State and Parameter Estimation for Linear Systems with Nonlinearly Parameterized Perturbations

Håvard Fjær Grip⋆ Ali Saberi⋆⋆ Tor A. Johansen⋆

Abstract— **We consider systems that can be described by a linear part with a nonlinear perturbation, where the perturbation is parameterized by a vector of unknown, constant parameters. Under a set of technical assumptions about the perturbation and its relationship to the outputs, we present a modular observer design technique. The observers produced by this design technique consist of a modified high-gain observer that estimates the states of the system together with the full perturbation, and a parameter estimator. The parameter estimator is constructed by the designer to identify the unknown parameters by dynamically inverting a nonlinear equation. We illustrate the design technique by constructing an observer for a DC motor with friction modeled by the dynamic LuGre friction model.**

I. Introduction

A common problem in model-based estimation and control is the handling of unknown or partially unknown perturbations to 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. In this paper, we consider observer design for systems that can be described by a linear part with a nonlinear, time-varying perturbation that is parameterized by a vector of unknown, constant parameters. Our design extends the results of [1], [2], where parameter identification for systems with full state measurement was considered.

In the literature, many of the results that deal with the handling of unknown or partially unknown perturbations are based either on high-gain designs, where the perturbations are suppressed by using a sufficiently high gain, or adaptive designs, where online estimates of unknown parameters are used to cancel the perturbations. Many of the results on high-gain observer design for nonlinear systems, for example [3], [4], rely on placing the system in some variation of the normal form, where it is represented as a chain of integrators from the nonlinearities to the outputs, possibly with a separate subsystem representing the zero dynamics. An alternative to the normal form is the special coordinate basis (SCB) introduced in [5], where the system is divided into subsystems that reflect its zero structure and invertibility properties. While high-gain designs based on the normal form are limited to systems that are square-invertible, minimum-phase, and of uniform relative degree with respect to the nonlinearities, [6] shows that high-gain observers can be designed based on the SCB for more general multiple-input multiple-output (MIMO) systems that are left-invertible and minimum-phase with respect to the nonlinearities.

In the presence of an unknown perturbation, high-gain observer designs can be extended by considering the perturbation as an additional state to be estimated. The perturbation estimate can be used for control, for example, to cancel the actual perturbation. This type of approach has been adopted recently in the context of performance recovery, in [7], where the perturbation appears in the first derivatives, and in [8], where it appears at the end of an integrator chain in a single-input single-output (SISO) system, and is estimated by an *extended high-gain observer*. A strength of the high-gain designs mentioned above is that little structural information about the perturbation is needed for implementation. Conversely, however, these methods do not fully exploit structural information about uncertainties in the perturbation, even when such information is available.

Adaptive techniques are based on online estimation of unknown parameters that can be considered constant or slowly varying. Most of the adaptive literature deals with systems that are linear with respect to the unknown parameters (see, e.g., [9]). Ways of handling nonlinear parameterizations include the extended Kalman filter (see [10]), various methods for handling nonlinearities with convex or concave parameterizations [11]–[14], and methods for first-order systems with fractional parameterizations [15]–[17]. Adaptive techniques related to our work include [18]–[20]. In [18], perturbations with matrix fractional parameterizations are considered, by introducing an auxiliary estimate of the full perturbation that is used in estimation of the unknown parameters. In [19], adaptive update laws for a class of nonlinearly parameterized perturbations in the first derivatives of the outputs are designed, by first creating virtual algorithms based on information from the derivatives, and then transforming these into implementable form without explicit differentiation. In [20], adaptive control for linear parameterizations is designed for the case of full state measurement, and then extended to the case of partial state measurement by using high-gain estimates of the output derivatives.

In this paper, we introduce a modular design consisting of two separate but connected modules. The first module is a *modified high-gain observer* that estimates the states

⋆Department of Engineering Cybernetics, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway. The work of Håvard Fjær Grip and Tor A. Johansen is supported by the Research Council of Norway.

⋆⋆School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164-2752. The work of Ali Saberi is partially supported by NAVY grants ONR KKK777SB001 and ONR KKK760SB0012, and National Aeronautics and Space Administration grant NNA06CN26A.

of the system and the unknown perturbation. The second module is a *parameter estimator* that uses the state and perturbation estimates to produce estimates of the unknown parameters, which are in turn fed back to the modified high-gain observer. This approach allows us to exploit structural information about the perturbation instead of suppressing all uncertainties using high gain. The price we pay is that the method is restricted to a limited class of perturbations that satisfy persistency-of-excitation requirements allowing for exponentially stable estimation of the unknown parameters. We base our design on the SCB and the theory from [6], enabling us to deal comfortably with a large class of MIMO systems.

A. Preliminaries

For a set of vectors $z_1, \dots, z_n$, we denote by $\mathrm{col}(z_1, \dots, z_n)$ 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 positive-semidefinite matrix P, the maximum and minimum eigenvalues are denoted $\lambda_{\max}(P)$ and $\lambda_{\min}(P)$. For a set $E \subset \mathbb{R}^n$, we write $(E - E) := \{z_1 - z_2 \in \mathbb{R}^n \mid z_1, z_2 \in E\}$. When considering systems of the form $\dot{z} = F(t, z)$, we assume that all functions involved are sufficiently smooth to guarantee that $F\colon \mathbb{R}_{\geq 0} \times \mathbb{R}^n \to \mathbb{R}^n$ 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 $z(t_0)$ is denoted $z(t)$. We shall often simplify notation by omitting function arguments.

II. Problem Formulation

We consider systems of the following type:

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

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

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 $\nu := \mathrm{col}(u, y)$ of known signals. The perturbation is given by the expression $\phi = g(\nu, x, \theta)$, where $g\colon \mathbb{R}^{m+r} \times \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^k$ is differentiable in ν and continuously differentiable in (x, θ), and $\theta \in \mathbb{R}^p$ is a vector of constant, unknown parameters. As we construct an estimator, additional smoothness requirements may be needed for g, as well as other functions, to guarantee that the piecewise continuity and local Lipschitz conditions in Section I-A hold for all systems involved. It is left to the designer to check these requirements.

We shall design an observer to estimate both the state x and the unknown parameter vector θ. The technicalities of this observer 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 ν and $\dot{\nu}$ 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 in [9], restricting the parameter estimates to a compact, convex set $\Theta \subset \mathbb{R}^p$, defined slightly larger than the set of possible parameter values. The parameter projection is denoted $\mathrm{Proj}(\cdot)$. 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, ν, and $\dot{\nu}$ belong, and by Φ the compact image of $V \times X \times \Theta$ under g.

For ease of notation, we define

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

representing the time derivative of the perturbation ϕ.

Assumption 1: The triple (C, A, E) is left-invertible and minimum-phase.

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

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 $g(\nu, x, \theta)$ outside $V \times X \times \Theta$ 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 can be satisfied by requiring that $g(\nu, x, \theta)$ is sufficiently smooth.

In the following sections, we shall first present the modified high-gain observer, and then the parameter estimator. While the modified high-gain observer has a fixed structure, the parameter estimator can be designed in a number of different ways to solve the equation $\phi = g(\nu, x, \theta)$ with respect to θ. In Section IV we discuss several ways to design the parameter estimator.

III. Modified High-Gain Observer

Our goal is to estimate both the state x and the perturbation ϕ. To accomplish this goal we extend the system by introducing ϕ as a state. The extended system then becomes

$$\begin{bmatrix} \dot{x} \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} A & E \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \phi \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ I_k \end{bmatrix} d(\nu, \dot{\nu}, x, \theta, \phi), \tag{2}$$

with $y = Cx$. We shall implement an observer for this system of the following form:

$$\dot{\hat{x}} = A\hat{x} + Bu + E\hat{\phi} + K_x(\varepsilon)(y - C\hat{x}), \tag{3a}$$

$$\dot{z} = -\frac{\partial g}{\partial \theta}(\nu, \hat{x}, \hat{\theta})\dot{\hat{\theta}} - \frac{\partial g}{\partial x}(\nu, \hat{x}, \hat{\theta})K_x(\varepsilon)(y - C\hat{x}) + K_\phi(\varepsilon)(y - C\hat{x}), \tag{3b}$$

$$\hat{\phi} = g(\nu, \hat{x}, \hat{\theta}) + z. \tag{3c}$$

where $K_x(\varepsilon) \in \mathbb{R}^{n \times r}$ and $K_\phi(\varepsilon) \in \mathbb{R}^{k \times r}$ are gain matrices parameterized by a number $0 < \varepsilon \leq 1$ to be specified later. In (3), we have made use of a parameter estimate $\hat{\theta}$ and its time derivative. These values are produced by the parameter estimation module, which we shall discuss in Section IV. The perturbation estimate $\hat{\phi}$ is seen to consist of $g(\nu, \hat{x}, \hat{\theta})$ plus an internal variable z. The variable z can be viewed as an adjustment made to $g(\nu, \hat{x}, \hat{\theta})$ to produce a more accurate estimate of the perturbation. Taking the time derivative of $\hat{\phi}$ yields $\dot{\hat{\phi}} = d(\nu, \dot{\nu}, \hat{x}, \hat{\theta}, \hat{\phi}) + K_\phi(\varepsilon)(y - C\hat{x})$. Defining the errors $\tilde{x} = x - \hat{x}$, $\tilde{\phi} = \phi - \hat{\phi}$, and $\tilde{y} = y - C\hat{x}$, we may therefore write the error dynamics of the observer as

$$\begin{bmatrix} \dot{\tilde{x}} \\ \dot{\tilde{\phi}} \end{bmatrix} = \begin{bmatrix} A & E \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{\phi} \end{bmatrix} + \begin{bmatrix} 0 \\ I_k \end{bmatrix} \tilde{d} - \begin{bmatrix} K_x(\varepsilon) \\ K_\phi(\varepsilon) \end{bmatrix} \tilde{y}, \quad (4)$$

where $\tilde{d} := d(\nu, \dot{\nu}, x, \theta, \phi) - d(\nu, \dot{\nu}, \hat{x}, \hat{\theta}, \hat{\phi})$. In the error dynamics (4), the term $\tilde{d}$ acts as an unknown disturbance, and $\tilde{y}$ is an available output. Our goal is to design a family of gains $K(\varepsilon) := [K_x^\top(\varepsilon), K_\phi^\top(\varepsilon)]^\top$ such that, as the number ε becomes small, the modified high-gain observer produces stable estimates with a diminishing effect from the parameter error $\tilde{\theta} := \theta - \hat{\theta}$.

Lemma 1: The error dynamics (4) with input $\tilde{d}$ and output $\tilde{y}$, and gain $K(\varepsilon) = 0$, is minimum-phase and left-invertible.

Proof: See Appendix. ■

A. SISO Systems

Suppose first that $k = r = 1$, and consider the error dynamics (4) with $K(\varepsilon) = 0$. Let n_a denote the number of invariant zeros in the system, and let $n_q = n - n_a$ denote the relative degree. Let Λ_1, Λ_2, and Λ_3 denote the nonsingular state, output, and input transformations that take the system (4) with input $\tilde{d}$ and output $\tilde{y}$, and with $K(\varepsilon) = 0$, to the SCB. We apply these transformations to the system (4) (without setting $K(\varepsilon) = 0$), by writing $\mathrm{col}(\tilde{x}, \tilde{\phi}) = \Lambda_1 \chi$, $\tilde{y} = \Lambda_2 \gamma$, and $\tilde{d} = \Lambda_3 \delta$. From [6, Theorem 2.2], the transformed state vector χ is partitioned as $\chi = \mathrm{col}(\chi_a, \chi_q)$, where χ_a has dimension n_a and χ_q has dimension n_q, and the resulting system is written as

$$\dot{\chi}_a = A_a \chi_a + L_{aq}\gamma - K_a(\varepsilon)\gamma, \quad (5a)$$

$$\dot{\chi}_q = A_q \chi_q + B_q(\delta + D_a \chi_a + D_q \chi_q) - K_q(\varepsilon)\gamma, \quad (5b)$$

$$\gamma = C_q \chi_q, \quad (5c)$$

where

$$A_q = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \; B_q = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \; C_q = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}.$$

The matrices $K_a(\varepsilon)$ and $K_q(\varepsilon)$ in (5) are observer gains, related to the gain $K(\varepsilon)$ by $K(\varepsilon) = \Lambda_1 [K_a^\top(\varepsilon), K_q^\top(\varepsilon)]^\top \Lambda_2^{-1}$. Once we have chosen $K_a(\varepsilon)$ and $K_q(\varepsilon)$, we can therefore implement the modified high-gain observer (3).

1) SCB Structure: In (5), the χ_a subsystem represents the zero dynamics of (4). In particular, the eigenvalues of the matrix A_a correspond to the invariant zeros of (4). Since the system is minimum-phase, this implies that A_a is Hurwitz. The χ_q subsystem represents the infinite zero structure of (4), and consists of a single chain of integrators, from the input δ to the output γ, with an interconnection to the zero dynamics at the lowest level of the integrator chain.

2) Design of SISO Gains: To design the observer gains, let $\bar{K}_q := \mathrm{col}(\bar{K}_{q_1}, \ldots, \bar{K}_{q_{n_q}})$ be chosen such that the matrix $H := A_q - \bar{K}_q C_q$ is Hurwitz. Because of the special structure of A_q, this is always possible by using regular pole-placement techniques. Then, let $K_a(\varepsilon) = L_{aq}$ and $K_q(\varepsilon) = \mathrm{col}(\bar{K}_{q_1}/\varepsilon, \ldots, \bar{K}_{q_{n_q}}/\varepsilon^{n_q})$.

B. MIMO Systems

Consider again the error dynamics (4) with $K(\varepsilon) = 0$. Let n_a and n_q denote the number of invariant zeros and infinite zeros in the system, respectively, and define $n_b = n - n_a - n_q$. Let Λ_1, Λ_2, and Λ_3 denote the nonsingular state, output, and input transformations that take the system (4) with input $\tilde{d}$ and output $\tilde{y}$, and with $K(\varepsilon) = 0$, to the SCB. We apply these transformations to the system (4) (without setting $K(\varepsilon) = 0$), by writing $\mathrm{col}(\tilde{x}, \tilde{\phi}) = \Lambda_1 \chi$, $\tilde{y} = \Lambda_2 \gamma$, and $\tilde{d} = \Lambda_3 \delta$. From [6, Th. 2.6], the transformed state χ is partitioned as $\chi = \mathrm{col}(\chi_a, \chi_b, \chi_q)$, where χ_a has dimension n_a, χ_b has dimension n_b, and χ_q has dimension n_q. The subsystem χ_q can be further partitioned as $\chi_q = \mathrm{col}(\chi_{q_1}, \ldots, \chi_{q_k})$, where each χ_{q_j}, $j = 1, \ldots, k$, has dimension n_{q_j}. The transformed output is partitioned as $\gamma = \mathrm{col}(\gamma_q, \gamma_b)$, where γ_q has dimension k and is further partitioned as $\gamma_q = \mathrm{col}(\gamma_{q_1}, \ldots, \gamma_{q_k})$, and γ_b has dimension $r_b := r - k$. The transformed input has dimension k and is further partitioned as $\delta = \mathrm{col}(\delta_1, \ldots, \delta_k)$. The resulting system is written as

$$\dot{\chi}_a = A_a \chi_a + L_{aq}\gamma_q + L_{ab}\gamma_b - K_{aq}(\varepsilon)\gamma_q - K_{ab}(\varepsilon)\gamma_b, \quad (6a)$$

$$\dot{\chi}_b = A_b \chi_b + L_{bq}\gamma_q - K_{bq}(\varepsilon)\gamma_q - K_{bb}(\varepsilon)\gamma_b, \quad (6b)$$

$$\dot{\chi}_{q_j} = A_{q_j}\chi_{q_j} + L_{q_j q}\gamma_q - K_{q_j q}(\varepsilon)\gamma_q - K_{q_j b}(\varepsilon)\gamma_b + B_{q_j}(\delta_j + D_{a_j}\chi_a + D_{b_j}\chi_b + D_{q_j}\chi_q), \quad (6c)$$

$$\gamma_b = C_b \chi_b, \quad \gamma_{q_j} = C_{q_j}\chi_{q_j}, \quad (6d)$$

where A_{q_j}, B_{q_j}, and C_{q_j} have the same special structure as A_q, B_q, and C_q in Section III-A, and where (C_b, A_b) is an observable pair. The gains $K_{aq}(\varepsilon)$, $K_{ab}(\varepsilon)$, $K_{bq}(\varepsilon)$, $K_{bb}(\varepsilon)$, $K_{q_j q}(\varepsilon)$, and $K_{q_j b}(\varepsilon)$, $j = 1, \ldots, k$, are related to $K(\varepsilon)$ by

$$K(\varepsilon) = \Lambda_1 \begin{bmatrix} K_{aq}(\varepsilon) & K_{ab}(\varepsilon) \\ K_{bq}(\varepsilon) & K_{bb}(\varepsilon) \\ K_{qq}(\varepsilon) & K_{qb}(\varepsilon) \end{bmatrix} \Lambda_2^{-1}, \quad (7)$$

where $K_{qq}(\varepsilon) = [K_{q_1 q}^\top(\varepsilon), \ldots, K_{q_k q}^\top(\varepsilon)]^\top$ and $K_{qb}(\varepsilon) = [K_{q_1 b}^\top(\varepsilon), \ldots, K_{q_k b}^\top(\varepsilon)]^\top$. Once we have chosen $K_{aq}(\varepsilon)$, $K_{ab}(\varepsilon)$, $K_{bq}(\varepsilon)$, $K_{bb}(\varepsilon)$, $K_{qq}(\varepsilon)$, and $K_{qb}(\varepsilon)$, we can therefore implement the modified high-gain observer (3) with the gains given by (7).

1) SCB Structure: As in the SISO case, the χ_a subsystem in (6) represents the zero dynamics of (4), with the eigenvalues of A_a corresponding to the invariant zeros of (4). The system (6) has a χ_b subsystem that is not present in the SISO case. This subsystem represents states that are observable from the output γ_b, but that are not directly affected by any inputs. As in the SISO case, the χ_q subsystem represents the infinite zero structure of (4). It is divided into k integrator chains, from δ_j to γ_{q_j}, with interconnections to other subsystems at the lowest level of each integrator chain.

2) Design of MIMO Gains: Let $K_{bb}(\varepsilon) = K_{bb}$ be chosen independently from ε such that the matrix $A_b - K_{bb}C_b$ is Hurwitz. This is always possible and can be carried out using standard pole-placement techniques, since the pair (C_b, A_b) is observable. For each $j \in 1, \dots, k$, select $\bar{K}_{q_j} := \mathrm{col}(\bar{K}_{q_j 1}, \dots, \bar{K}_{q_j n_{q_j}})$ such that the matrix $H_j := A_{q_j} - \bar{K}_{q_j} C_{q_j}$ is Hurwitz. Then, let $K_{aq}(\varepsilon) = L_{aq}$, $K_{ab}(\varepsilon) = L_{ab}$, $K_{bq}(\varepsilon) = L_{bq}$, $K_{qb}(\varepsilon) = 0$, and let $K_{q_j q}(\varepsilon) = K'_{q_j q} + L_{q_j q}$, with $K'_{q_j q}$ given by

$$\left[0_{n_{q_j} \times (j-1)} \quad \mathrm{col}\left(\frac{\bar{K}_{q_j 1}}{\varepsilon^{\bar{n}_q - n_{q_j} + 1}}, \dots, \frac{\bar{K}_{q_j n_{q_j}}}{\varepsilon^{\bar{n}_q}}\right) \quad 0_{n_{q_j} \times (k-j)}\right],$$

where $\bar{n}_q := \max_{j=1,\dots,k} n_{q_j}$.

Lemma 2: If the gains are chosen according to Section III-A.2 (SISO) or Section III-B.2 (MIMO), then there exists $0 < \varepsilon^* \le 1$ such that for all $0 < \varepsilon \le \varepsilon^*$, the error dynamics (4) is input-to-state stable (ISS) with respect to $\tilde{\theta}$.

Proof: See Appendix. ■

Remark 2: By selecting the gains as described above, we place n_a poles of the linear part of the observer error dynamics at the locations of the invariant zeros of (4), and we place n_b poles freely, as the poles of $A_b - K_{bb}C_b$. The last n_q poles are placed far into the left-half complex plane, asymptotically as ε becomes small.

IV. Parameter Estimator

As previously mentioned, the goal of the parameter estimation module is to produce an estimate of θ based on the perturbation estimate $\hat{\phi}$ and the state estimate $\hat{x}$. For this to work, we require an update law

$$\dot{\hat{\theta}} = u_\theta(\nu, \hat{x}, \hat{\phi}, \hat{\theta}), \tag{8}$$

which, in the hypothetical case of perfect state and perturbation estimates ($\hat{\phi} = \phi$ and $\hat{x} = x$), would provide an unbiased asymptotic estimate of θ. This requirement is formally stated by the following assumption on the dynamics of the error variable $\tilde{\theta}$.

Assumption 3: There exist a differentiable function $V_\mathrm{u}\colon \mathbb{R}_{\ge 0} \times (\Theta - \Theta) \to \mathbb{R}_{\ge 0}$ and positive constants $a_1, \dots, a_4$ such that for all $(t, \tilde{\theta}) \in \mathbb{R}_{\ge 0} \times (\Theta - \Theta)$,

$$a_1 \|\tilde{\theta}\|^2 \le V_\mathrm{u}(t, \tilde{\theta}) \le a_2 \|\tilde{\theta}\|^2, \tag{9}$$

$$\frac{\partial V_\mathrm{u}}{\partial t}(t, \tilde{\theta}) - \frac{\partial V_\mathrm{u}}{\partial \tilde{\theta}}(t, \tilde{\theta}) u_\theta(\nu, x, \phi, \theta - \tilde{\theta}) \le -a_3 \|\tilde{\theta}\|^2, \tag{10}$$

$$\left\| \frac{\partial V_\mathrm{u}}{\partial \tilde{\theta}}(t, \tilde{\theta}) \right\| \le a_4 \|\tilde{\theta}\|. \tag{11}$$

Furthermore, the update law (8) ensures that if $\hat{\theta}(t_0) \in \Theta$, then for all $t \ge t_0$, $\hat{\theta}(t) \in \Theta$.

Satisfying Assumption 3 constitutes the greatest challenge in applying the method presented this paper, and this issue is therefore discussed in the next section.

A. Satisfying Assumption 3

Assumption 3 guarantees that the origin of the error dynamics

$$\dot{\tilde{\theta}} = -u_\theta(\nu, \hat{x}, \hat{\phi}, \theta - \tilde{\theta}), \tag{12}$$

is uniformly exponentially stable with $(\Theta - \Theta)$ contained in the region of attraction whenever $\hat{\phi} = \phi$ and $\hat{x} = x$. Essentially, this amounts to asymptotically solving the inversion problem of finding θ given $\phi = g(\nu, x, \theta)$. In the following, we shall discuss some possibilities for how to satisfy Assumption 3. As a useful reference, we point to [21], which deals with the use of state observers for inversion of nonlinear maps. The material in this section is a straightforward adaptation of [2, Sec. 3.2], which also contains examples of each of the propositions below applied to particular perturbations.

The most obvious way to satisfy Assumption 3 is to invert the equality $\phi = g(\nu, x, \theta)$ algebraically, and to let $\hat{\theta}$ be attracted to the solution. This may be possible the whole time (Proposition 1), or just part of the time (Proposition 2).

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

Proof: The proof follows from using the function $V_\mathrm{u} = \frac{1}{2}\tilde{\theta}^\top \Gamma^{-1} \tilde{\theta}$ when $\hat{\phi} = \phi$ and $\hat{x} = x$. ■

The proofs of the remaining propositions in this section are found in the Appendix.

Proposition 2: Suppose that there exists a known, piecewise continuous function $l\colon V \times \mathbb{R}^n \times \mathbb{R}^k \to [0, 1]$, and that for all $(\nu, x, \phi) \in V \times \mathbb{R}^n \times \mathbb{R}^k$, $l(\nu, x, \phi) > 0$ implies that we can find a unique solution $\theta = \theta^*(\nu, x, \phi)$ from the equation $\phi = g(\nu, x, \theta)$. Suppose furthermore that there exist $T > 0$ and $\sigma > 0$ such that for all $t \in \mathbb{R}_{\ge 0}$, $\int_t^{t+T} l(\nu(\tau), x(\tau), \phi(\tau))\, \mathrm{d}\tau \ge \sigma$. Then Assumption 3 is satisfied with the update law $u_\theta(\nu, \hat{x}, \hat{\phi}, \hat{\theta}) = \mathrm{Proj}(l(\nu, \hat{x}, \hat{\phi})\Gamma(\theta^*(\nu, \hat{x}, \hat{\phi}) - \hat{\theta}))$, 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 update function as a numerical search for the solution.

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

$$M(\nu, x, \theta_1) \frac{\partial g}{\partial \theta}(\nu, x, \theta_2) + \frac{\partial g}{\partial \theta}^\top (\nu, x, \theta_2) M^\top(\nu, x, \theta_1) \ge 2P. \tag{13}$$

Then Assumption 3 is satisfied with the update law $u_\theta(\nu, \hat{x}, \hat{\phi}, \hat{\theta}) = \mathrm{Proj}(\Gamma M(\nu, \hat{x}, \hat{\theta})(\hat{\phi} - g(\nu, \hat{x}, \hat{\theta})))$, where Γ is a symmetric, positive-definite gain matrix.

Proposition 3 applies to certain monotonic perturbations for which a solution can be found arbitrarily fast by increasing the gain Γ. In many cases this is not possible, because the inversion problem is singular the whole time or part of the time. The following proposition applies to cases where a solution is only available by using data over longer periods of time, by incorporating a persistency-of-excitation condition.

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

$$M(\nu, x, \theta_1)\frac{\partial g}{\partial \theta}(\nu, x, \theta_2) + \frac{\partial g}{\partial \theta}^{\mathsf{T}}(\nu, x, \theta_2)M^{\mathsf{T}}(\nu, x, \theta_1) \geq 2S(\nu, x). \quad (14)$$

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

V. Stability of Interconnected System

When connecting the two modules, we need one additional assumption about the parameter update law.

Assumption 4: The parameter update law $u_\theta(\nu, \hat{x}, \hat{\phi}, \hat{\theta})$ is Lipschitz continuous in $(\hat{x}, \hat{\phi})$, uniformly in $(\nu, \hat{\theta})$, on $V \times \mathbb{R}^n \times \mathbb{R}^k \times \Theta$.

Remark 3: The Lipschitz condition in Assumption 4 is a global one, in the sense that $\hat{x}$ and $\hat{\phi}$ are not presumed to be bounded. Such a condition may fail to hold in many cases. However, if a local Lipschitz condition holds, then the update law is easily modified to satisfy Assumption 4 by introducing a saturation on $\hat{x}$ and $\hat{\phi}$ outside X and Φ. We also remark that when checking Assumption 4, the projection in the update law may be disregarded, since the Lipschitz property is retained under projection [2, App. A].

Theorem 1: If the gains are chosen according to Section III-A.2 (for SISO systems) or Section III-B.2 (for MIMO systems), then there exists $0 < \varepsilon^* \leq 1$ such that for all $0 < \varepsilon \leq \varepsilon^*$, the origin of the error dynamics (4), (12) is exponentially stable with $\mathbb{R}^{n+k} \times (\Theta - \Theta)$ contained in the region of attraction.

Proof: See Appendix. ■

VI. Redesign for Improved State Estimates

The modified high-gain observer is designed to estimate not only the states x, but also the perturbation ϕ. This configuration may not result in optimal estimates of x. In many cases, less noisy estimates may be obtained by implementing a second observer that does not include a perturbation estimate. Instead, this observer has the form $\dot{\hat{x}} = A\hat{x} + Bu + Eg(\nu, \hat{x}, \hat{\theta}) + K_x(\varepsilon)(y - C\hat{x})$, with $\hat{\theta}$ obtained from the first observer. The second observer can be designed using the same high-gain methodology as the modified high-gain observer, by transforming the error dynamics $\dot{\tilde{x}} = A\tilde{x} + E\tilde{g} - K_x(\varepsilon)\tilde{y}$ to the form (6) and designing gains. In this case, $\tilde{g} := g(\nu, x, \theta) - g(\nu, \hat{x}, \hat{\theta})$, rather than $\tilde{d}$, is the input that is transformed to δ, and we need to impose the same Lipschitz-like assumptions on g as we previously did on d. The resulting ISS property of the second observer with respect to the parameter error $\tilde{\theta}$, and the boundedness of $\tilde{\theta}$, justifies this type of cascaded design.

VII. Discussion

The purpose of the modified high-gain observer described in Section III is to estimate the states and the perturbation with sufficient accuracy to guarantee stability in closed loop with the parameter estimator. As implied by the name, a high gain may sometimes be required to achieve stability. Such cases are of little practical interest, because a high gain usually results in unacceptable noise amplification. Often, however, only a moderate gain is required. Since the required gain depends on quantities that are difficult to determine analytically, such as Lipschitz constants, the observer is typically tuned by starting with $\varepsilon = 1$ and decreasing ε in small decrements until satisfactory performance is achieved.

The zero structure of the system is of vital importance in the high-gain design methodology presented in Section III. It is well-known that small perturbations to the system matrices can lead to dramatic changes in a system's zero structure. In a SISO system, for example, an arbitrarily small perturbation of the system matrices can reduce the system's relative degree, resulting in singularly perturbed zero dynamics [22]. The result may be a poorly conditioned system description, with very large elements occurring in the SCB system matrices. In these situations, it is often better to design the observer gains with respect to the unperturbed system matrices. In Section VIII, we discuss a problem for which this precise situation is encountered.

VIII. Simulation Example

We consider the example of a DC motor with friction modeled by the LuGre friction model, borrowed from [23]. 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 torque, and the parameter $J = 0.0023\,\mathrm{kg\,m^2}$ 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 $\sigma_0 = 260.0\,\mathrm{Nm/rad}$, $\sigma_1 = 0.6\,\mathrm{Nm\,s/rad}$, $\alpha_0 = 0.28\,\mathrm{Nm}$, $\alpha_1 = 0.05\,\mathrm{Nm}$, $\alpha_2 = 0.176\,\mathrm{Nm\,s/rad}$, and $\omega_0 = 0.01\,\mathrm{rad/s}$. We shall assume that these parameters are known, except for the uncertain parameter $\theta := \alpha_0$, which represents static Coloumb friction. To indicate that ζ depends on the unknown parameter, we shall henceforth write $\zeta(\omega, \theta)$. We assume that θ is known *a priori* to belong to the range $\Theta := [0.05\,\mathrm{Nm}, 1\,\mathrm{Nm}]$. Following the notation from previous sections, we write $x = \mathrm{col}(\omega, \eta)$, $y = \omega$, and $\nu = \mathrm{col}(u, y)$. Let us define the perturbation $\phi = g(\nu, x, \theta) := \sigma_0\eta|y|/\zeta(\omega, \theta)$. It is straightforward to confirm that the system with input ϕ and output y is left-invertible and minimum-phase, as required by Assumption 1. By extending the state space as described in Section III,

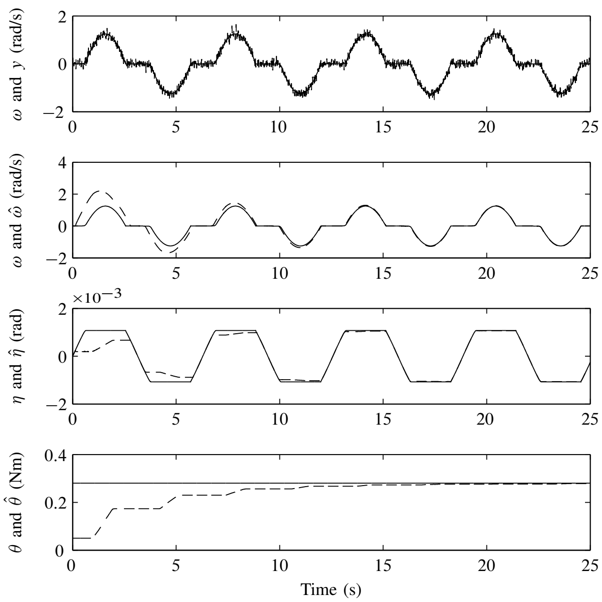

Fig. 1. From top: angular velocity ω (solid) and noisy measurement y (dashed); angular velocity ω (solid) and estimate $\hat{\omega}$ (dashed); internal friction state η (solid) and estimate $\hat{\eta}$ (dashed); unknown parameter θ (solid) and parameter estimate $\hat{\theta}$ (dashed)

we obtain the system $J\dot{\omega} = u - (\alpha_2 + \sigma_1)\omega - \sigma_0\eta + \sigma_1\phi$, $\dot{\eta} = \omega - \phi$, $\dot{\phi} = d(\nu, x, \theta, \phi)$.

Remark 4: In the definition of g, we have used the absolute value of the output y, rather than ω, which implies that $g(\nu, \hat{x}, \hat{\theta})$ will be implemented as $\sigma_0\eta|y|/\zeta(\hat{\omega}, \hat{\theta})$, rather than $\sigma_0\eta|\hat{\omega}|/\zeta(\hat{\omega}, \hat{\theta})$. The reason for defining g in this way is that the presence of $|\hat{\omega}|$ in $g(\nu, \hat{x}, \hat{\theta})$ would make it impossible to satisfy the Lipschitz-like condition on d in Assumption 2.

A. Observer Design

We start the observer design by creating a modified high-gain observer according to (3). We then obtain error dynamics of the form (4):

$$\begin{bmatrix} \dot{\tilde{\omega}} \\ \dot{\tilde{\eta}} \\ \dot{\tilde{\phi}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{J}(\alpha_2 + \sigma_1) & -\frac{1}{J}\sigma_0 & \frac{1}{J}\sigma_1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\omega} \\ \tilde{\eta} \\ \tilde{\phi} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tilde{d} - \begin{bmatrix} K_x(\varepsilon) \\ K_\phi(\varepsilon) \end{bmatrix} \tilde{y},$$

where $K_x(\varepsilon)$ and $K_\phi(\varepsilon)$ are to be designed. Using the software from [24] to transform the error dynamics to the SCB, we obtain

$$\begin{aligned} \dot{\chi}_a &= -\frac{\sigma_0}{\sigma_1}\chi_a + \frac{\sigma_0(-\sigma_1\alpha_2 + J\sigma_0)}{\sigma_1^3}\gamma - K_a(\varepsilon)\gamma, \\ \dot{\chi}_{q_1} &= \chi_{q_2} - K_{q_1}(\varepsilon)\gamma, \\ \dot{\chi}_{q_2} &= -\frac{\sigma_0(-\sigma_1\alpha_2 + J\sigma_0)}{J\sigma_1^2}\chi_{q_1} + \frac{J\sigma_0 - \sigma_1\alpha_2 - \sigma_1^2}{\sigma_1 J}\chi_{q_2} \\ &\quad + \frac{\sigma_0}{J}\chi_a + \delta - K_{q_2}(\varepsilon)\gamma, \quad \gamma = \chi_{q_1}. \end{aligned}$$

We may now proceed to design the gains according to Section III-A.2. However, we quickly discover that the problem is poorly conditioned, and that we consequently require an unacceptably large gain. The reason for the poor conditioning is that the parameter σ_1 acts as a small perturbation in the system matrices that reduces the relative degree of the system and results in singularly perturbed zero dynamics. This is precisely the situation discussed in Section VII. To eliminate this problem, we design the gains with respect to the simplified design system obtained by setting $\sigma_1 = 0$:

$$\begin{bmatrix} \dot{\tilde{\omega}} \\ \dot{\tilde{\eta}} \\ \dot{\tilde{\phi}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{J}\alpha_2 & -\frac{1}{J}\sigma_0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tilde{d} - \begin{bmatrix} K_x(\varepsilon) \\ K_\phi(\varepsilon) \end{bmatrix} \tilde{y}.$$

This system is transformed to the SCB representation $\dot{\chi}_{q_1} = \chi_{q_2} - K_{q_1}(\varepsilon)\gamma$, $\dot{\chi}_{q_2} = \chi_{q_3} - K_{q_2}(\varepsilon)\gamma$, $\dot{\chi}_{q_3} = -\frac{\sigma_0}{J}\chi_{q_2} - \frac{\alpha_2}{J}\chi_{q_3} + \delta - K_{q_3}(\varepsilon)\gamma$, $\gamma = \chi_{q_1}$. We now design the gains according to Section III-A.2, placing the poles of $H = A_q - \bar{K}_q C_q$ at -1, -2, and -3.

To design the parameter estimator, we note that whenever $\phi \neq 0$, we have $\theta = \sigma_0\eta|y|/\phi - \alpha_1\exp(-(\omega/\omega_0)^2)$. We may therefore apply Proposition 2, by defining $\dot{\hat{\theta}} = \mathrm{Proj}(l(\nu, \hat{x}, \hat{\phi})\Gamma(\theta^*(\nu, \hat{x}, \hat{\phi}) - \hat{\theta}))$, where θ^* represents the algebraic solution $\theta^*(\nu, \hat{x}, \hat{\phi}) := \sigma_0\hat{\eta}|y|/\hat{\phi} - \alpha_1\exp(-(\hat{\omega}/\omega_0)^2)$, found whenever $l(\nu, \hat{x}, \hat{\phi}) > 0$. We define $l(\nu, \hat{x}, \hat{\phi})$ as $l(\nu, \hat{x}, \hat{\phi}) = 1$ when $|\hat{\phi}| \geq 1$, and $l(\nu, \hat{x}, \hat{\phi}) = 0$ otherwise, and choose the gain $\Gamma = 1$. According to Proposition 2, this approach is valid, assuming that $|\phi| > 1$ is guaranteed to occur some portion of the time.

Remark 5: For simplicity and due to space constraints, we have ignored a couple of technicalities in this example: we have not used any projections or saturations; and we have defined $l(\nu, \hat{x}, \hat{\phi})$ to be discontinuous in $\hat{\phi}$, thereby breaking the Lipschitz condition in Assumption 4. It is easy to redefine $l(\nu, \hat{x}, \hat{\phi})$ to fix this problem.

B. Simulation Results

We simulate the system with the output y corrupted by noise. The noisy output can be seen together with the actual angular rate ω in the top plot in Fig. 1. We find that $\varepsilon = 1$, which translates to $K(\varepsilon) \approx [6, 0, 6]^{\mathsf{T}}$, ensures stability. The resulting estimates of ω and η are somewhat noisy, not primarily due to the injection term, but due to the use of the noisy output y in constructing $g(\nu, \hat{x}, \hat{\theta})$. To improve the state estimates, we create a second observer that uses the parameter estimate $\hat{\theta}$ from the first observer, as described in Section VI. In the second observer, Lipschitz-like conditions must only be placed on g, and not on the derivative d, and we can therefore implement $g(\nu, \hat{x}, \hat{\theta})$ in the second observer using only the state estimates, and not the noisy output y. We follow the same high-gain methodology to construct the second observer, placing the poles of the matrix H at -1 and -2 and selecting $\varepsilon = 1$, which translates to $K(\varepsilon) \approx [3, 0]^{\mathsf{T}}$. The resulting state estimates can be seen in Fig. 1, together with the parameter estimate.

APPENDIX

Proof of Lemma 1: The error dynamics (4) with $K(\varepsilon) = 0$ consists of a system described by the left-invertible, minimum-phase triple (C, A, E), augmented by adding an integrator at each input point. Because integrators are left-invertible, it follows from the definition of left-invertibility that the augmented system is also left-invertible. The invariant zeros of a left-invertible triple (C, A, E) are the values of z for which the Rosenbrock system matrix $\left[\begin{smallmatrix} zI-A & -E \\ -C & 0 \end{smallmatrix}\right]$ loses rank. It is easy to confirm that when the triple (C, A, E) is augmented with an integrator at each input point, the resulting Rosenbrock system matrix loses rank for precisely the same values of z as before, and hence the invariant zeros (and the minimum-phase property) remain the same. ■

Proof of Lemma 2: This proof is based on the theory of [6]. The proof is stated for the MIMO case, but is valid for SISO systems as a special case. Define $\xi_a = \chi_a$, $\xi_b = \chi_b$, and $\xi_q = \mathrm{col}(\xi_{q_1}, \ldots, \xi_{q_k})$, where for each $j \in 1, \ldots, k$, $\xi_{q_j} := S_j \chi_{q_j}$, where $S_j = \varepsilon^{\bar{n}_q - n_{q_j}} \mathrm{diag}(1, \ldots, \varepsilon^{n_{q_j}-1})$. We then obtain the following equations:

$$\dot{\xi}_a = A_a \xi_a, \quad \dot{\xi}_b = (A_b - K_{bb} C_b)\xi_b, \tag{15a}$$

$$\begin{aligned} \varepsilon \dot{\xi}_{q_j} &= H_j \xi_{q_j} \\ &\quad + B_{q_j} \varepsilon^{\bar{n}_q} (D_{a_j} \xi_a + D_{b_j} \xi_b + \delta_j + D^{\varepsilon}_{q_j} \xi_q), \end{aligned} \tag{15b}$$

where $D^{\varepsilon}_{q_j} = D_{q_j} \mathrm{diag}(S_1^{-1}, \ldots, S_k^{-1})$. Let P_a, P_b, and P_{q_j}, $j = 1, \ldots, k$, be the symmetric, positive-definite solutions of the Lyapunov equations $P_a A_a + A_a^{\mathsf{T}} P_a = -I_{n_a}$, $P_b(A_b - K_{bb}C_b) + (A_b - K_{bb}C_b)^{\mathsf{T}} P_b = -I_{n_b}$, and $P_{q_j} H_j + H_j^{\mathsf{T}} P_{q_j} = -I_{n_{q_j}}$, respectively. Define $W = \xi_a^{\mathsf{T}} P_a \xi_a + \xi_b^{\mathsf{T}} P_b \xi_b + \varepsilon \sum_{j=1}^{k} \xi_{q_j}^{\mathsf{T}} P_{q_j} \xi_{q_j}$. We then have

$$\begin{aligned} \dot{W} &\le -\|\xi_a\|^2 - \|\xi_b\|^2 - \sum_{j=1}^{k} \|\xi_{q_j}\|^2 + \varepsilon^{\bar{n}_q} \sum_{j=1}^{k} 2\|P_{q_j}\| \\ &\times (\|D_{a_j}\|\|\xi_a\| + \|D_{b_j}\|\|\xi_b\| + \|\delta_j\| + \|D^{\varepsilon}_{q_j}\|\|\xi_q\|)\|\xi_{q_j}\|. \end{aligned}$$

From the Lipschitz-like condition on d from Assumption 2, we know that for each $j \in 1, \ldots, k$, there exist constants β_{a_j}, β_{b_j}, β_{θ_j}, and β_{q_j} such that $\|\delta_j\| \le \beta_{a_j}\|\chi_a\| + \beta_{b_j}\|\chi_b\| + \beta_{\theta_j}\|\tilde{\theta}\| + \beta_{q_j}\|\chi_q\|$, which means that $\|\delta_j\| \le \beta_{a_j}\|\xi_a\| + \beta_{b_j}\|\xi_b\| + \beta_{\theta_j}\|\tilde{\theta}\| + \varepsilon^{-(\bar{n}_q-1)}\beta_{q_j}\|\xi_q\|$. We furthermore have $\|D^{\varepsilon}_{q_j}\| \le \varepsilon^{-(\bar{n}_q-1)}\|D_{q_j}\|$.

Let $\rho_a = \sum_{j=1}^{k} 2\|P_{q_j}\|(\|D_{a_j}\| + \beta_{a_j})$, $\rho_b = \sum_{j=1}^{k} 2\|P_{q_j}\|(\|D_{b_j}\| + \beta_{b_j})$, $\rho_q = \sum_{j=1}^{k} 2\|P_{q_j}\|(\|D_{q_j}\| + \beta_{q_j})$, and $\rho_\theta = \sum_{j=1}^{k} 2\|P_{q_j}\|\beta_{\theta_j}$. Then we may write

$$\begin{aligned} \dot{W} &\le -\|\xi_a\|^2 - \|\xi_b\|^2 - (1 - \varepsilon\rho_q)\|\xi_q\|^2 \\ &+ \varepsilon^{\bar{n}_q}\rho_a\|\xi_a\|\|\xi_q\| + \varepsilon^{\bar{n}_q}\rho_b\|\xi_b\|\|\xi_q\| + \varepsilon^{\bar{n}_q}\rho_\theta\|\tilde{\theta}\|\|\xi_q\|. \end{aligned}$$

Note that ρ_q is multiplied by ε. Furthermore, note that the cross terms between $\|\xi_a\|$ and $\|\xi_q\|$, and between $\|\xi_b\|$ and $\|\xi_q\|$, are multiplied by $\varepsilon^{\bar{n}_q}$ relative to the stabilizing quadratic terms in $\|\xi_a\|^2$, $\|\xi_b\|^2$, and $\|\xi_q\|^2$. It is therefore straightforward to show that, by decreasing ε, the cross terms are dominated, and there exist positive constants c_a, c_b, and c_q such that $\dot{W} \le -c_a\|\xi_a\|^2 - c_b\|\xi_b\|^2 - c_q\|\xi_q\|^2 + \varepsilon^{\bar{n}_q}\rho_\theta\|\tilde{\theta}\|\|\xi_q\|$. This expression shows that $\dot{W}$ is negative outside a ball around the origin, the size of which is proportional to $\tilde{\theta}$. From [25, Th. 4.19], we conclude that (15) is ISS with respect to $\tilde{\theta}$. Since (15) is obtained through a nonsingular transformation of (4), the same holds for (4). ■

Proof of Theorem 1: We first note that Assumption 3 ensures that if $\hat{\theta}(t_0) \in \Theta$, then for all $t \in \mathbb{R}_{\ge 0}$, $\hat{\theta}(t) \in \Theta$. From Lemma 2, the error dynamics (4) is ISS with respect to $\tilde{\theta}$. Hence, any trajectory of (4), (12) originating in $\mathbb{R}^{n+k} \times (\Theta - \Theta)$ remains in a compact subset of $\mathbb{R}^{n+k} \times (\Theta - \Theta)$ for all future time. In the remainder of this proof, we shall deal with the transformed system (15) from the proof of Lemma 2, and prove that this system, together with (12), is exponentially stable with $\mathbb{R}^{n+k} \times (\Theta - \Theta)$ contained in the region of attraction. This will imply that the theorem holds with respect to the error dynamics (4), (12). Based on Assumption 4, we know, by following the same argument as for $\|\delta_j\|$ in the proof of Lemma 2, that there exist positive constants $\bar{\beta}_a$, $\bar{\beta}_b$, $\bar{\beta}_q$ such that $\|u_\theta(\nu, x, \phi, \hat{\theta}) - u_\theta(\nu, \hat{x}, \hat{\phi}, \hat{\theta})\| \le \bar{\beta}_a\|\xi_a\| + \bar{\beta}_b\|\xi_b\| + \varepsilon^{-(\bar{n}_q-1)}\bar{\beta}_q\|\xi_q\|$. Define $V := W + \varepsilon^{2\bar{n}_q-1}V_u$, where W is from the proof of Lemma 2. We then obtain

$$\begin{aligned} \dot{V} &\le -c_a\|\xi_a\|^2 - c_b\|\xi_b\|^2 - c_q\|\xi_q\|^2 - a_3\varepsilon^{2\bar{n}_q-1}\|\tilde{\theta}\|^2 \\ &+ \varepsilon^{\bar{n}_q}[(\rho_\theta + a_4\bar{\beta}_q)\|\xi_q\| + a_4\bar{\beta}_a\|\xi_a\| + a_4\bar{\beta}_b\|\xi_b\|]\|\tilde{\theta}\|. \end{aligned}$$

To show that we may dominate all the cross terms in the expression above by decreasing ε, note that for arbitrary positive constants α_1, α_2, and α_3, the expression $-\alpha_1\zeta_1^2 - \alpha_2\varepsilon^{2\bar{n}_q-1}\zeta_2^2 + \alpha_3\varepsilon^{\bar{n}_q}\zeta_1\zeta_2$ can be made negative definite by selecting $\varepsilon < 4\alpha_1\alpha_2/\alpha_3^2$. Since we may split out expressions like the one above for each of the cross terms (by letting α_1 and α_2 be small fractions of the negative quadratic terms, and α_3 the constant in the cross term) and make them negative definite by decreasing ε, it is clear that there exists ε^* and a constant $c > 0$ such that for all $0 < \varepsilon < \varepsilon^*$, $\dot{V} \le -c(\|\xi_a\|^2 + \|\xi_b\|^2 + \|\xi_q\|^2 + \|\tilde{\theta}\|^2)$. This expression holds on the set $\mathbb{R}^{n+k} \times (\Theta - \Theta)$. By invoking the comparison lemma [25, Lemma 3.4], we may therefore conclude that the origin of (15), (12) is exponentially stable with $\mathbb{R}^{n+k} \times (\Theta - \Theta)$ contained in the region of attraction, and the same holds for (4), (12). ■

Proof of Proposition 2: For ease of notation, we write $l(t) = l(\nu(t), x(t), \phi(t))$. Inspired by [26] we use the function $V_{\mathrm{u}} = \frac{1}{2}\tilde{\theta}^{\mathsf{T}}\left(\Gamma^{-1} - \mu\int_t^\infty \mathrm{e}^{t-\tau} I_p l(\tau)\,\mathrm{d}\tau\right)\tilde{\theta}$, where $\mu > 0$ is a constant yet to be specified. We first note that $\frac{1}{2}\tilde{\theta}^{\mathsf{T}}\left(\Gamma^{-1} - \mu I_p\right)\tilde{\theta} \le V_{\mathrm{u}} \le \frac{1}{2}\tilde{\theta}^{\mathsf{T}}\Gamma^{-1}\tilde{\theta}$. Hence, V_{u} is positive definite provided $\mu < \lambda_{\min}(\Gamma^{-1})$. With $\hat{\phi} = \phi$ and $\hat{x} = x$, we get $\dot{\tilde{\theta}} = -\mathrm{Proj}(l(t)\Gamma\tilde{\theta})$. Using the property [9, Lemma

E.1] that $-\tilde{\theta}^\top\Gamma^{-1}\mathrm{Proj}(\tau) \le -\tilde{\theta}^\top\Gamma^{-1}\tau$, we have

$$\begin{aligned}\dot{V}_{\mathrm{u}} &= \frac{1}{2}\mu\tilde{\theta}^\top I_p l(t)\tilde{\theta} - \frac{1}{2}\mu\tilde{\theta}^\top \int_t^\infty \mathrm{e}^{t-\tau} I_p l(\tau)\,\mathrm{d}\tau\,\tilde{\theta} \\ &\quad - \tilde{\theta}^\top\left(\Gamma^{-1} - \mu\int_t^\infty \mathrm{e}^{t-\tau} I_p l(\tau)\,\mathrm{d}\tau\right)\mathrm{Proj}(l(t)\Gamma\tilde{\theta}) \\ &\quad \le -(1-\frac{1}{2}\mu)l(t)\tilde{\theta}^\top\tilde{\theta} - \frac{1}{2}\mu\sigma\mathrm{e}^{-T}\tilde{\theta}^\top\tilde{\theta} \\ &\quad + \mu\|\tilde{\theta}\|\left\|\int_t^\infty \mathrm{e}^{t-\tau} I_p l(\tau)\,\mathrm{d}\tau\right\|\left\|\mathrm{Proj}(l(t)\Gamma\tilde{\theta})\right\| \\ &\le -(1-\frac{1}{2}\mu - \mu\sqrt{\kappa}\|\Gamma\|)l(t)\|\tilde{\theta}\|^2 - \frac{1}{2}\mu\sigma\mathrm{e}^{-T}\|\tilde{\theta}\|^2, \quad (16)\end{aligned}$$

where κ is the ratio of the largest to the smallest eigenvalue of Γ^{-1}. Above, we have used the property [9, Lemma E.1] that $\mathrm{Proj}(\tau)^\top\Gamma^{-1}\mathrm{Proj}(\tau) \le \tau^\top\Gamma^{-1}\tau$, which implies that $\|\mathrm{Proj}(\tau)\| \le \sqrt{\kappa}\|\tau\|$. We have also used that $\int_t^\infty \mathrm{e}^{t-\tau}l(\tau)\,\mathrm{d}\tau \ge \int_t^{t+T}\mathrm{e}^{t-\tau}l(\tau)\,\mathrm{d}\tau \ge \mathrm{e}^{-T}\int_t^{t+T} l(\tau)\,\mathrm{d}\tau \ge \mathrm{e}^{-T}\sigma$. From the calculation above, we see that the time derivative is negative definite provided $\mu < 1/(\frac{1}{2} + \sqrt{\kappa}\|\Gamma\|)$. ■

Proof of Proposition 3: For the sake of brevity, we write $M = M(\nu, x, \hat{\theta})$. With $\hat{\phi} = \phi$ and $\hat{x} = x$, we get $\dot{\hat{\theta}} = -\mathrm{Proj}(\Gamma M(g(\nu,x,\theta) - g(\nu,x,\hat{\theta})))$. We use the function $V_{\mathrm{u}} = \frac{1}{2}\tilde{\theta}^\top\Gamma^{-1}\tilde{\theta}$. Using the property [9, Lemma E.1] that $-\tilde{\theta}^\top\Gamma^{-1}\mathrm{Proj}(\tau) \le -\tilde{\theta}^\top\Gamma^{-1}\tau$, we have $\dot{V}_{\mathrm{u}} \le -\frac{1}{2}\tilde{\theta}^\top M(g(\nu,x,\theta) - g(\nu,x,\hat{\theta})) - \frac{1}{2}(g(\nu,x,\theta) - g(\nu,x,\hat{\theta}))^\top M^\top\tilde{\theta}$. Since $g(\nu,x,\theta)$ is continuously differentiable with respect to θ, we may write, according to Taylor's theorem (see, e.g., [27, Theorem 11.1]), $g(\nu,x,\theta) - g(\nu,x,\hat{\theta}) = \int_0^1 [\partial g/\partial\theta](\nu,x,\hat{\theta}+p\tilde{\theta})\tilde{\theta}\,\mathrm{d}p$. Hence, we have $\dot{V}_{\mathrm{u}} \le -\frac{1}{2}\int_0^1 \tilde{\theta}^\top(M[\partial g/\partial\theta](\nu,x,\hat{\theta}+p\tilde{\theta}) + [\partial g/\partial\theta](\nu,x,\hat{\theta}+p\tilde{\theta})^\top M^\top)\tilde{\theta}\,\mathrm{d}p \le -\int_0^1 \tilde{\theta}^\top P\tilde{\theta}\,\mathrm{d}p = -\tilde{\theta}^\top P\tilde{\theta}$, which proves that Assumption 3 holds. ■

Proof of Proposition 4: We use the function $V_{\mathrm{u}} = \frac{1}{2}\tilde{\theta}^\top\left(\Gamma^{-1} - \mu\int_t^\infty \mathrm{e}^{t-\tau}S(\nu(\tau),x(\tau))\,\mathrm{d}\tau\right)\tilde{\theta}$, where $\mu > 0$ is a constant yet to be specified. First, we confirm that V_{u} is positive definite. We have $\frac{1}{2}(\lambda_{\min}(\Gamma^{-1}) - \mu\lambda'_S)\|\tilde{\theta}\|^2 \le V_{\mathrm{u}} \le \frac{1}{2}\lambda_{\min}(\Gamma^{-1})\|\tilde{\theta}\|^2$, where $\lambda'_S = \sup_{(\nu,x)\in V\times X}\lambda_{\max}(S(\nu,x))$. It follows that V_{u} is positive definite provided $\lambda_{\min}(\Gamma^{-1}) - \mu\lambda'_S > 0$, which is guaranteed if $\mu < \lambda_{\min}(\Gamma^{-1})/\lambda'_S$. When we insert $\hat{\phi} = \phi$ and $\hat{x} = x$, we get the same error dynamics as in the proof of Proposition 3. Following a calculation similar to the proof of Proposition 2, we get $\dot{V}_{\mathrm{u}} \le -(1-\frac{1}{2}\mu)\tilde{\theta}^\top S(\nu,x)\tilde{\theta} - \frac{1}{2}\mu\sigma\mathrm{e}^{-T}\|\tilde{\theta}\|^2 + \mu\sqrt{\kappa}M_S\|\Gamma\|M_M L_g\|\tilde{\theta}\|(\tilde{\theta}^\top S(\nu,x)\tilde{\theta})^{1/2}$, where M_S and M_M are bounds on $\|S(\nu,x)\|$ and $\|M(\nu,x,\hat{\theta})\|$, respectively, and κ is ratio of the largest to the smallest eigenvalue of Γ^{-1}. We may write this as a quadratic expression with respect to $\mathrm{col}((\tilde{\theta}^\top S(\nu,x)\tilde{\theta})^{1/2}, \|\tilde{\theta}\|)$. It is then easily confirmed that $\dot{V}_{\mathrm{u}}$ is negative definite if $\mu < 2/(1 + \kappa M_S^2\|\Gamma\|^2 M_M^2 L_g^2\sigma^{-1}\mathrm{e}^T)$. ■

References

[1] H. F. Grip, T. A. Johansen, and L. Imsland, "Estimation and control for systems with nonlinearly parameterized perturbations," in *Proc. 17th IFAC World Congress*, Seoul, South Korea, 2008.

[2] H. F. Grip, T. A. Johansen, L. Imsland, and G.-O. Kaasa, "Parameter estimation and compensation in systems with nonlinearly parameterized perturbations," *Automatica*, 2010, to appear.

[3] H. Khalil and F. Esfandiari, "Semiglobal stabilization of a class of nonlinear systems using output feedback," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 1412–1415, 1993.

[4] A. Teel and L. Praly, "Global stabilizability and observability imply semi-global stabilizability by output feedback," *Syst. Contr. Lett.*, vol. 22, pp. 313–325, 1994.

[5] P. Sannuti and A. Saberi, "Special coordinate basis for multivariable linear systems—finite and infinite zero structure, squaring down and decoupling," *Int. J. Contr.*, vol. 45, no. 5, pp. 1655–1704, 1987.

[6] A. Saberi and P. Sannuti, "Observer design for loop transfer recovery and for uncertain dynamical systems," *IEEE Trans. Automat. Contr.*, vol. 35, no. 8, pp. 878–897, 1990.

[7] A. Chakrabortty and M. Arcak, "Time-scale separation redesigns for stabilization and performance recovery of uncertain nonlinear systems," *Automatica*, vol. 45, no. 1, pp. 34–44, 2009.

[8] L. B. Freidovich and H. K. Khalil, "Performance recovery of feedback-linearization-based designs," *IEEE Trans. Automat. Contr*, vol. 53, no. 10, pp. 2324–2334, 2008.

[9] M. Krstić, I. Kanellakopoulos, and P. Kokotović, *Nonlinear and Adaptive Control Design*. New York: Wiley, 1995.

[10] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, 3rd ed. New York: Wiley, 1997.

[11] V. Fomin, A. Fradkov, and V. Yakubovich, Eds., *Adaptive Control of Dynamical Systems*. Moskow: Nauka, 1981.

[12] R. Ortega, "Some remarks on adaptive neuro-fuzzy systems," *Int. J. Adapt. Contr. Signal Process.*, vol. 10, pp. 79–83, 1996.

[13] A. M. Annaswamy, F. P. Skantze, and A.-P. Loh, "Adaptive control of continuous time systems with convex/concave parametrization," *Automatica*, vol. 34, no. 1, pp. 33–49, 1998.

[14] A.-P. Loh, A. M. Annaswamy, and F. P. Skantze, "Adaptation in the presence of a general nonlinear parameterization: An error model approach," *IEEE Trans. Automat. Contr.*, vol. 44, no. 9, pp. 1634–1652, 1999.

[15] J. D. Bošković, "Stable adaptive control of a class of first-order nonlinearly parameterized plants," *IEEE Trans. Automat. Contr.*, vol. 40, no. 2, pp. 347–350, 1995.

[16] ——, "Adaptive control of a class of nonlinearly parameterized plants," *IEEE Trans. Automat. Contr.*, vol. 43, no. 7, pp. 930–934, 1998.

[17] T. Zhang, S. S. Ge, C. C. Hang, and T. Y. Chai, "Adaptive control of first-order systems with nonlinear parameterization," *IEEE Trans. Automat. Contr.*, vol. 45, no. 8, pp. 1512–1516, 2000.

[18] Z. Qu, "Adaptive and robust controls of uncertain systems with nonlinear parameterization," *IEEE Trans. Automat. Contr.*, vol. 48, no. 10, pp. 1817–1823, 2003.

[19] I. Y. Tyukin, D. V. Prokhorov, and C. van Leeuwen, "Adaptation and parameter estimation in systems with unstable target dynamics and nonlinear parametrization," *IEEE Trans. Automat. Contr.*, vol. 52, no. 9, pp. 1543–1559, 2007.

[20] H. K. Khalil, "Adaptive output feedback control of nonlinear systems represented by input-output models," *IEEE Trans. Automat. Contr.*, vol. 41, no. 2, pp. 177–188, 1996.

[21] S. Nicosia, A. Tornambè, and P. Valigi, "Nonlinear map inversion via state observers," *Circ. Syst. Signal Process.*, vol. 13, no. 5, pp. 571–589, 1994.

[22] S. Sastry, J. Hauser, and P. Kokotović, "Zero dynamics of regularly perturbed systems may be singularly perturbed," *Syst. Contr. Lett.*, vol. 13, no. 4, pp. 299–314, 1989.

[23] C. Canudas de Wit and P. Lischinsky, "Adaptive friction compensation with partially known dynamic friction model," *Int. J. Adapt. Contr. Signal Process.*, vol. 11, pp. 65–80, 1997.

[24] H. F. Grip and A. Saberi, "Structural decomposition of linear multivariable systems using symbolic computations," in *Proc. Asian Contr. Conf.*, Hong Kong, 2009.

[25] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2002.

[26] A. Loría, E. Panteley, D. Popović, and A. R. Teel, "A nested Matrosov theorem and persistency of excitation for uniform convergence in stable nonautonomous systems," *IEEE Trans. Automat. Contr.*, vol. 50, no. 2, pp. 183–198, 2005.

[27] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York: Springer-Verlag, 1999.

Corrigendum to “State and Parameter Estimation for Linear Systems with Nonlinearly Parameterized Perturbations”

Håvard Fjær Grip, Ali Saberi, and Tor A. Johansen

In Section III-B.2, there is a scaling error in the expression for $K'_{q_j q}$. The correct expression is

$$K'_{q_j q} = \left[0_{n_{q_j} \times (j-1)} \quad \operatorname{col}\left(\frac{\bar{K}_{q_j 1}}{\varepsilon}, \dots, \frac{\bar{K}_{q_j n_{q_j}}}{\varepsilon^{n_{q_j}}} \right) \quad 0_{n_{q_j} \times (k-j)} \right].$$

Left uncorrected, the error affects MIMO systems of non-uniform rank, but not SISO systems or MIMO systems of uniform rank.