

State and Parameter Estimation for Nonlinearly Parameterized Systems: An \mathcal{H}_∞ -Based Approach [★]

Håvard Fjær Grip ^{*} Ali Saberi ^{*} Tor A. Johansen ^{**}

^{*} School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164-2752, USA

^{**} Department of Engineering Cybernetics, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway

Abstract: We present a framework for estimating states and parameters in systems that can be described by a detectable linear system perturbed by a nonlinear function of the states, exogenous signals, and a vector of unknown, constant parameters. The estimators designed in this framework consist of two interconnected modules: a parameter estimator that is constructed as though the states and the nonlinear perturbation were directly available for measurement; and an observer/perturbation estimator that estimates the states as well as the perturbation. The design methodology is based on satisfying an \mathcal{H}_∞ condition, which leaves the designer with a wide range of options for carrying out the design. We discuss the use of LMI-based techniques in particular, and illustrate their application on a simulation example. We also discuss how the results of the paper can be applied to a more general class of cascaded nonlinear systems.

Keywords: parameter estimation; nonlinear parameterization; observer design

1. INTRODUCTION

A problem frequently encountered in model-based control and estimation is the presence of uncertain perturbations in the system equations. Such perturbations can be the result of external disturbances or internal plant changes, such as a configuration change, system fault, or changes in physical plant characteristics. The uncertainty associated with these perturbations can often be characterized in terms of a vector of unknown, constant parameters.

Unknown parameters are often dealt with by introducing parameter estimates that are updated online. For linearly parameterized systems there is a large body of work on adaptive control and estimation (see, e.g., Krstić, Kanellakopoulos, and Kokotović, 1995). For systems that are nonlinearly parameterized, however, only a few specialized techniques have been developed. Ortega (1996) presented a simplified version of a result by Fomin, Fradkov, and Yakubovich (1981), which applies to nonlinear parameterizations with a convexity property. Another method for convex or concave parameterizations was proposed by Annaswamy, Skantze, and Loh (1998) and extended to general nonlinear parameterizations by Loh, Annaswamy, and Skantze (1999). The results of Bošković (1995, 1998) and Zhang, Ge, Hang, and Chai (2000) focus on first-order systems with fractional parameterizations. Qu (2003) presented an approach for a class of higher-order systems with matrix fractional parameterizations, where an auxiliary estimate of the full perturbation to the system equations was used in the estimation of the unknown parameters. An approach for more general nonlinear parameterizations was presented by Qu, Hull, and Wang (2006), using a control law based on a parameter estimate that is biased by an

appropriately chosen vector function. Tyukin, Prokhorov, and van Leeuwen (2007) developed an adaptive design framework for monotonically parameterized perturbations based on the idea of *virtual algorithms* that are designed as though the time derivative of the measurements were available. Liu, Ortega, Su, and Chu (2011) have recently presented methods based on *Immersion & Invariance*.

The authors have previously presented a methodology for estimating unknown parameters in systems of the form $\dot{x} = f(t, x) + B(t, x)v(t, x) + \phi$, where $\phi = B(t, x)g(t, x, \theta)$ is a perturbation parameterized by the vector θ (Grip, Johansen, and Imsland, 2008; Grip, Johansen, Imsland, and Kaasa, 2010). This design methodology is based on the observation that, if the perturbation ϕ were directly available, then identifying θ would be a matter of inverting the nonlinear equation $\phi = B(t, x)g(t, x, \theta)$ with respect to θ . The overall design is therefore modular, consisting of a *parameter estimator* and a *perturbation estimator*. The parameter estimator is designed as though ϕ were known, to dynamically invert the expression $\phi = B(t, x)g(t, x, \theta)$ with respect to θ . The perturbation estimator is designed to produce an estimate of ϕ . The two modules are then connected, so that the parameter estimator uses the estimate of ϕ provided by the perturbation estimator instead of the actual perturbation. The parameter estimate is in turn fed back to the perturbation estimator.

The techniques cited above are all based on the assumption of full-state measurement. The authors have recently presented an extension of their previous work to the case of partial-state measurements, for systems that can be described as a linear system with a nonlinear perturbation (Grip, Saberi, and Johansen, 2009, 2011). This extension consists of replacing the perturbation estimator with a high-gain observer for an *extended system*, which estimates both the states of the system and the perturbation. Although the

[★] The work of Håvard Fjær Grip is supported by the Research Council of Norway. The work of Ali Saberi is partially supported by NAVY grants ONR KKK777SB001 and ONR KKK760SB0012.

extension greatly expands the class of systems that can be handled, there are some drawbacks to the way the high-gain observer is designed. Principal among these is the issue of complexity: the design is carried out by transforming the linear part of the system to a *special coordinate basis* (SCB) (Sannuti and Saberi, 1987); assigning gains using a pole-placement technique; and then scaling the gains in certain directions to achieve stability. This process requires considerable effort from the designer, not least in understanding the structural properties displayed by the SCB. In a subspace of the state space the gain is designed based only on the infinite zero structure of the system, which can be dramatically altered by infinitesimally small changes to the system matrices. This sensitivity can in some cases lead to poor conditioning that must be rectified through further effort by the designer. Finally, the design methodology implicitly assumes that a high-gain design is needed, and a minimum-phase condition is therefore imposed to ensure that high gain does not cause instability.

1.1 Contributions of This Paper

The goal of this paper is to present an alternative methodology for the case of partial-state measurements, which is conceptually simpler than the one presented by Grip et al. (2009), and which rectifies some of the drawbacks mentioned above. In particular, we shall replace our previous high-gain design with an *observer/perturbation estimator*. The observer/perturbation estimator serves the same purpose, and is implemented in the same way, as the high-gain observer. However, the gains can be chosen freely by the designer, as long as they ensure that the \mathcal{H}_∞ norm of a particular transfer matrix is sufficiently small. Consequently, a wide range of \mathcal{H}_∞ design techniques can be applied, including Riccati-based techniques, LMI-based techniques, and direct methods (see Saberi, Stoorvogel, and Sannuti, 2006).

1.2 Preliminaries

For vectors z_1, \dots, z_n , we denote by $\text{col}(z_1, \dots, z_n)$ the column vector obtained by stacking the elements of z_1, \dots, z_n . The operator $\|\cdot\|$ denotes the Euclidean norm for vectors and the induced Euclidean norm for matrices. For a symmetric matrix P , the maximum and minimum eigenvalues are denoted by $\lambda_{\max}(P)$ and $\lambda_{\min}(P)$. For a set $E \subset \mathbb{R}^n$, we write $(E - E) := \{z_1 - z_2 \in \mathbb{R}^n \mid z_1, z_2 \in E\}$. Throughout the paper, we assume that all signals and functions are sufficiently smooth to permit differentiation when necessary. When considering systems of the form $\dot{z} = F(t, z)$, we assume that all functions involved are sufficiently smooth to guarantee that $F: \mathbb{R}_{\geq 0} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is piecewise continuous in t and locally Lipschitz continuous in z , uniformly in t , on $\mathbb{R}_{\geq 0} \times \mathbb{R}^n$. The solution of this system, initialized at time $t = 0$ with initial condition $z(0)$ is denoted by $z(t)$. We shall generally simplify the notation by omitting function arguments.

2. PROBLEM FORMULATION

We consider systems of the following type:

$$\dot{x} = Ax + Bu + E\phi, \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m, \phi \in \mathbb{R}^k, \quad (1a)$$

$$y = Cx, \quad y \in \mathbb{R}^r, \quad (1b)$$

where x is the state; y is the output; ϕ is a perturbation to the system equations; and u is a time-varying input that is

well-defined for all $t \in \mathbb{R}_{\geq 0}$ and may include control inputs, reference signals, measured disturbances, or other known influences. For ease of notation, we introduce the vector $v := \text{col}(u, \gamma)$ of known signals. The perturbation is given by the expression $\phi = g(v, x, \theta)$, where g is a continuously differentiable function and $\theta \in \mathbb{R}^p$ is a vector of constant, unknown parameters.

We shall design an estimator for both the state x and the unknown parameter vector θ . The technicalities of this design are most easily overcome if the time derivative \dot{u} is well-defined and piecewise continuous, and x , u , \dot{u} , and θ are known *a priori* to belong to compact sets. We shall therefore make this assumption throughout the paper. We note that this assumption also implies that v and \dot{v} belong to compact sets. In most estimation problems, the restriction of the variables to compact sets is reasonable, because the states and inputs are typically derived from physical quantities with natural bounds. When designing update laws for parameter estimates, we also assume that a parameter projection can be implemented as described by Krstić et al. (1995, App. E), restricting the parameter estimates to a compact, convex set $\Theta \subset \mathbb{R}^p$, defined slightly larger than the set of possible parameter values. Throughout the paper, we assume that θ is always initialized from Θ , that is, $\hat{\theta}(0) \in \Theta$, and we shall make use of this assumption when constructing our proofs.

We denote by $X \subset \mathbb{R}^n$, $V \subset \mathbb{R}^{m+r}$, and $V' \subset \mathbb{R}^{m+r}$ the compact sets to which x , v , and \dot{v} belong, and by Φ the compact image of $V \times X \times \Theta$ under g .

For ease of notation, we define

$$d(v, \dot{v}, x, \theta, \phi) := \frac{\partial g}{\partial v}(v, x, \theta)\dot{v} + \frac{\partial g}{\partial x}(v, x, \theta)(Ax + Bu + E\phi),$$

which represents the time derivative of the perturbation ϕ .

Assumption 1. The pair (C, A) is detectable; and the triple (C, A, E) is left-invertible with no invariant zeros at the origin.

Assumption 2. There exists a number $\beta > 0$ such that for all $(v, \dot{v}, x, \theta, \phi) \in V \times V' \times X \times \Theta \times \Phi$ and for all $(\hat{x}, \hat{\theta}, \hat{\phi}) \in \mathbb{R}^n \times \Theta \times \mathbb{R}^k$, $\|d(v, \dot{v}, x, \theta, \phi) - d(v, \dot{v}, \hat{x}, \hat{\theta}, \hat{\phi})\| \leq \beta \|\text{col}(x - \hat{x}, \theta - \hat{\theta}, \phi - \hat{\phi})\|$.

Remark 1. In Assumption 2, we specify a Lipschitz-like condition on d , which is global in the sense that there are no bounds on \hat{x} and $\hat{\phi}$. Although this condition may appear restrictive, we are free to redefine $g(v, x, \theta)$ outside the compact set $V \times X \times \Theta$ without altering the accuracy of the system description (1). We may, for example, introduce a smooth saturation on x outside X , in which case the condition is satisfied if $g(v, x, \theta)$ is sufficiently smooth.

3. PARAMETER ESTIMATOR

When designing the parameter estimator, we pretend that x and ϕ are available, so that an estimator of the form

$$\dot{\hat{\theta}} = u_\theta(v, x, \phi, \hat{\theta}) \quad (2)$$

can be implemented. The function u_θ is designed so that (2) dynamically inverts the expression $\phi = g(v, x, \theta)$ with respect to θ , causing $\hat{\theta}$ to converge exponentially to θ . The formal requirement is given by the following assumption, which is stated in terms of the estimation error $\tilde{\theta} = \theta - \hat{\theta}$.

Assumption 3. There exist a differentiable function $V_u: \mathbb{R}_{\geq 0} \times (\Theta - \Theta) \rightarrow \mathbb{R}_{\geq 0}$ and positive constants a_1, \dots, a_4 such that for all $(t, \tilde{\theta}) \in \mathbb{R}_{\geq 0} \times \Theta$,

$$a_1 \|\tilde{\theta}\|^2 \leq V_u(t, \tilde{\theta}) \leq a_2 \|\tilde{\theta}\|^2, \quad (3)$$

$$\frac{\partial V_u}{\partial t}(t, \tilde{\theta}) - \frac{\partial V_u}{\partial \tilde{\theta}}(t, \tilde{\theta}) u_\theta(v, x, \phi, \theta - \tilde{\theta}) \leq -a_3 \|\tilde{\theta}\|^2, \quad (4)$$

$$\left\| \frac{\partial V_u}{\partial \tilde{\theta}}(t, \tilde{\theta}) \right\| \leq a_4 \|\tilde{\theta}\|. \quad (5)$$

In reality, x and ϕ are not available, and hence (2) is not implementable. Instead, the parameter estimator that we actually implement is

$$\dot{\hat{\theta}} = u_\theta(v, \hat{x}, \hat{\phi}, \hat{\theta}), \quad (6)$$

where \hat{x} and $\hat{\phi}$ are estimates of x and ϕ provided by the observer/perturbation estimator to be presented in Section 4. Thus, the dynamics of the estimation error becomes

$$\dot{\tilde{\theta}} = -u_\theta(v, \hat{x}, \hat{\phi}, \theta - \tilde{\theta}). \quad (7)$$

As mentioned in Section 1, we assume that the parameter estimates are restricted by projection to a set Θ , encompassing all possible values of θ . This requirement is formally stated by the following assumption:

Assumption 4. The update law (6) ensures that if $\hat{\theta}(0) \in \Theta$, then for all $t \geq 0$, $\hat{\theta}(t) \in \Theta$.

3.1 Constructing the Parameter Estimator

Constructing the parameter estimator to satisfy the above assumptions constitutes the greatest challenge in applying our methodology. Grip et al. (2010) discussed this issue in detail, and provided a series of propositions with accompanying examples that showed how the parameter estimator may be constructed for certain types of nonlinearly parameterized perturbations. The same propositions were presented with slight modifications for the partial-state measurement case by Grip et al. (2009). We repeat the propositions again in this section, but without proofs, examples, and some of the surrounding discussion.

The first two propositions apply to the case when the equality $\phi = g(v, x, \theta)$ can be explicitly inverted, either the whole time (Proposition 1), or just part of the time (Proposition 2).

Proposition 1. Suppose that for all $(v, x, \phi) \in V \times \mathbb{R}^n \times \mathbb{R}^k$, we can find a unique solution $\theta = \theta^*(v, x, \phi)$ from the equation $\phi = g(v, x, \theta)$. Then Assumptions 3 and 4 are satisfied with the update law $u_\theta(v, \hat{x}, \hat{\phi}, \hat{\theta}) = \text{Proj}(\Gamma(\theta^*(v, \hat{x}, \hat{\phi}) - \hat{\theta}))$, where Γ is a symmetric, positive-definite gain matrix.

Proposition 2. Suppose that there exists a known function $l: V \times \mathbb{R}^n \times \mathbb{R}^k \rightarrow [0, 1]$ such that $t \mapsto l(v(t), x(t), \phi(t))$ is piecewise continuous and that for all $(v, x, \phi) \in V \times \mathbb{R}^n \times \mathbb{R}^k$, $l(v, x, \phi) > 0$ implies that we can find a unique solution $\theta = \theta^*(v, x, \phi)$ from the equation $\phi = g(v, x, \theta)$. Suppose furthermore that there exist $T > 0$ and $\sigma > 0$ such that for all $t \in \mathbb{R}_{\geq 0}$, $\int_t^{t+T} l(v(\tau), x(\tau), \phi(\tau)) d\tau \geq \sigma$. Then Assumptions 3 and 4 are satisfied with the update law $u_\theta(v, \hat{x}, \hat{\phi}, \hat{\theta}) = \text{Proj}(l(v, \hat{x}, \hat{\phi})\Gamma(\theta^*(v, \hat{x}, \hat{\phi}) - \hat{\theta}))$, where Γ is a symmetric, positive-definite gain matrix.

When it is not possible or desirable to solve the inversion problem explicitly, it is often possible to implement the parameter estimator as a numerical search for the solution.

Proposition 3. Suppose that there exist a positive-definite matrix P and a function $M: V \times \mathbb{R}^n \times \Theta \rightarrow \mathbb{R}^{p \times k}$ such that for all $(v, x) \in V \times \mathbb{R}^n$, and for all pairs $\theta_1, \theta_2 \in \Theta$,

$$M(v, x, \theta_1) \frac{\partial g}{\partial \theta}(v, x, \theta_2) + \frac{\partial g}{\partial \theta}^\top(v, x, \theta_2) M^\top(v, x, \theta_1) \geq P.$$

Then Assumptions 3 and 4 are satisfied with the update law $u_\theta(v, \hat{x}, \hat{\phi}, \hat{\theta}) = \text{Proj}(\Gamma M(v, \hat{x}, \hat{\theta})(\hat{\phi} - g(v, \hat{x}, \hat{\theta})))$, where Γ is a symmetric, positive-definite gain matrix.

Proposition 4. Suppose that there exist a function $S: V \times \mathbb{R}^n \rightarrow \mathbb{S}_{p+}$ such that each element of $t \mapsto S(v(t), x(t))$ is piecewise continuous, where \mathbb{S}_{p+} is the cone of $p \times p$ positive-semidefinite matrices; and a function $M: V \times \mathbb{R}^n \times \Theta \rightarrow \mathbb{R}^{p \times k}$, both bounded for bounded x , such that for all $(v, x) \in V \times \mathbb{R}^n$ and for all pairs $\theta_1, \theta_2 \in \Theta$,

$$M(v, x, \theta_1) \frac{\partial g}{\partial \theta}(v, x, \theta_2) + \frac{\partial g}{\partial \theta}^\top(v, x, \theta_2) M^\top(v, x, \theta_1) \geq S(v, x).$$

Suppose furthermore that there exist $T > 0$ and $\sigma > 0$ such that for all $t \in \mathbb{R}_{\geq 0}$, $\int_t^{t+T} S(v(\tau), x(\tau)) d\tau \geq \sigma I$, and a number $L_g > 0$ such that for all $(v, x, \theta) \in V \times \mathbb{R}^n \times \Theta$, $\|g(v, x, \theta) - g(v, x, \hat{\theta})\| \leq L_g(\hat{\theta}^\top S(v, x)\hat{\theta})^{1/2}$. Then Assumptions 3 and 4 are satisfied with the update law $u_\theta(v, \hat{x}, \hat{\phi}, \hat{\theta}) = \text{Proj}(\Gamma M(v, \hat{x}, \hat{\theta})(\hat{\phi} - g(v, \hat{x}, \hat{\theta})))$, where Γ is a symmetric, positive-definite gain matrix.

In Section 5 we shall analyze the stability properties of the parameter estimator (6) together with the observer/perturbation estimator that will be introduced in the next section. To do so we shall need an additional assumption regarding u_θ .

Assumption 5. The parameter update law $u_\theta(v, \hat{x}, \hat{\phi}, \hat{\theta})$ is Lipschitz continuous with respect to $(\hat{x}, \hat{\phi})$, uniformly in $(v, \hat{\theta})$, on $V \times \mathbb{R}^n \times \mathbb{R}^k \times \Theta$.

Remark 2. The Lipschitz condition in Assumption 5 is global in the sense that \hat{x} and $\hat{\phi}$ are not presumed to be bounded, and it may therefore fail to hold in many cases. However, if a local version of the Lipschitz condition holds, then the update law is easily modified to satisfy Assumption 5 by introducing a saturation on \hat{x} and $\hat{\phi}$ outside the compact sets X and Φ . When checking Assumption 5, the projection in the update law may be disregarded, since the Lipschitz property is retained under projection (Grip et al., 2010, App. A).

4. OBSERVER/PERTURBATION ESTIMATOR

The purpose of the observer/perturbation estimator is to estimate both the state x and the perturbation ϕ . To accomplish this we extend the system (1) by introducing ϕ as a state, and introduce an extended system vector $\chi = [x^\top, \phi^\top]^\top$. The dynamics of the extended system is given by

$$\dot{\chi} = \bar{A}\chi + \bar{B}u + \bar{E}d(v, \dot{v}, x, \theta, \phi), \quad y = \bar{C}\chi, \quad (8)$$

where

$$\bar{A} = \begin{bmatrix} A & E \\ 0 & 0 \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad \bar{E} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad \bar{C} = [C \ 0].$$

We construct the following observer:

$$\dot{\hat{x}} = A\hat{x} + Bu + E\hat{\phi} + K_x(y - C\hat{x}), \quad (9a)$$

$$\dot{z} = -\frac{\partial g}{\partial \theta}(v, \hat{x}, \hat{\theta})\hat{\theta} - \frac{\partial g}{\partial x}(v, \hat{x}, \hat{\theta})K_x(y - C\hat{x}) + K_\phi(y - C\hat{x}), \quad (9b)$$

$$\dot{\hat{\phi}} = g(v, \hat{x}, \hat{\theta}) + z. \quad (9c)$$

where $K_x \in \mathbb{R}^{n \times r}$ and $K_\phi \in \mathbb{R}^{k \times r}$ are gain matrices. In (9), we have made use of the parameter estimate $\hat{\theta}$, produced by the parameter estimation module, and its time derivative $\dot{\hat{\theta}} = u_\theta(v, \hat{x}, \hat{\phi}, \hat{\theta})$. The variables \hat{x} and $\hat{\phi}$ are estimates of x and ϕ , and they are gathered in a vector $\hat{\chi} = [\hat{x}^\top, \hat{\phi}^\top]^\top$. It is convenient to analyze the observer in terms of $\hat{\chi}$, which

constitutes a nonsingular transformation from the original observer states (\hat{x}, z) .

Taking the time derivative of $\hat{\phi}$ yields $\dot{\hat{\phi}} = d(v, \dot{v}, \hat{x}, \hat{\theta}, \hat{\phi}) + K_\phi(\gamma - C\hat{x})$. Defining the error $\tilde{\chi} = \chi - \hat{\chi}$ it is therefore easily verified that the error dynamics of the observer becomes

$$\dot{\tilde{\chi}} = (\bar{A} - \bar{K}\bar{C})\tilde{\chi} + \bar{E}\bar{d}, \quad (10a)$$

$$\bar{d} = d(v, \dot{v}, x, \theta, \phi) - d(v, \dot{v}, \hat{x}, \hat{\theta}, \hat{\phi}), \quad (10b)$$

where $\bar{K} = [K_x^\top, K_\phi^\top]^\top$. In (10), \bar{d} can be viewed as a nonlinear disturbance term, which must be suppressed in order to ensure stability. Considering the dynamics (10a), the transfer matrix from the input point of \bar{d} to the state $\tilde{\chi}$ is given by

$$H(s) = (sI - \bar{A} + \bar{K}\bar{C})^{-1}\bar{E}.$$

Our stability results in the next section will be presented in terms of an \mathcal{H}_∞ condition on $H(s)$.

5. STABILITY

Before stating our main stability result for the overall estimator, we give a preliminary result for the observer/perturbation estimator alone. The proof of this and the other formal results in this section are given in Appendix A.

Lemma 1. There exists a $\gamma > 0$ such that, if \bar{K} is chosen such that $\bar{A} - \bar{K}\bar{C}$ is Hurwitz and $\|H(s)\|_\infty < \gamma$, then (assuming $\hat{\theta} \in \Theta$) the error dynamics (10) is input-to-state stable with respect to $\tilde{\theta}$.

Next, we state our main stability result for the combined parameter estimator and observer/perturbation estimator.

Theorem 1. There exists a $\gamma > 0$ such that, if \bar{K} is chosen such that $\bar{A} - \bar{K}\bar{C}$ is Hurwitz and $\|H(s)\|_\infty < \gamma$, then the error dynamics (10), (7) is exponentially stable with all initial conditions such that $\hat{\theta}(0) \in \Theta$ contained in the region of attraction.

Theorem 1 specifies that \bar{K} must be chosen to satisfy certain conditions. The question of whether such a \bar{K} exists is answered by the following theorem.

Theorem 2. There exists a $\gamma^* > 0$ such that, for all $\gamma > \gamma^*$, \bar{K} can be chosen such that $\bar{A} - \bar{K}\bar{C}$ is Hurwitz and $\|H(s)\|_\infty < \gamma$. Moreover, if (C, A, E) is minimum-phase, then $\gamma^* = 0$.

6. GAIN SYNTHESIS AND TUNING

Theorem 1 tells us that we can design our observer by ensuring that $\bar{A} - \bar{K}\bar{C}$ is Hurwitz and that $\|H(s)\|_\infty < \gamma$, for some $\gamma > 0$. This condition may not be achievable in general; however, Theorem 2 shows that it is always achievable if the system satisfies an additional minimum-phase condition. From the proof of Theorem 1, it is possible to compute an explicit numerical value of γ for use in constructing the gains. Such a computation is, however, likely to be conservative and lead to poor performance. It is therefore preferable in practice to tune the observer by starting with a large value of γ and decreasing it gradually until the desired stability and performance is achieved.

As a practical matter, ensuring that $\bar{A} - \bar{K}\bar{C}$ is Hurwitz and that $\|H(s)\|_\infty < \gamma$ can be achieved using several different \mathcal{H}_∞ design methods; specifically, Riccati-based methods, direct methods, and LMI-based methods (see Saberi et al., 2006).

The use of LMIs is attractive, because it allows for easy incorporation of additional performance criteria in the design process. For a given γ , it follows from the bounded-real lemma (see, e.g., Saberi et al., 2006, Th. 11.45) that $\|H(s)\|_\infty < \gamma$ is satisfied by choosing $\bar{K} = P^{-1}X$, where X and $P = P^\top > 0$ are solutions of the LMI

$$\begin{bmatrix} P\bar{A} + \bar{A}^\top P - X\bar{C} - \bar{C}^\top X^\top + I & P\bar{E} \\ \bar{E}^\top P & -\gamma^2 I \end{bmatrix} < 0. \quad (11)$$

Such a solution is far from unique—there are additional degrees of freedom in choosing \bar{K} that can be used to improve performance. In particular, it was shown by Chilali and Gahinet (1996) that by including additional LMIs based on a common Lyapunov matrix P , it is possible to constrain the closed-loop poles to some convex LMI region (assuming the region is feasible for the given \mathcal{H}_∞ objective), or to incorporate additional \mathcal{H}_∞ or \mathcal{H}_2 minimization objectives.

6.1 Limiting the Observer Gain

Of particular concern in observer design is the effect of measurement noise, which is severely amplified if the observer gains are chosen too large. Using LMIs, we can incorporate an additional objective to ensure that the gains are not chosen much larger than what is necessary to achieve $\|H(s)\|_\infty < \gamma$. Suppose that γ is affected by additive noise n ; that is, $\gamma = Cx + n$. Equation (10a) then becomes $\dot{\tilde{\chi}} = (\bar{A} - \bar{K}\bar{C})\tilde{\chi} + \bar{E}\bar{d} - \bar{K}n$. If we view this as a linear system with inputs \bar{d} and \tilde{n} , the transfer matrix from n to $\tilde{\chi}$ is $G(s) := -(sI - \bar{A} + \bar{K}\bar{C})^{-1}\bar{K}$. In the spirit of Chilali and Gahinet (1996), we can minimize a bound on $\|G(s)\|_\infty$, while at the same time satisfying (11), by minimizing a value $\gamma_n^2 > 0$ subject to the LMIs (11) and

$$\begin{bmatrix} P\bar{A} + \bar{A}^\top P - X\bar{C} - \bar{C}^\top X^\top + I & -X \\ -X^\top & -\gamma_n^2 I \end{bmatrix} < 0. \quad (12)$$

This LMI minimization problem can be solved using commonly available software (see Boyd, El Ghaoui, and Feron, 1994).

Remark 3. Note that the overall dynamics (10) is nonlinear, and thus there is no well-defined transfer function from n to the estimation error $\tilde{\chi}$. Minimizing $\|G(s)\|_\infty$ should therefore be viewed as a technique for limiting the magnitude of \bar{K} rather than a precise performance objective.

Remark 4. As in other \mathcal{H}_∞ -based design problems, it is beneficial to pre-scale the input and output channels of the transfer function to the same order of magnitude. Details on this topic are beyond the scope of this paper, and we refer instead to other sources (e.g., Skogestad and Postlethwaite, 2005).

7. SIMULATION EXAMPLE

In this section we revisit an example previously considered by Grip et al. (2009), concerning a DC motor with friction modeled by the LuGre friction model. The model is borrowed from Canudas de Wit and Lischinsky (1997). When this example was considered by Grip et al. (2009), a problem was encountered due to a small parameter σ_1 that altered the infinite zero structure of the system and resulted in excessively high gains. To achieve stability with more moderate gains, the design was carried out based on a manually altered model with $\sigma_1 = 0$. The \mathcal{H}_∞ -based methodology does not require such an alteration when the gains are synthesized using LMIs.

The model is described by $J\dot{\omega} = u - F$, where ω is the measured angular velocity, u is the motor torque, F is the friction

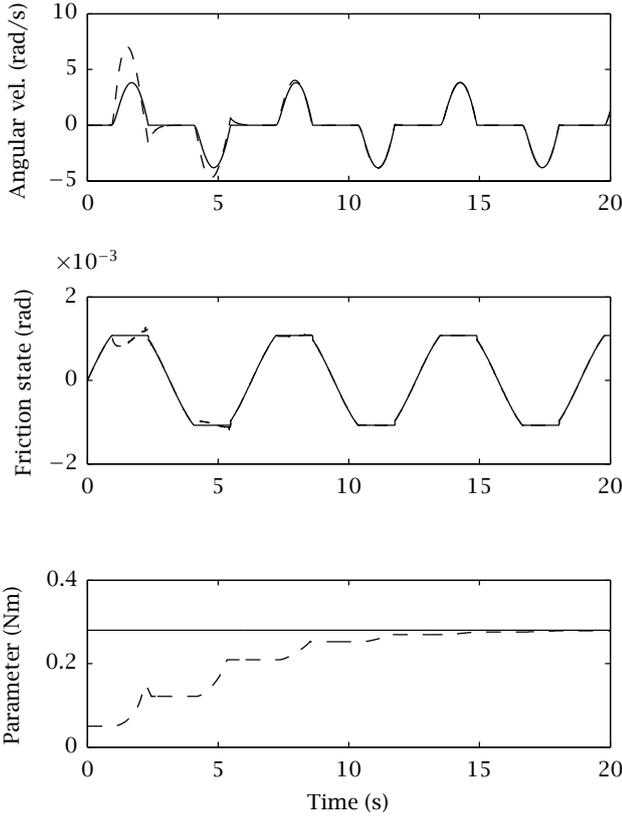


Fig. 1. Actual (solid) and estimated (dashed) values for simulation example

torque, and the parameter $J = 0.0023 \text{ kg m}^2$ is the motor and load inertia. The friction torque is given by the dynamic LuGre friction model: $F = \sigma_0 \eta + \sigma_1 \dot{\eta} + \alpha_2 \omega$, where the internal friction state η is given by $\dot{\eta} = \omega - \sigma_0 \eta |\omega| / \zeta(\omega)$, with $\zeta(\omega) = \alpha_0 + \alpha_1 \exp(-(\omega/\omega_0)^2)$. The model is parameterized by $\sigma_0 = 260.0 \text{ Nm/rad}$, $\sigma_1 = 0.6 \text{ Nm s/rad}$, $\alpha_0 = 0.28 \text{ Nm}$, $\alpha_1 = 0.05 \text{ Nm}$, $\alpha_2 = 0.0176 \text{ Nm s/rad}$, and $\omega_0 = 0.01 \text{ rad/s}$. As in the previous paper, we shall assume that these parameters are known, except for the uncertain parameter $\theta := \alpha_0$, which represents Coloumb friction. To indicate that ζ depends on the unknown parameter, we write $\zeta(\omega, \theta)$. We assume that θ is known *a priori* to belong to the range $\Theta := [0.05 \text{ Nm}, 1 \text{ Nm}]$. Following the notation from previous sections, we write $x = \text{col}(\omega, \eta)$, $y = \omega$, and $v = \text{col}(u, y)$. We furthermore define the perturbation $\phi = g(v, x, \theta) := \sigma_0 \eta |y| / \zeta(\omega, \theta)$. By extending the state space as described in Section 4, we obtain the system $J \dot{\omega} = u - (\alpha_2 + \sigma_1) \omega - \sigma_0 \eta + \sigma_1 \phi$, $\dot{\eta} = \omega - \phi$, $\dot{\phi} = d(v, x, \theta, \phi)$.

We design the parameter estimator as before (Grip et al., 2009), by noting that the equation $\phi = g(v, x, \theta)$ can be explicitly inverted with respect to θ when $\phi \neq 0$. Proposition 2 applies to such situations and yields the parameter estimator

$$\dot{\hat{\theta}} = \text{Proj}(l(v, \hat{x}, \hat{\phi}) \Gamma (\theta^*(v, \hat{x}, \hat{\phi}) - \hat{\theta})),$$

where θ^* represents the algebraic solution $\theta^*(v, \hat{x}, \hat{\phi}) := \sigma_0 \eta |y| / \hat{\phi} - \alpha_1 \exp(-(\hat{\omega}/\omega_0)^2)$. The function $l(v, \hat{x}, \hat{\phi})$ is defined as $l(v, \hat{x}, \hat{\phi}) = 0$ when $|\hat{\phi}| \leq 0.9$ and $l(v, \hat{x}, \hat{\phi}) = 1$ when $|\hat{\phi}| \geq 1$, with a linear transition between 0 and 1 for $0.9 < |\hat{\phi}| < 1$. We choose the gain $\Gamma = 1$.

To understand why this choice of parameter estimator makes sense, consider the error dynamics of $\tilde{\theta}$ in the hypothetical

situation that $\hat{x} = x$ and $\hat{\phi} = \phi$:

$$\dot{\tilde{\theta}} = -\text{Proj}(l(v, x, \phi) \tilde{\theta}).$$

It is easy to see that $\tilde{\theta}$ is drawn exponentially toward the origin when $l(v, x, \phi) > 0$ (i.e., $|\phi| > 0.9$), and that it remains constant otherwise. Thus, if $|\phi| > 0.9$ occurs a certain portion of the time, then the conditions specified by Proposition 2 are satisfied.

Next, we design the observer/perturbation estimator according to (9). It is straightforward to confirm that the triple (C, A, E) corresponding to the model is left-invertible and minimum-phase. Thus, by Theorem 2, we know that we can make $\|H(s)\|_\infty$ as small as is necessary to achieve stability.

We design the gains using the formulation described in Section 6, by minimizing $\gamma_n^2 > 0$ subject to the LMIs (11) and (12), for some sufficiently small choice of $\gamma > 0$. We find that $\gamma = 0.5$ yields stability and good performance, resulting in the gain $\bar{K} \approx [-0.70, -1.13 \cdot 10^{-5}, 4.95]^T$. The simulated states and the unknown parameter are shown together with the estimated values in Fig. 1.

8. CONCLUDING REMARKS

The methodology presented in this paper is focused on the case where θ represents a vector of unknown, constant parameters. However, the methodology can be applied in a straightforward way to the case when θ is a state variable described by $\dot{\theta} = f(v, \theta)$, provided an observer for θ of the form (6), satisfying Assumption 3, can be constructed. This observation effectively extends the results to a class of nonlinear cascaded systems.

REFERENCES

- Annaswamy, A. M., Skantze, F. P., Loh, A.-P., 1998. Adaptive control of continuous time systems with convex/concave parametrization. *Automatica* 34 (1), 33-49.
- Bošković, J. D., 1995. Stable adaptive control of a class of first-order nonlinearly parameterized plants. *IEEE Trans. Automat. Contr.* 40 (2), 347-350.
- Bošković, J. D., 1998. Adaptive control of a class of nonlinearly parameterized plants. *IEEE Trans. Automat. Contr.* 43 (7), 930-934.
- Boyd, S., El Ghaoui, L., Feron, E., 1994. *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial Mathematics, Philadelphia, PA.
- Canudas de Wit, C., Lischinsky, P., 1997. Adaptive friction compensation with partially known dynamic friction model. *Int. J. Adapt. Contr. Signal Process.* 11 (1), 65-80.
- Chilali, M., Gahinet, P., 1996. H_∞ design with pole placement constraints: An LMI approach. *IEEE Trans. Automat. Contr.* 41 (3), 358-367.
- Fomin, V., Fradkov, A., Yakubovich, V. (Eds.), 1981. *Adaptive Control of Dynamical Systems*. Nauka, Moscow.
- Grip, H. F., Johansen, T. A., Imslund, L., 2008. Estimation and control for systems with nonlinearly parameterized perturbations. In: *Proc. IFAC World Congr. Seoul, South Korea*, pp. 11178-11183.
- Grip, H. F., Johansen, T. A., Imslund, L., Kaasa, G.-O., 2010. Parameter estimation and compensation in systems with nonlinearly parameterized perturbations. *Automatica* 46 (1), 19-28.

- Grip, H. F., Saberi, A., Johansen, T. A., 2009. State and parameter estimation for linear systems with nonlinearly parameterized perturbations. In: Proc. IEEE Conf. Dec. Contr. Shanghai, China, pp. 8218-8225.
- Grip, H. F., Saberi, A., Johansen, T. A., 2011. Estimation of states and parameters for linear systems with nonlinearly parameterized perturbations. Syst. Contr. Lett. Accepted.
- Khalil, H. K., 2002. Nonlinear Systems, 3rd Edition. Prentice-Hall, Upper Saddle River, NJ.
- Krstić, M., Kanellakopoulos, I., Kokotović, P. V., 1995. Nonlinear and Adaptive Control Design. Wiley, New York.
- Liu, X., Ortega, R., Su, H., Chu, J., 2011. On adaptive control of nonlinearly parameterized nonlinear systems: Towards a constructive procedure. Syst. Contr. Lett. 60 (1), 36-43.
- Loh, A.-P., Annaswamy, A. M., Skantze, F. P., 1999. Adaptation in the presence of a general nonlinear parameterization: An error model approach. IEEE Trans. Automat. Contr. 44 (9), 1634-1652.
- Ortega, R., 1996. Some remarks on adaptive neuro-fuzzy systems. Int. J. Adapt. Contr. Signal Process. 10 (1), 79-83.
- Qu, Z., 2003. Adaptive and robust controls of uncertain systems with nonlinear parameterization. IEEE Trans. Automat. Contr. 48 (10), 1817-1823.
- Qu, Z., Hull, R. A., Wang, J., 2006. Globally stabilizing adaptive control design for nonlinearly-parametrized systems. IEEE Trans. Automat. Contr. 51 (6), 1073-1079.
- Saberi, A., Chen, B. M., Sannuti, P., 1991. Theory of LTR for non-minimum phase systems, recoverable target loops, and recovery in a subspace Part 1. Analysis. Int. J. Contr. 53 (5), 1067-1115.
- Saberi, A., Stoorvogel, A. A., Sannuti, P., 2006. Filtering Theory. Birkhäuser, Boston.
- Sannuti, P., Saberi, A., 1987. Special coordinate basis for multivariable linear systems—Finite and infinite zero structure, squaring down and decoupling. Int. J. Contr. 45 (5), 1655-1704.
- Skogestad, S., Postlethwaite, I., 2005. Multivariable Feedback Control. Wiley, Chichester, UK.
- Tyukin, I. Y., Prokhorov, D. V., van Leeuwen, C., 2007. Adaptation and parameter estimation in systems with unstable target dynamics and nonlinear parametrization. IEEE Trans. Automat. Contr. 52 (9), 1543-1559.
- Zhang, T., Ge, S. S., Hang, C. C., Chai, T. Y., 2000. Adaptive control of first-order systems with nonlinear parameterization. IEEE Trans. Automat. Contr. 45 (8), 1512-1516.

Appendix A. PROOFS

Proof of Lemma 1 By the bounded-real lemma (Saberi et al., 2006, Th. 11.45), $\|H(s)\|_\infty < \gamma$ and the fact that $\bar{A} - \bar{K}\bar{C}$ is Hurwitz implies that the LMI (11) with $X = P\bar{K}$ is satisfied for some positive definite P . Define the Lyapunov-like function $W = \tilde{x}^T P \tilde{x}$. The time derivative of W along the trajectories of the error dynamics (10) is

$$\begin{aligned} \dot{W} &= \tilde{x}^T (P(\bar{A} - \bar{K}\bar{C}) + (\bar{A} - \bar{K}\bar{C})^T P) \tilde{x} + 2\tilde{x}^T P \bar{E} \tilde{d} \\ &= \tilde{x}^T (P\bar{A} + \bar{A}^T P - X\bar{C} - \bar{C}^T X^T) \tilde{x} + 2\tilde{x}^T P \bar{E} \tilde{d} \\ &= \begin{bmatrix} \tilde{x} \\ \tilde{d} \end{bmatrix}^T \begin{bmatrix} P\bar{A} + \bar{A}^T P - X\bar{C} - \bar{C}^T X^T + I & P\bar{E} \\ \bar{E}^T P & -\gamma^2 I \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{d} \end{bmatrix} \\ &= -\|\tilde{x}\|^2 + \gamma^2 \|\tilde{d}\|^2. \end{aligned}$$

Using (11) and Assumption 2, we therefore have $\dot{W} \leq -(1 - \gamma^2 \beta^2) \|\tilde{x}\|^2 + \gamma^2 \beta^2 \|\tilde{\theta}\|^2$. Let γ be small enough that $1 - \gamma^2 \beta^2 \geq$

$2\varepsilon^2 > 0$. Then $\dot{W} \leq -\varepsilon^2 \|\tilde{x}\|^2$ whenever $\|\tilde{x}\| \geq \gamma\beta\|\tilde{\theta}\|/\varepsilon$, which implies ISS (Khalil, 2002, Th. 4.19). \square

Proof of Theorem 1 By Assumption 4, the parameter estimate $\hat{\theta}(t)$ cannot escape from Θ . It therefore follows from Lemma 1 that $\tilde{x}(t)$ is well-defined and remains in a compact subset of \mathbb{R}^{n+k} . Define the Lyapunov function $V = W + \gamma V_u$. Let $\tilde{\beta}$ denote the Lipschitz constant in Assumption 5, such that $\|u_\theta(v, x, \phi, \hat{\theta}) - u_\theta(v, \hat{x}, \hat{\phi}, \hat{\theta})\| \leq \tilde{\beta} \|\tilde{x}\|$. Using Assumption 3, we find that the time derivative of V along the trajectories of (10), (7) satisfies

$$\begin{aligned} \dot{V} &\leq -(1 - \gamma^2 \beta^2) \|\tilde{x}\|^2 + \gamma^2 \beta^2 \|\tilde{\theta}\|^2 \\ &\quad + \gamma \left(\frac{\partial V_u}{\partial t} - \frac{\partial V_u}{\partial \theta} u_\theta(v, x, \phi, \theta - \tilde{\theta}) \right) \\ &\quad + \gamma \frac{\partial V_u}{\partial \theta} (u_\theta(v, x, \phi, \theta - \tilde{\theta}) - u_\theta(v, \hat{x}, \hat{\phi}, \theta - \tilde{\theta})) \\ &\leq -(1 - \gamma^2 \beta^2) \|\tilde{x}\|^2 + \gamma^2 \beta^2 \|\tilde{\theta}\|^2 - \gamma a_3 \|\tilde{\theta}\|^2 + \gamma a_4 \tilde{\beta} \|\tilde{\theta}\| \|\tilde{x}\|. \end{aligned}$$

It is straightforward to confirm that the above quadratic expression is negative definite if γ is small enough that $\gamma < 4(1 - \gamma^2 \beta^2)(a_3 - \gamma \beta^2)/(a_4^2 \tilde{\beta}^2)$. The proof now follows from the comparison lemma (Khalil, 2002, Lemma 3.4). \square

Proof of Theorem 2 We start by showing that the pair (\bar{C}, \bar{A}) is detectable. Consider any eigenvalue λ of \bar{A} that is unobservable with respect to (\bar{C}, \bar{A}) . There exist $x \in \mathbb{R}^n$ and $p \in \mathbb{R}^k$, not both zero, such that

$$\begin{bmatrix} \bar{A} - \lambda I \\ \bar{C} \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} = 0 \iff \begin{bmatrix} (A - \lambda I)x + Ep \\ -\lambda p \\ Cx \end{bmatrix} = 0.$$

Clearly, either $\lambda = 0$ or $p = 0$. If $p = 0$, then it follows that $x \neq 0$ and $\begin{bmatrix} A - \lambda I \\ C \end{bmatrix} x = 0$, which implies that λ is an unobservable eigenvalue of (C, A) . Since (C, A) is detectable, λ must be in the open left-half complex plane. If $p \neq 0$, then $\lambda = 0$ and $\begin{bmatrix} A & E \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} = 0$, which implies that the Rosenbrock system matrix $R(z) = -\begin{bmatrix} A - zI & E \\ C & 0 \end{bmatrix}$ is rank-deficient for $z = 0$. Since the triple (C, A, E) is left-invertible, this rank-deficiency implies that (C, A, E) has an invariant zero at the origin (Saberi, Chen, and Sannuti, 1991), which contradicts Assumption 1. It follows that all unobservable eigenvalues of the pair (\bar{C}, \bar{A}) are in the open left-half complex plane, and hence it is detectable.

Since (\bar{C}, \bar{A}) is detectable, there exists a \bar{K} such that $\bar{A} - \bar{K}\bar{C}$ is Hurwitz, which implies that $\|H(s)\|_\infty < \tilde{\gamma}$ for some $\tilde{\gamma} > 0$. It follows that there exists a $\gamma^* \leq \tilde{\gamma}$ such that $\|H(s)\|_\infty < \gamma$ can be achieved for all $\gamma > \gamma^*$.

The error dynamics (10a) is identical to the error dynamics obtained by constructing a strictly proper filter of the CSS architecture (see Saberi et al., 2006, eq. (9.12)), for a system $\dot{x} = \bar{A}x + \bar{E}u$, $y = \bar{C}x$, where x is the state to be estimated, u is an unknown input, and y is the available output. Hence, according to Saberi et al. (2006, Th. 9.22), the gain matrix \bar{K} can be chosen such that $\bar{A} - \bar{K}\bar{C}$ is Hurwitz and $\|H(s)\|_\infty < \gamma$ for arbitrarily small $\gamma > 0$, provided the triple $(\bar{C}, \bar{A}, \bar{E})$ has no invariant zeros on the imaginary axis and the subspaces $S^{-0}(\bar{A}, \bar{E}, \bar{C}, 0)$ and $\mathcal{V}^*(\bar{A}, \bar{E}, \bar{C}, 0)$, defined by Saberi et al. (2006, Ch. 3.2.5), intersect only at the origin. According to Grip et al. (2009), left-invertibility and minimum phase of (C, A, E) implies left-invertibility and minimum phase of $(\bar{C}, \bar{A}, \bar{E})$. This implies that $(\bar{C}, \bar{A}, \bar{E})$ has no invariant zeros on the imaginary axis, and also that $S^{-0}(\bar{A}, \bar{E}, \bar{C}, 0) \cap \mathcal{V}^*(\bar{A}, \bar{E}, \bar{C}, 0) = \{0\}$ (see Saberi et al., 2006, Ch. 3.2.5). \square