\in\mathcal{K}} \sum_{t=1}^T l_t(K)$. Online Gradient Descent, Algorithm 3, has expected regret

$$\mathbb{E}\left[\sum_{t=1}^{T}\left(l_t(K_t) - l_t(K^\star)\right)\right] \leq \frac{1}{2}\sum_{t=1}^{T}\frac{b_t^2}{\alpha t}. \tag{7}$$

□

The proof follows the outline in [Hazan, 2019], but involves some extra bookkeeping:

Proof. Let $\mathcal{F}_t = \sigma(J_1, \dots, J_{t-1})$. Then K_t is a stochastic sequence adapted to $\mathcal{F}_t$. Define for simplicity $\nabla_t = \frac{\partial}{\partial K}\mathbb{E}[l_t(K_t)]$. By strong convexity

$$2\mathbb{E}\left[l_t(K_t) - l_t(K^\star)|\mathcal{F}_t\right] \leq 2\langle \nabla_t, K_t - K^\star\rangle_F - \alpha\|K^\star - K_t\|_F^2.$$

Algorithm 5 Learning with partial gradient information

Input: initial guess K_0, bound b_K, step-sizes $\{\eta_t\}$
Each agent plays $u_i^t = K_i^t y_i^t$
The team incurs cost $l_t(K_t) = z_t^\top z_t$
for $t = 0$ **to** $T - 1$ **do**
  for $i = 1$ **to** N **do**
    Observe the partial gradient $G_i^t = 2D_i^\top z_t(y_i^t)^\top$
    Update $L_i^{t+1} = K_i^t - \eta_t G_i^t$
    if $\|L_i^{t+1}\|_2 > b_K$ **then**
      $K_i^{t+1} = L_i^{t+1}/b_K$
    else
      $K_i^{t+1} = L_i^{t+1}$
    end if
  end for
end for

To bound $\langle \nabla_t, K_t - K^\star \rangle_F$, consider

$$\mathbb{E}\left[\|K_{t+1} - K^\star\|_F^2 | \mathcal{F}_t\right] = \mathbb{E}\left[\|\Pi_{\mathcal{K}}(K_t - \eta_t \tilde{\nabla}_t) - K^\star\|_F^2 | \mathcal{F}_t\right]$$
$$\leq \|K_t - K^\star\|_F^2 + \eta_t^2 b_t^2 - 2\eta_t \langle \nabla_t, K_t - K^\star \rangle_F.$$

Taking $\eta_t = \frac{1}{\alpha t}$ and defining $\frac{1}{\eta_0} = 0$, we get $2\mathbb{E}\left[\sum_{t=1}^T \left(l_t(K_t) - l_t(K^\star)\right)\right] \leq \sum_{t=1}^T \frac{b_t^2}{\alpha t}$. □

We are now ready to apply online gradient descent to learn distributed team decisions.

5.1 Learning Team Decisions with Partial Gradient Information

We assume that the designer is aware of a lower bound on the strong convexity parameter, λ, and upper bound on the operator norm of the optimal policy b_K. The resulting algorithm, Algorithm 5, is a direct extension of online gradient descent. Its behavior is summarized in Theorem 20.

THEOREM 20—PARTIAL-GRADIENT FEEDBACK

Assume that Assumptions 1, 2 and 3 hold. Then, Algorithm 5 with step-size $\eta_t = \frac{1}{\lambda t}$ for any $0 < \lambda \leq \alpha$, where α is the strong-convexity parameter in Proposition 7, has bounded expected regret against the optimal policy $K^\star$ defined in (4). The bound is given by

$$\sum_{t=1}^T \mathbb{E}[l_t(K_t) - l_t(K^\star)] \leq \frac{b_G^2}{2\lambda}(1 + \log(T)). \quad (8)$$

The constant b_G in (8) is given by

$$b_G^2 = 4\|D\|_2^2(\|H\|_2 + b_K\|D\|_2(\|C\|_2+1))^2(\|C\|_2+1)^2 \\ (\kappa_x + 2\operatorname{Tr} V_{xx}\operatorname{Tr} V_{vv} + \kappa_v) \qquad \square$$

The regret bound is equivalent to that of online gradient descent in the convex optimization setting, where b_G takes the place of the bound on the gradient. The difference is that the bound holds in expectation and that b_G^2 is a bound on the second moment of the gradient estimator. Before proving Theorem 20 we need the following lemma to characterize the gradient estimate.

LEMMA 7
For $\|K\|_2 \leq b_K$, the gradient estimate $G_i^t := 2[D^\top]_i z_t(y_i^t)^\top$ is **consistent:** $\mathbb{E}[G_i^t] = \frac{\partial}{\partial K_i}\mathbb{E}[l_t(K)]$, and has **bounded variance**: $\mathbb{E}\left[\|G_t\|_F^2\right] \leq b_G^2$, where $G_t = \operatorname{Diag}(G_1^t, \ldots, G_N^t)$ and b_G satisfies

$$b_G^2 = 4\|D\|_2^2(\|H\|_2 + b_K\|D\|_2(\|C\|_2+1))^2(\|C\|_2+1)^2 \\ (\kappa_x + 2\operatorname{Tr} V_{xx}\operatorname{Tr} V_{vv} + \kappa_v). \quad (9)$$

$\square$

Proof Proof of Theorem 20.. Since all agents have the same loss functions, the partial gradient update is equivalent to a full gradient update. The result thus follows directly from Theorem 19 with the covariance-bounded gradient oracle in Lemma 7 and the strong convexity coefficient from Proposition 7. $\square$

5.2 Learning Team Decisions with Bandit Feedback

Towards constructing an estimator for the derivative, in addition to requiring the estimate to be consistent and have bounded variance, we insist that each agent must be able to compute her estimate independently. The last requirement invalidates the one-point estimate used in [Flaxman et al., 2004] as sampling from the unit sphere would require communication between agents. In [Shamir, 2013], the authors found that sampling uniformly and independently from the unit hypercube leads to consistent and bounded estimators for quadratic problems. Sampling from the hypercube reduces to sampling independent Rademacher variables coordinate-wise and can be done in a distributed fashion. Algorithm 6 is constructed by applying a matrix version of the estimate from [Shamir, 2013] and shrinking the exploration parameter ϵ_t each time-step. The regret properties of Algorithm 6 is summarized in Theorem 21.

THEOREM 21—BANDIT FEEDBACK
Assume that Assumptions 1, 2 and 3 hold. Then, Algorithm 6 with step-sizes $\eta_t = \frac{1}{\lambda t}$ for any $0 < \lambda \leq \alpha$ where α is the strong-convexity parameter in Proposition 7, and

Algorithm 6 Learning with bandit feedback

Input: initial guess K_0, bound b_K, step-sizes $\{\eta_t\}$ and exploration parameters $\{\epsilon_t\}$
for $t = 1$ **to** T **do**
  for $i = 1$ **to** N **do**
    Sample $R_i^t \in \{-1, 1\}^{m_i \times p_i} \sim$ Uniform
    Play $u_i^t = (K_i^t + R_i^t \mathcal{E}_i^t) y_i^t$ where $\mathcal{E}_i^t = \frac{\epsilon_t}{\sqrt{m_i p_i}}$
  end for
  The team incurs loss $l_t(K_t + R^t \mathcal{E}^t) = z_t^\top z_t$
  for $i = 1$ **to** N **do**
    Observe loss $l_t = z_t^\top z_t$
    Set gradient estimate $\tilde{G}_i^t = l_t R_i^t (\mathcal{E}_i^t)^{-1}$
    Update $L_i^{t+1} = K_i^t - \eta_t \tilde{G}_i^t$
    if $\|L_i^{t+1}\|_2 > b_K$ **then**
      $K_i^{t+1} = L_i^{t+1} / b_K$
    else
      $K_i^{t+1} = L_i^{t+1}$
    end if
  end for
end for

exploration parameters $\epsilon_t = t^{-1/4} \left(\sum_{i=1}^N m_i^2 p_i^2 \right)^{-1/4}$ has bounded expected regret against the optimal policy $K^\star$ defined in (4). The bound is given by

$$\sum_{t=1}^{T} \mathbb{E}[l_t(\tilde{K}_t) - l_t(K^\star)] \leq 2 \left(M_1 + \frac{M_2}{\lambda} \right) \left(\sum_{i=1}^{N} m_i^2 p_i^2 \right)^{1/2} \sqrt{T}. \tag{10}$$

In (10) $\tilde{K}_t = K_t + R_t \mathcal{E}_t$ is the policy played by the agents at time-step t. The problem-dependent constants M_1 and M_2 are given by

$$\begin{aligned} M_1 &= \|D\|_2^2 \left(\|C\|_2^2 \operatorname{Tr} V_{xx} + \operatorname{Tr} V_{vv} \right) \\ M_2 &= \left(\|H\|_2 + \|D\|_2 (b_K + 1)(\|C\|_2 + 1) \right)^4 \\ &\quad \times (\kappa_x + 2 \operatorname{Tr} V_{xx} \operatorname{Tr} V_{vv} + \kappa_v). \end{aligned}$$

□

To prove 21 we need the following lemma, which states that the gradient estimator used in Algorithm 6 is consistent and has bounded variance.

Lemma 8—Variant of lemma 2 in [Shamir, 2013]
Let $R_i \in \{-1, 1\}^{m_i \times p_i}$ be independent random variables following uniform distributions. Let $R = \operatorname{Diag}\{R_1, \ldots, R_N\}$, and $\mathcal{E} = \epsilon \operatorname{Diag}\{I/\sqrt{m_1 p_1}, \ldots, I/\sqrt{m_N p_N}\}$. Define the zeroth-order gradient estimator

$$\tilde{G}_i^t := l_t (K + R\mathcal{E}) R_i \mathcal{E}_i^{-1},$$

and let $\tilde{G}_t := \mathrm{Diag}\{\tilde{G}_1^t, \dots, \tilde{G}_N^t\}$. Under assumptions 1–3, $\tilde{G}_t$ is **consistent:** $\mathbb{E}_{R,x,v}\left[\tilde{G}_i^t\right] = \frac{\partial}{\partial K_i} J(K)$ and has **bounded variance** $\mathbb{E}_{R,x,v}\left[\|\tilde{G}_t\|_F^2\right] \leq \tilde{b}_G^2$. The bound, $\tilde{b}_G$ can be taken as

$$\tilde{b}_G^2 = \left(\|H\|_2 + \|D\|_2(b_K + \epsilon))(\|C\|_2 + 1)\right)^4 \times (\kappa_x + 2\,\mathrm{Tr}\,V_{xx}\,\mathrm{Tr}\,V_{vv} + \kappa_v) \sum_{i=1}^{N} m_i^2 p_i^2/\epsilon^2. \quad (11)$$

□

Note that the bound $\tilde{b}_G$ is decreasing in the exploration parameter ϵ_t, leading to an exploration/exploitation trade-off. The choice of ϵ_t minimizes the regret asymptotic upper bound.

Proof Proof of Theorem 21.. We will first quantify the added loss due to the perturbation term $R_t\mathcal{E}_t$. Let $\tilde{K}_t = K_t + R_t\mathcal{E}_t$, then

$$\begin{aligned}\mathbb{E}[l_t(\tilde{K}_t)] &= \mathbb{E}\left[\|Hx_t + D(K_t + R_t\mathcal{E}_t)Cy_t\|_2^2\right] \\ &= \mathbb{E}\Big[\|Hx_t + DKy_t\|_2^2 + \|DR_t\mathcal{E}_t y_t\|_2^2 \\ &\quad + 2(Hx_t + DKy_t)^\top R_t\mathcal{E}_t y_t\Big].\end{aligned}$$

By the first property of Lemma 9 we know that $\mathbb{E}[R_t] = 0$. Applying the fifth property we conclude that

$$\mathbb{E}[l_t(\tilde{K}_t)] \leq \mathbb{E}[l_t(K_t)] + \epsilon_t^2\|D\|_2^2\left(\|C\|_2^2\,\mathrm{Tr}\,V_{xx} + \mathrm{Tr}\,V_{vv}\right).$$

Combining this with Lemma 19, we get

$$\sum_{t=1}^{T} \mathbb{E}[J(\tilde{K}_t) - J(K^\star)] \leq \sum_{t=1}^{T} \frac{(\tilde{b}_G^t)^2}{\lambda t} + M_1 \sum_{t=1}^{T} \epsilon_t^2.$$

Substituting ϵ_t into $\tilde{b}_G^t$ from Lemma 8 and the inequality $\sum_{t=1}^T \frac{1}{\sqrt{t}} \leq 2\sqrt{T}$ completes the proof. □

6. Numerical example

In Fig. 1 we apply the algorithms to [Gattami, 2007, Example 4.1] for two players, where $C_1 = C_2 = 1$, $x, v_1, v_2 \sim \mathcal{N}(0, 1)$ and

$$H = \begin{bmatrix}1 & 0 & 0\end{bmatrix}^\top, \quad D = \begin{bmatrix}1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}^\top.$$

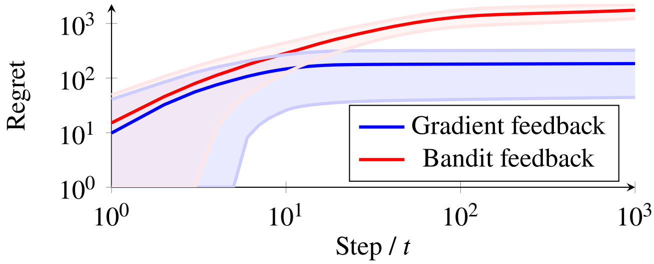


Figure 1. The average (solid lines) ± one standard deviation (shaded area) from 1280 simulations of Example 4.1 in [Gattami, 2007] using Algorithm 5 (gradient feedback, blue) and Algorithm 6 (bandit feedback, red).

Regret is bounded for Algorithm 5 by $46000(1 + \log(t))$ and for Algorithm 6 by $1.42 \cdot 10^6 \sqrt{T}$. The results from the 1280 simulations[2] in Fig. 1 indicates far better performance.

7. Conclusions and Future Research

We have proposed algorithms that efficiently learn optimal team decisions in a decentralized manner without knowing the problem parameters. The exploration required with bandit feedback gives worse asymptotic regret, both with respect to time and the number of parameters to be learned. Our work gives a first approach, and there are several interesting open questions to answer. Interesting directions for future research include learning when the covariance matrices change over time, applications to feedback control of dynamical systems, and empirical convergence studies.

8. Appendix

LEMMA 9
Let $R_i \in \{-1, 1\}^{m_i \times p_i}$ for $i = 1, \dots, N$ be independent random variables following uniform distributions, and take $R = \text{Diag}\{R_1, R_2, \dots, R_N\}$. Define $m = m_1 + \cdots + m_N$ and $p = p_1 + \cdot p_N$. Define the set

$$\mathcal{I}_R := \Big\{(i, k, l) \in \mathbb{N}^3 \,:\, i \in \{1, \dots, N\}, k \in \{1, \dots, m_i\}, l \in \{1, \dots, p_i\}\Big\}.$$

Let (i, k, l), (i', k', l') and $(\hat{i}, \hat{k}, \hat{l}) \in \mathcal{I}_R$. Them the following hold

[2] For a Julia implementation, See `https://github.com/kjellqvist/LearningTeamDecisions.jl`

1. $\mathbb{E}[R_i(k,l)] = 0$,
2. $\mathbb{E}[R_i(k,l)R_{i'}(k',l')] = \delta_{(i,k,l)=(i',k',l')}$,
3. $\mathbb{E}[R_i(k,l)R_{i'}(k',l')R_{\hat{i}}(\hat{k},\hat{l})] = 0$,
4. $\mathbb{E}\left[\mathrm{Tr}(AR^\top)R_i\right] = [A]_i$ for all $A \in \mathbb{R}^{m\times p}$,
5. $\|R_i\|_F = \sqrt{m_i p_i}$. □

References

Bubeck, S. (2015). "Convex optimization: algorithms and complexity". *Foundations and Trends® in Machine Learning* **8**, pp. 231–357. DOI: 10.1561/2200000050.

Chen, X. and E. Hazan (2021). "Black-box control for linear dynamical systems". In: Belkin, M. et al. (Eds.). *Proceedings of Thirty Fourth Conference on Learning Theory*. Vol. 134. Proceedings of Machine Learning Research. PMLR, pp. 1114–1143. URL: https://proceedings.mlr.press/v134/chen21c.html.

Cohen, A., A. Hassidim, T. Koren, N. Lazic, Y. Mansour, and K. Talwar (2018). "Online linear quadratic control". In: *ICML*.

Flaxman, A., A. T. Kalai, and H. B. McMahan (2004). "Online convex optimization in the bandit setting: gradient descent without a gradient". *CoRR* **cs.LG/0408007**. URL: http://arxiv.org/abs/cs.LG/0408007.

Gattami, A. (2007). *Optimal Decisions with Limited Information*. eng. PhD thesis. Lund University. URL: https://lup.lub.lu.se/search/files/4812314/26865.pdf.

Gattami, A., B. M. Bernhardsson, and A. Rantzer (2012). "Robust team decision theory". *IEEE Transactions on Automatic Control* **57**:3, pp. 794–798. DOI: 10.1109/TAC.2011.2168071.

Gupta, A., S. Yuksel, T. Başar, and C. Langbort (2014). "On the existence of optimal policies for a class of static and sequential dynamic teams". *SIAM Journal on Control and Optimization* **53**. DOI: 10.1137/14096534X.

Hazan, E. (2019). "Introduction to online convex optimization". *CoRR* **abs/1909.05207**. arXiv: 1909.05207. URL: http://arxiv.org/abs/1909.05207.

Hazan, E., S. Kakade, and K. Singh (2020). "The nonstochastic control problem". In: Kontorovich, A. et al. (Eds.). *Proceedings of the 31st International Conference on Algorithmic Learning Theory*. Vol. 117. Proceedings of Machine Learning Research. PMLR, pp. 408–421. URL: https://proceedings.mlr.press/v117/hazan20a.html.

Ho, Y.-C. and K. Chu (1972). “Team decision theory and information structures in optimal control problems–part i”. *IEEE Transactions on Automatic Control* **17**:1, pp. 15–22. DOI: 10.1109/TAC.1972.1099850.

Kim, S.-J. and G. B. Giannakis (2014). “Real-time electricity pricing for demand response using online convex optimization”. In: *ISGT 2014*, pp. 1–5. DOI: 10.1109/ISGT.2014.6816447.

Li, Y., S. Das, and N. Li (2021). “Online optimal control with affine constraints”. *Proceedings of the AAAI Conference on Artificial Intelligence* **35**:10, pp. 8527–8537. URL: https://ojs.aaai.org/index.php/AAAI/article/view/17035.

Mahajan, A., N. C. Martins, M. C. Rotkowitz, and S. Yüksel (2012). “Information structures in optimal decentralized control”. In: *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 1291–1306. DOI: 10.1109/CDC.2012.6425819.

Marschak, J. (1955). “Elements for a theory of teams”. *Management Science* **1**:2, pp. 127–137. URL: https://EconPapers.repec.org/RePEc:inm:ormnsc:v:1:y:1955:i:2:p:127-137.

Narayanaswamy, B., V. K. Garg, and T. S. Jayram (2012). “Online optimization for the smart (micro) grid”. In: *2012 Third International Conference on Future Systems: Where Energy, Computing and Communication Meet (e-Energy)*, pp. 1–10. DOI: 10.1145/2208828.2208847.

Radner, R. (1962). “Team Decision Problems”. *The Annals of Mathematical Statistics* **33**:3, pp. 857–881. DOI: 10.1214/aoms/1177704455. URL: https://doi.org/10.1214/aoms/1177704455.

Rakhlin, A., O. Shamir, and K. Sridharan (2012). “Making gradient descent optimal for strongly convex stochastic optimization.” In: *ICML*. icml.cc / Omnipress. URL: http://dblp.uni-trier.de/db/conf/icml/icml2012.html#RakhlinSS12.

Sanjari, S., T. Başar, and S. Yüksel (2021). “Policy-dependent and policy-independent static reduction of stochastic dynamic teams and games and fragility of equivalence properties”. In: *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 6231–6236. DOI: 10.1109/CDC45484.2021.9683260.

Shamir, O. (2013). “On the complexity of bandit and derivative-free stochastic convex optimization”. In: *Conference on Learning Theory*. PMLR, pp. 3–24.

Witsenhausen, H. S. (1968). “A counterexample in stochastic optimum control”. *SIAM Journal on Control* **6**:1, pp. 131–147. DOI: 10.1137/0306011.

Witsenhausen, H. S. (1988). “Equivalent stochastic control problems”. *Mathematics of Control, Signals and Systems* **1**, pp. 3–11.

Zanini, F., D. Atienza, G. Micheli, and S. Boyd (2010). "Online convex optimization-based algorithm for thermal management of mpsocs". *Proceedings of the ACM Great Lakes Symposium on VLSI, GLSVLSI*. DOI: 10.1145/1785481.1785532.

Zinkevich, M. (2003). "Online convex programming and generalized infinitesimal gradient ascent". In: *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*. ICML'03. AAAI Press, Washington, DC, USA, pp. 928–935. ISBN: 1577351894.