Proceedings of Machine Learning Research vol 168:1–11, 2022 4th Annual Conference on Learning for Dynamics and Control

Learning-Enabled Robust Control with Noisy Measurements

Olle Kjellqvist OLLE.KJELLQVIST@CONTROL.LTH.SE
Anders Rantzer ANDERS.RANTZER@CONTROL.LTH.SE
Automatic Control LTH
Lund University
Box 118
SE-221 00 Lund
Sweden

Editors: R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, M. Kochenderfer

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.

Keywords: adaptive control, real-time learning

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 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:

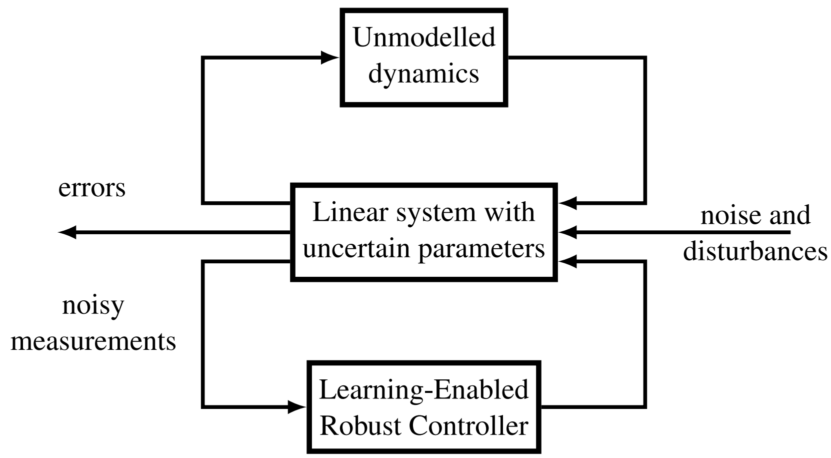


Figure 1: For a finite set of linear time-invariant models, the Learning-Enabled Robust Controller minimizes the ℓ_2-gain from noise and disturbances to errors for any realization of the unknown model parameters. This gain bound guarantees robustness to unmodelled dynamics.

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 7. 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. We mostly give sketches of proofs in this version of the article and some have been omitted altogether due to space constraints; complete proofs of all theorems and lemmata can be found in the extended version of the article[1].

1. https://github.com/kjellqvist/noisy-lerc

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 $\mathcal{M} \in \mathbb{R}^3$, 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 1 *The cases $b = -1$ and $b = 1$ cannot be simultaneously stabilized by a static feedback controller when $a \geq 1$.*

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

4. An information-state condition

In this section we will apply a slight modification to the $\mathcal{H}_\infty$-observer from Basar and Bernhard (1995) to bound (3) in a way which leads itself to recursive computation. We need the following lemma:

Lemma 3 (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*

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

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

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

Remark 4 *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 $\hat{g}_M = \gamma^2 ac/(P_M + \gamma^2 c^2 - 1)$.

Lemma 3 lets us express the worst-case accumulated cost compatible with the dynamics as a function of the past trajectory (u, y) and the next state $x(t+1)$, if the dynamics M of the system (1) are known. As $x(t+1)$ changes, so does the set of trajectories w, v that are compatible with $x(t+1)$. In particular, the entire sequence of a maximizing trajectory will change as $x(t+1)$ is varied. With that in mind, it is remarkable that the effect to the accumulated cost is captured completely by the term $-P\left(x(t+1) - \hat{x}(t+1)^2\right)$. The second term $l(t+1)$ contains the terms of the cost that depend only on past inputs and outputs and is independent of $x(t+1)$.

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

$$\alpha_M(t) := \sup_{w,v\in\ell_2[0,t],x_0\in\mathbb{R}} \{\alpha(t) : (a, b, c) = M, \text{subject to (1) and (2)}\} \quad (7)$$

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

Theorem 5 (Information-state condition) *Given a causal output-feedback control policy μ, a positive quantity γ, and an uncertainty set $\mathcal{M}$. Assume that for all $(a, b, c) = M \in \mathcal{M}$ the Riccati equation*

$$P_M = \left(\frac{a^2}{P_M + \gamma^2 c^2 - 1} + \gamma^{-2}\right)^{-1} \quad (8)$$

has a positive solution P_M and let

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

Further let

$$\hat{x}_M(t+1) = \hat{a}_M \hat{x}_M(t) + bu(t) + \hat{g}_M y(t), \ \hat{x}_M(0) = 0, \tag{9}$$

$$l_M(t+1) = l_M(t) - P_M \hat{x}_M(t)^2 - \gamma^2 y(t)^2 + \frac{(P_M \hat{x}_M(t) + \gamma^2 c y(t))^2}{P_M + \gamma^2 c^2 - 1}, \quad l_M(0) = 0. \tag{10}$$

Then the closed-loop system (1), (2) with control policy 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, then γ is not an upper bound of the ℓ_2-gain from disturbance to error.

Remark 6 *The quantities $l_M(t+1)$ are known at time t and depending only on measurements up until time t. It is thus available to the controller at time t. Further, $\alpha_M(t) = \ell_M(t+1)$.*

Proof [Sketch] Let $\alpha_M(t)$ be defined as in (7). 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 3 to express $\alpha_M(t)$ in the known quantities $\hat{x}_M(t)$, P_M and $l_M(t)$[2]:

$$\begin{aligned}
\alpha_M(t) &= \sup_{x \in \mathbb{R}, v \in \mathbb{R}} \left\{ x^2 - \gamma^2 v^2 - P_M \left(x - \hat{x}_M(t)\right)^2 + l_M(t) \right\} \\
&= \left(P_M \hat{x}_M(t) + \gamma^2 c y(t)\right)^2 / (P_M + \gamma^2 c^2 - 1) - P_M \hat{x}_M^2(t) - \gamma^2 y(t)^2 + l_M(t) \\
&= l_M(t+1).
\end{aligned}$$

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. ■

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?*. In 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 7 (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 (9), (10) and an observer-based control policy $\eta^\star$

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

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

2. We let subscript M denote quantities computed using $(a, b, c) = M$.

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

$$\mu_t^\star\left(y(0), y(1)\ldots, y(t)\right) = \eta^\star\left\{(\hat{x}_M(t), l_M(t+1), y(t)) : m \in \mathcal{M}\right\} \tag{11}$$

Proof [Sketch] *(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 admissible. By assumption, $l_M(T) \le 0$ for all T, M and $y \in \ell_2[0,T]$ for the controller (11), which we know implies that the system is finite gain by Theorem 5.

(i) implies (ii): Assume that the policy $\mu^\star$ 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 (9) and (10). At time $t = 0$ we have that $\hat{x}_M(0) = 0$ and that $l_M(1)$ are known functions of the known measurement $y(0)$. We can thus pick $\eta^\star\{0, l_M(1)(y(0)), y(0)\} = \mu^\star(y(0))$. Further, given $\hat{x}_M(t), l_M(t+1)$ and $y(t)$, assume that there exist a trajectory $\breve{y}$ so that $\breve{y}$ together with the controller $u(\tau) = \mu^\star(y(0), \ldots, y(\tau))$ could have generated $\hat{x}_M(t), l_M(t+1)$ and that $\breve{y}(t) = y(t)$. Then taking $\eta\{\hat{x}_M(t), l_M(t+1), y(t)\} = \mu^\star(\breve{y}(0), \ldots \breve{y}(t))$ ensures that the property holds for $t+1$. By induction, it holds for all t. By assumption (3) is fulfilled, thus $l_M(t+1) \le 0$. ■

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 7 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 (8) reduce to

$$P_i = P = \frac{1}{2}(1 - \gamma^2 a^2) + \sqrt{\gamma^2(-1+\gamma^2) + (\gamma^2 a^2 - 1)^2/4}. \tag{12}$$

Construct the observers $\hat{x}_i$ and cumulative objectives l_i using (9) and (10) 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

$$u(t) = \begin{cases} -(\hat{a}\hat{x}_1(t) + \hat{g}y(t)) & \text{if } l_1(t+1) \ge l_{-1}(t+1) \\ \hat{a}\hat{x}_{-1}(t) + \hat{g}y(t) & \text{if } l_1(t+1) < l_{-1}(t+1). \end{cases} \tag{13}$$

The dead-beat controller[3] 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.

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

Proposition 8 *With $\hat{a}$, $\hat{g}$, P as above, $\hat{x}_i$ and l_i as in (9) and (10), and the control signal given by (13), let*

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

Then the following is true:

$$\begin{aligned}
&1: \quad \hat{x}_1(t) = \begin{cases} 0, & \text{if } l_1(t) \geq l_{-1}(t) \\ \hat{x}(t), & \text{if } l_1(t) < l_{-1}(t) \end{cases}, \quad \hat{x}_{-1}(t) = \begin{cases} \hat{x}(t), & \text{if } l_1(t) \geq l_{-1}(t) \\ 0, & \text{if } l_1(t) < l_{-1}(t), \end{cases} \\
&2: \quad \begin{cases} l_1(t+1) &= \begin{cases} l_1(t) - \gamma^2 y(t)^2 + \frac{(\gamma^2 y(t))^2}{P+\gamma^2-1} & \text{if } l_1(t) \geq l_{-1}(t) \\ l_1(t) - P\hat{x}(t)^2 - \gamma^2 y(t)^2 + \frac{(P\hat{x}(t)+\gamma^2 y(t))^2}{P+\gamma^2-1}, & \text{if } l_1(t) < l_{-1}(t) \end{cases} \\ l_{-1}(t+1) &= \begin{cases} l_{-1}(t) - P\hat{x}(t)^2 - \gamma^2 y(t)^2 + \frac{(P\hat{x}(t)+\gamma^2 y(t))^2}{P+\gamma^2-1}, & \text{if } l_1(t) \geq l_{-1}(t) \\ l_{-1}(t) - \gamma^2 y(t)^2 + \frac{(\gamma^2 y(t))^2}{P+\gamma^2-1}, & \text{if } l_1(t) < l_{-1}(t) \end{cases} \end{cases}
\end{aligned} \tag{14}$$

Proof We start by proving the first claim. Consider the case when $l_1(t+1) \geq l_{-1}(t+1)$. Then $\hat{x}_1(t+1) = 0$ and $\hat{x}_{-1}(t+1) = \hat{a}(\hat{x}_1(t) + \hat{x}_{-1}(t)) + 2\hat{g}y(t)$. The case when $l_1(t+1) < l_{-1}(t+1)$ is similar. Taking $\hat{x}(t) = \hat{x}_1(t) + \hat{x}_{-1}(t)$ completes the proof. To see that the second claim is true, note that if $l_1(t) \geq l_{-1}(t)$ then $\hat{x}_1(t) = 0$ and $\hat{x}_{-1}(t) = \hat{x}(t)$. The claim follows by substitution into (10). ■

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

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

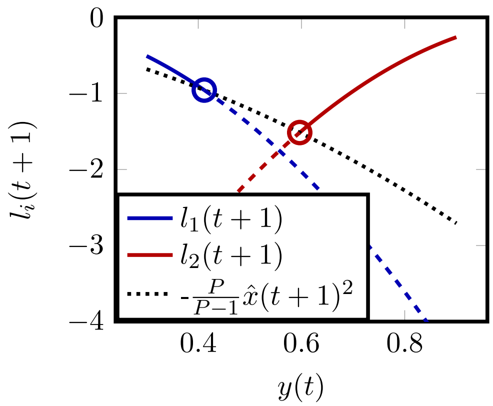
Then with $l_i(t+1)$ as in (14), it holds that $l_i(t+1) \leq 0$ for $i \in \{1, -1\}$.

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

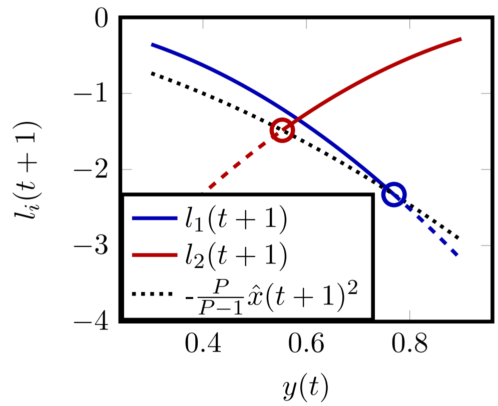
$$\frac{1}{\gamma^2} \geq \frac{1}{P+\gamma^2-1} \iff P+\gamma^2-1 \geq \gamma^2,$$

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

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



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



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

Figure 2: 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 9 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.

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 9 are fulfilled for all t. This is illustrated in Figure 2, 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 2 (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 10 (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 (15) *and* strong negativity condition (16) *below, then the closed-loop system* (1) *controlled with the certainty-equivalence deadbeat controller* (13) *has gain from $(w, v) \to x$ bounded above by γ.*

$$P > 2\gamma - 1 \tag{15}$$

$$(P + 2\gamma^2 - 1)\left(P - 1 - 2\sqrt{\gamma^2 - P})^2\right) \geq (P-1)\left((P+1)^2 - 4\gamma^2\right) \tag{16}$$

Remark 11 *We can solve* (16) *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 3.*

Remark 12 *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 3.

Proof [Sketch] 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 9 will hold for all t. Then, by Theorem 7 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 9 $\max_i\{l_i(t+1)\} \leq 0$. It remains to show that

$$\min_i\{l_i(t+1)\} \leq -\frac{P}{P-1}\hat{x}(t+1)^2. \tag{17}$$

Let $z(t) := y(t) - \frac{P}{2\gamma^2}\hat{x}(t)$. Then $\hat{x}(t+1) = 2\hat{g}z(t)$. For (17) 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 $P > 2\gamma - 1$, i.e. if the *curvature condition* (15) holds.

Strong negativity: Define the upper bounds

$$\bar{l}_1(t+1) := l_1(t+1) - l_1(t), \quad \bar{l}_{-1}(t+1) := l_{-1}(t+1) - l_{-1}(t) - \frac{P}{P-1}\hat{x}(t)^2, \tag{18}$$

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 (17) is satisfied if and only if $\#\left(\mathcal{I}_1 \cap \mathcal{I}_{-1}\right) \leq 1$, i.e. the intersection contains at most one point.. Since $\mathcal{I}_i \subseteq \bar{\mathcal{I}}_i$, a sufficient condition is $\#\left(\bar{\mathcal{I}}_1 \cap \bar{\mathcal{I}}_{-1}\right) \leq 1$. We can determine the boundary points of these sets by solving $\bar{l}_i(t+1) = -4\frac{P}{P-1}\hat{g}^2 z(t)^2$. One can the express $\#\left(\bar{\mathcal{I}}_1 \cap \bar{\mathcal{I}}_{-1}\right) \leq 1$ as

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

which simplifies to (16). ■

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 7 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

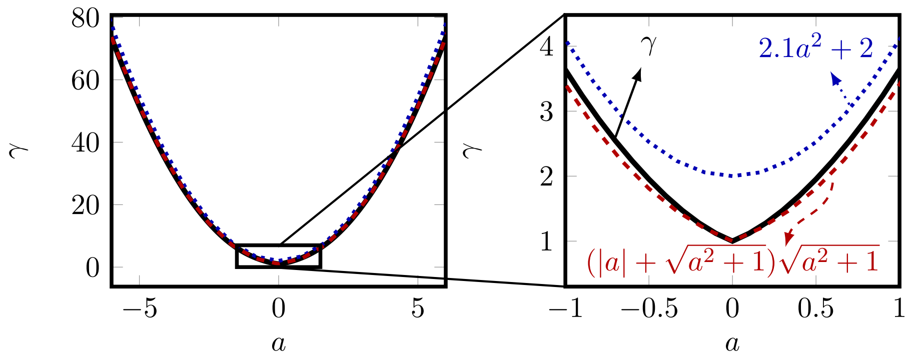


Figure 3: Guaranteed bound on the ℓ_2-gain from disturbances to error under feedback with the certainty equivalence controller with respect to a. We note that experimentally γ is lower bounded by $(|a| + \sqrt{a^2+1})\sqrt{a^2+1}$ and upper bounded by $\leq 2.1a^2 + 2$. The lower bound becomes tighter as a increases.

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 7 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.

Acknowledgments

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

Naman Agarwal, Brian Bullins, Elad Hazan, Sham Kakade, and Karan Singh. Online control with adversarial disturbances. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 111–119. PMLR, 09–15 Jun 2019. URL https://proceedings.mlr.press/v97/agarwal19c.html.

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

Sarah Dean, Stephen Tu, Nikolai Matni, and Benjamin Recht. Safely learning to control the constrained linear quadratic regulator. In *2019 American Control Conference (ACC)*, pages 5582–5588, 2019. doi: 10.23919/ACC.2019.8814865.

G. Didinsky and T. Basar. Minimax adaptive control of uncertain plants. In *Proceedings of 1994 33rd IEEE Conference on Decision and Control*, volume 3, pages 2839–2844 vol.3, 1994. doi: 10.1109/CDC.1994.411368.

Nikolai Matni, Alexandre Proutiere, Anders Rantzer, and Stephen Tu. From self-tuning regulators to reinforcement learning and back again. pages 3724–3740, 12 2019. doi: 10.1109/CDC40024.2019.9029916.

Anders Rantzer. Minimax adaptive control for a finite set of linear systems, 2021.

Max Simchowitz. Making non-stochastic control (almost) as easy as stochastic. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 18318–18329. Curran Associates, Inc., 2020.

Glenn Vinnicombe. Examples and counterexamples in finite l2-gain adaptive control. 04 2004.

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