Parameter Identification for Uncertain Linear Systems with Partial State Measurements Under an H^∞ Criterion

Zigang Pan, *Member, IEEE*, and Tamer Başar, *Fellow, IEEE*

***Abstract*—This paper addresses the worst-case parameter identification problem for uncertain single-input/single-output (SISO) and multi-input/multi-output (MIMO) linear systems under partial state measurements and derives worst-case identifiers using the *cost-to-come* function method. In the SISO case, the worst-case identifier obtained subsumes the Kreisselmeier observer as part of its structure with parameters set at some optimal values. Its structure is different from the common least-squares (LS) identifier, however, in the sense that there is additional dynamics for the state estimate, coupled with the dynamics of the parameter estimate in a nontrivial way. In the MIMO case as well, the worst-case identifier has additional dynamics for the state estimate which do not appear in the conventional LS-based schemes. Also for both SISO and MIMO problems, approximate identifiers are obtained which are numerically much better conditioned when the disturbances in the measurement equations are "small." The theoretical results are then illustrated on an extensive numerical example to demonstrate the effectiveness of the identification schemes developed.**

I. Introduction

Parameter identification is a vital part of any successful design for controlling systems with unknown parameters. For continuous-time linear systems, the Kalman filter-based identifier is the prominent design method for online parameter identification. When the state variable and its derivative are both available for identification purposes, the Kalman filter-based identification scheme can be applied directly [1]. On the other hand, if the available information contains only noise-corrupted output measurements, the identification scheme must utilize the well-known Kreisselmeier observer in the single-input/single-output (SISO) case [2] or a general prefiltering-based design for the multi-input/multi-output (MIMO) case [3], [4] so that the problem can be converted to one that can be solved by a Kalman filter-based method.

Recently, worst-case parameter identification for general nonlinear systems where the unknown parameters enter linearly in the system dynamics has been studied in [5] under noisy full state measurements. It has been shown there that the worst-case parameter identification problem without the measurement of the state derivative can be dealt with by introducing small measurement noise in the optimization process. The resulting identification scheme can achieve a performance level arbitrarily close to the one where the state derivative is available, as the intensity of the measurement noise, say $\epsilon > 0$, decreases to zero. Also, in [5] a reduced-order identifier was constructed which is much simpler in structure and has asymptotically the same performance as that of the full-order worst-case identifier. The structure of the worst-case identifier closely resembles that of a least-squares (LS) identifier except for additional state estimate dynamics and an extra negative definite term in the differential equation for the error covariance matrix.

In this paper, we study the worst-case parameter identification problem under *partial* state measurements, but for *linear* systems only. We construct worst-case (full-order) identifiers as well as their appropriate approximations which are numerically better conditioned for implementation. We will first study the SISO case in detail, which will allow us to introduce the solution concept, and discuss the main ideas in the derivation and verification of full-order and reduced-order identifiers without introducing cumbersome notation (which unfortunately cannot be escaped from in the MIMO case). By representing the system in output injection form in state space, we can apply the *cost-to-come function* method [6]–[9] as in [5] to obtain the full-order worst-case identifier, provided that the desired performance level is achievable. It will be shown that the worst-case identifier contains the Kreisselmeier observer as part of its structure. Our derivation leads to an optimum choice for the observer parameters which can be obtained from the solution of an algebraic Riccati equation. As in [5], the worst-case identifier structure obtained contains additional dynamics for the state estimates which are not present in the conventional LS identifiers. As the noise intensity in the measurements decreases, the full-order identifier becomes numerically ill-conditioned. To alleviate this, we construct an approximate (reduced-order) identifier that is simpler in structure by two integrators and is numerically much better conditioned for implementation. In contradistinction to the LS or LMS algorithms, where the cost functions are restricted to some fixed quadratic weight on the identification error, the worst-case identifier presented here is derived under an

Manuscript received August 4, 1995; revised March 14, 1996. Recommended by Associate Editor, V. Solo. This work was supported in part by the U.S. Department of Energy under Grant DOE-DEFG-02-94ER13939 and in part by the National Science Foundation under Grant NSF-ECS-93-12807.

Z. Pan is with the Center for Control Engineering and Computation, Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA.

T. Başar is with the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801 USA (e-mail: tbasar@decision.csl.uiuc.edu).

Publisher Item Identifier S 0018-9286(96)06778-5.

arbitrary quadratic cost function. This freedom in the choice of the cost function permits further fine-tuning of the identifier to any specific application at hand. Since obtained under a worst-case analysis, the identifiers presented here possess guaranteed robustness properties which may not be present in LS or LMS identifiers. In particular, they can tolerate unmodeled dynamics and generate parameter estimates which satisfy prespecified attenuation bounds. Barring the complexity in notation, these results are easily generalizable to the MIMO case which is done here without providing detailed proofs.

The balance of the paper is organized as follows. In the next section, we provide a precise problem formulation for worst-case identification in SISO linear systems. Then, we briefly present in Section III, the well-known Kreisselmeier observer design that allows for the application of Kalman filter-based identification schemes. In Section IV, a worst-case identifier for the SISO problem is derived, and two special cases are discussed where the structure of the identifier can be simplified. We prove the optimality of the identifiers derived and discuss the similarities as well as the differences between these and the conventional ones presented in the previous section. Counterparts of these results in the MIMO case are summarized in Section V. An extensive simulation study of a third-order five-parameter SISO example is presented in Section VI to illustrate the theory. The paper ends with the concluding remarks of Section VII.

II. Problem Formulation

We consider the class of SISO linear systems described by an nth order transfer function

$$H(s) = \frac{b_m s^m + \cdots + b_0}{s^n + a_{n-1}s^{n-1} + \cdots a_0}$$

where $m < n$, and all coefficient terms are unknown. We write this system in the state-space representation form as

$$\dot{x} = A_2 x - ay + bu \tag{1a}$$

$$y = e_1' x \tag{1b}$$

where e_i denotes the ith coordinate vector in $\mathbb{R}^n$

$$A_2 = \begin{bmatrix} \mathbf{0}_{(n-1)\times 1} & I_{n-1} \\ 0 & \mathbf{0}_{1\times(n-1)} \end{bmatrix}$$

$$a = \begin{bmatrix} a_{n-1} \\ \vdots \\ a_0 \end{bmatrix} \qquad b = \begin{bmatrix} \mathbf{0}_{(n-m-1)\times 1} \\ b_m \\ \vdots \\ b_0 \end{bmatrix}$$

I_i denotes the $i \times i$ identity matrix (the subscript will be suppressed if $i = n$), and $\mathbf{0}_{i\times j}$ denotes the $i \times j$ matrix of zeros.

To study the robustness of an identification scheme associated with this system, we will take the plant dynamics to be further perturbed by an unknown additive noise w, leading to the following general model:

$$\dot{x} = A_2 x + (A_{11} y + A_{12} u)\theta + w; \quad x(0) = x_0 \tag{2a}$$

$$y = e_1' x. \tag{2b}$$

Here θ is an r-dimensional unknown constant parameter, A_{11} and A_{12} are some known constant matrices (both of dimensions $n \times r$), A_2 is as defined above, and w is an n-dimensional unknown system noise whose components belong to $\mathcal{L}^\infty$. Note that the allowable class of disturbances is fairly general and includes in particular noises generated by unmodeled dynamics.

For the particular SISO system described above, the appropriate parameters and coefficient matrices are

$$\theta = \begin{bmatrix} -a \\ b \end{bmatrix}; \quad A_{11} = [I \quad \mathbf{0}_{n\times(m+1)}]$$

$$A_{12} = \begin{bmatrix} \mathbf{0}_{(n-m-1)\times n} & \mathbf{0}_{(n-m-1)\times(m+1)} \\ \mathbf{0}_{(m+1)\times n} & I_{m+1} \end{bmatrix}$$

and $r = m + 1 + n$. We will in fact study this problem with general (known) matrices A_{11} and A_{12} to allow for possible additional *a priori* information on the plant.

Even though this is the identification problem we wish to address in this paper, using a particular worst-case optimization technique (to be elucidated below), the singular nature of the underlying optimization problem forces us to embed this original problem in a larger class of parameterized identification problems, as in [5]. These are obtained by replacing the original measurement y by one that is contaminated by some small measurement noise, ϵv, where ϵ is a small positive scalar and v is an unknown $\mathcal{L}^\infty$ function. Hence this more general class, parameterized by ϵ, is described by the state and measurement equations

$$\dot{x} = A_2 x + (A_{11} y + A_{12} u)\theta + w; \quad x(0) = x_0 \tag{3a}$$

$$y = e_1' x + \epsilon v. \tag{3b}$$

Clearly, the original problem can be recovered as a special case of this more general one by either setting $\epsilon = 0$ or taking the measurement noise v to be identically zero. Furthermore, any robustness result (such as satisfaction of a disturbance attenuation bound) to be derived for this problem will apply (with additional margin of robustness) to the original problem. Of course, this measurement noise-perturbed problem would also be of independent interest, as any identification scheme developed for it would exhibit robustness to unmodeled measurement inaccuracies. It should be noted that when the measurement noise v is actually present, then the substitution of $x_1 \times y$ in the system dynamics (3a) is based on the understanding that the disturbance w includes, in part, the measurement noise v.

Henceforth, we work with the general perturbed model (3a) and (3b). For this system, we seek an estimate $\hat{\theta}(t)$ of the unknown parameter vector θ to be generated by

$$\hat{\theta}(t) = \delta(t, y_{[0,t]}, u_{[0,t]}) \tag{4}$$

where δ is an identifier (equivalently, estimator—yet to be determined) which is piecewise continuous in t and Lipschitz continuous in $y_{[0,t]}$ and $u_{[0,t]}$. Let us denote the class of these admissible identifiers by Δ.

Our objective, succinctly stated, is to find an identifier $\delta \in \Delta$ that minimizes an appropriate norm of the identification error under the worst choices of the uncertainty quadruple $(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)})$. One way of formulating this objective is to view the identification problem as a disturbance attenuation problem (as in H^∞ control), where the disturbance to be attenuated is the uncertainty quadruple above, and the output is the identification error (see [5] for further motivation for this approach to identification). Accordingly, we introduce as a natural candidate the performance index

$$L(\delta) = \sup_{x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}} \cdot \frac{\|\theta - \hat{\theta}(t)\|^2_{Q(t, y_{[0,t]}, u_{[0,t]})}}{\|w\|^2 + \|v\|^2 + |\theta - \overline{\theta}_0|^2_{Q_0} + |x_0 - \overline{x}_0|^2_{P_0}} \tag{5}$$

where $\|\cdot\|_R$ denotes an appropriate dimensional $\mathcal{L}^2$ (semi-) norm, weighted by R, and $|\cdot|_P$ denotes a Euclidean norm, weighted by P. We assume that $Q(\cdot,\cdot,\cdot) \geq 0$, $P_0 > 0$, and $Q_0 > 0$. $\overline{\theta}_0$ and $\overline{x}_0$ are some initial estimates for θ and x_0, respectively. The weighting matrix P_0 will be chosen in such a way as to reduce the order of the optimal identifier to be obtained, as to be discussed later.

The optimal performance level is the quantity γ^* defined by

$$\gamma^* := \inf_{\delta \in \Delta} \{L(\delta)\}^{1/2}. \tag{6}$$

As well known in linear and nonlinear H^∞ control [10], we can associate with this system a class of soft-constrained differential games, indexed by a parameter $\gamma > 0$, with a cost function

$$\begin{aligned} &J_\gamma(\delta; x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}) = \\ &\int_0^\infty \{|\theta - \delta|^2_{Q(t, y_{[0,t]}, u_{[0,t]})} - \gamma^2[|w(t)|^2 + |v(t)|^2]\}\, dt \\ &-\gamma^2(|\theta - \overline{\theta}_0|^2_{Q_0} + |x_0 - \overline{x}_0|^2_{P_0}) \end{aligned} \tag{7}$$

where δ is the minimizer and the quadruple $(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)})$ the maximizer. By [10], the quantity γ^* is the "smallest" value of γ such that this game has a zero upper value. Given an achievable performance level $\gamma > 0$, we will say that a particular estimator $\overline{\delta} \in \Delta$ achieves that level, if

$$\sup_{x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}} J_\gamma(\overline{\delta}; x_0, \theta_0, w_{[0,\infty)}, v_{[0,\infty)}) = 0. \tag{8}$$

This completes the formulation of the worst-case identification problem. We will present the solution to this problem in Section IV which will be parameterized by $\epsilon > 0$. In the same section, we will also study the limiting case as ϵ approaches zero which captures formally the situation when the measurement is almost noise free—the problem of real interest to us. To be able to compare our results with the existing identification schemes based on the Kreisselmeier observer design, we first present, in the next section, an overview of the latter.

III. Kreisselmeier Observer Design

For the identification of the parameter vectors a and b in (1), a general methodology was developed in [2], where stable prefilters for both the measurement y and control input u were introduced so as to reduce the problem to one where the LS method can be directly applied. These prefilters are known as the Kreisselmeier observer.

The prefilters are constructed as follows. First choose an n-dimensional vector f such that the matrix

$$A_f := A_2 + [f \quad \mathbf{0}_{n\times(n-1)}] \tag{9}$$

is Hurwitz. Then, filter the signals y and u through two n-dimensional prefilters

$$\dot{\eta} = A_f\eta + e_n y; \quad \eta(0) = 0 \tag{10}$$

$$\dot{\lambda} = A_f\lambda + e_n u; \quad \lambda(0) = 0. \tag{11}$$

As a result of this prefiltering operation, the state variable x for the system (1a) and (1b) satisfies the following algebraic equation in terms of η and λ:

$$x = -\sum_{i=0}^{n-1}(f_i + a_i)A_f^i\eta + \sum_{i=0}^{n-1} b_i A_f^i\lambda + \varepsilon$$

where A_f^i is the ith power of A_f, and ε satisfies the linear differential equation

$$\dot{\varepsilon} = A_f\varepsilon; \quad \varepsilon(0) = x_0.$$

Since the matrix A_f is stable, the term ε goes to zero exponentially fast as $t \to \infty$. Ignoring this exponentially decaying term, we have

$$y = -\sum_{i=0}^{n-1}(f_i + a_i)e_1' A_f^i\eta + \sum_{i=0}^{n-1} b_i e_1' A_f^i\lambda.$$

Hence, the parameter vectors a and b can be identified from the above equality using the standard LS method.

Although the Kreisselmeier observer provides a method to solve the SISO continuous-time identification problem, the choice of the vector f is still quite arbitrary apart from the requirement that A_f be stable, and it does not address any optimality property of the resulting identification scheme, nor its robustness to inaccuracies in modeling.

In the next section, we study the worst-case identification problem formulated in Section II. A byproduct of this analysis will be a verification of the fact that the Kreisselmeier observer structure with the parameters f fixed at certain "optimal" values is part of a minimax identifier.

IV. Worst-Case Identification

To apply the general framework of *affine quadratic minimax controller design* [5], [6] to the problem at hand, we first associate with (3a) and (3b) the natural simple dynamics

$$\dot{\theta} = 0.$$

In terms of $\xi := (\theta', x')'$, the system is now described by the following dynamic equations:

$$\begin{aligned} \dot{\xi} &= \begin{bmatrix} \mathbf{0}_{r\times r} & \mathbf{0}_{r\times n} \\ yA_{11} + uA_{12} & A_2 \end{bmatrix}\xi + \begin{bmatrix} \mathbf{0}_{r\times n} \\ I \end{bmatrix} w \\ &=: \overline{A}\xi + \overline{D}w; \qquad \xi(0) = \xi_0 := [\theta \quad x_0] \end{aligned} \tag{12}$$

$$\begin{aligned} y &= [\mathbf{0}_{1\times r} \quad e_1']\xi + \epsilon v \\ &=: \overline{C}\xi + \epsilon v. \end{aligned} \tag{13}$$

The soft-constrained game cost function, J_γ, can similarly be expressed as follows, in terms of the state variable ξ:

$$\begin{aligned} J_\gamma(\hat\xi;\, \xi_0,\, w_{[0,\infty)},\, v_{[0,\infty)}) = & \\ \int_0^\infty & \{|\xi-\hat\xi|^2_{\overline{Q}} - \gamma^2[|w(t)|^2+|v(t)|^2]\}\, dt \\ & - \gamma^2|\xi_0-\overline{\xi}_0|^2_{\overline{Q}_0} \end{aligned} \tag{14}$$

where

$$\overline{Q} := \begin{bmatrix} Q & \mathbf{0}_{r\times n} \\ \mathbf{0}_{n\times r} & \mathbf{0}_{n\times n}\end{bmatrix};\quad \overline{Q}_0 := \begin{bmatrix} Q_0 & \mathbf{0}_{r\times n} \\ \mathbf{0}_{n\times r} & P_0\end{bmatrix}$$

$$\hat\xi := \begin{bmatrix}\hat\theta \\ \hat x\end{bmatrix};\quad \overline{\xi}_0 := \begin{bmatrix}\overline{\theta}_0 \\ \overline{x}_0\end{bmatrix}$$

and $\hat x$ denotes an estimate for the state variable x, and $\hat\theta$ denotes an estimate for θ.

For the minimax problem (12)–(14), we introduce the *cost-to-come function* [6]–[9]

$$\begin{aligned} & W_\gamma(t,\, \xi,\, \hat\xi_{[0,t]},\, y_{[0,t]},\, u_{[0,t]}) := \\ & \quad \max_{\xi_0,\, w_{[0,t]},\, v_{[0,t]} | y_{[0,t]},\, u_{[0,t]},\, \xi(t)=\xi} \\ & \cdot \int_0^t \left[|\xi-\hat\xi|^2_{\overline{Q}} - \gamma^2\left(|w|^2 + \frac{1}{\epsilon^2}|y-\overline{C}\xi|^2\right)\right] dt \\ & - \gamma^2|\xi_0-\overline{\xi}_0|^2_{\overline{Q}_0} \end{aligned} \tag{15}$$

where the max operation is over all initial conditions ξ_0 and disturbance trajectories $w_{[0,t]}$ which, along with the input history $u_{[0,t]}$, generate the output $y_{[0,t]}$ and ensure that the terminal state satisfies the boundary condition $\xi(t)=\xi$. Using dynamic programming, this cost-to-come function can be rewritten in the simpler quadratic form

$$\begin{aligned} W_\gamma(t,\, \xi,\, \hat\xi_{[0,t]},\, y_{[0,t]},\, u_{[0,t]}) &= -m(t;\, \gamma,\, y_{[0,t]},\, u_{[0,t]}) \\ & - \gamma^2|\xi(t)-\overline{\xi}(t)|^2_{\overline{\Sigma}(t;\,\gamma,\, y_{[0,t]},\, u_{[0,t]})} \end{aligned} \tag{16}$$

where

$$\dot{\overline{\Sigma}} = -\overline{\Sigma}\,\overline{A} - \overline{A}'\overline{\Sigma} - \overline{\Sigma}\,\overline{D}\,\overline{D}'\overline{\Sigma} + \frac{1}{\epsilon^2}\overline{C}'\overline{C} - \frac{1}{\gamma^2}\overline{Q}$$

$$\overline{\Sigma}(0) = \overline{Q}_0 \tag{17}$$

$$\dot{\overline{\xi}} = \overline{A}\overline{\xi} + \frac{1}{\epsilon^2}\overline{\Sigma}^{-1}\overline{C}'(y-\overline{C}\overline{\xi}) + \frac{1}{\gamma^2}\overline{\Sigma}^{-1}\overline{Q}(\overline{\xi}-\hat\xi)$$

$$\hat\xi(0) = \overline{\xi}_0 \tag{18}$$

$$\dot m = \frac{\gamma^2}{\epsilon^2}|y-\overline{C}\overline{\xi}|^2 - |\overline{\xi}-\hat\xi|^2_{\overline{Q}};\quad m(0)=0 \tag{19}$$

and $\overline{A}$, $\overline{C}$, and $\overline{D}$ were defined through (12) and (13). Note that $\overline{\Sigma}$ satisfies a generalized Riccati differential equation (GRDE). Now, partitioning this matrix as

$$\overline{\Sigma} := \begin{bmatrix}\Sigma_1 & \Sigma_2 \\ \Sigma_2' & \Sigma_3\end{bmatrix} \tag{20}$$

where Σ_1 is an $r\times r$-dimensional symmetric matrix, and substituting this structure into the forward GRDE (17), we arrive at the following coupled differential equations for Σ_1, Σ_2, and Σ_3:

$$\dot\Sigma_1 = -\Sigma_2 A_1 - A_1'\Sigma_2' - \Sigma_2\Sigma_2' - \frac{1}{\gamma^2}Q;\quad \Sigma_1(0) = Q_0 \tag{21a}$$

$$\dot\Sigma_2 = -\Sigma_2 A_2 - A_1'\Sigma_3 - \Sigma_2\Sigma_3;\quad \Sigma_2(0) = \mathbf{0}_{r\times n} \tag{21b}$$

$$\dot\Sigma_3 = -\Sigma_3 A_2 - A_2'\Sigma_3 - \Sigma_3\Sigma_3 + \frac{1}{\epsilon^2}e_1e_1';\quad \Sigma_3(0) = P_0 \tag{21c}$$

where $A_1 := A_{11}y + A_{12}u$. Note that the last differential equation is decoupled from the first two. By standard results on Riccati equations, the matrix $\Sigma_3(t)$ is positive definite for all $t\geq 0$. Let us further introduce the following two matrices:

$$\Phi := \Sigma_3^{-1}\Sigma_2' \tag{22}$$

$$\Xi := \Sigma_1 - \Sigma_2\Sigma_3^{-1}\Sigma_2'. \tag{23}$$

It is straightforward to show that Φ and Ξ satisfy the following matrix differential equations:

$$\dot\Phi = \left(A_2 - \frac{1}{\epsilon^2}\Sigma_3^{-1}e_1e_1'\right)\Phi - A_1;\quad \Phi(0) = \mathbf{0}_{n\times r} \tag{24}$$

$$\dot\Xi = \frac{1}{\epsilon^2}\Phi'e_1e_1'\Phi - \frac{1}{\gamma^2}Q;\quad \Xi(0) = Q_0. \tag{25}$$

The cost-to-come function $W_\gamma(t,\, \xi,\, \hat\xi_{[0,t]},\, y_{[0,t]},\, u_{[0,t]})$ is nonpositive if $\hat\xi$ is chosen to be equal to $\overline{\xi}$ and if, furthermore, the matrix $\overline{\Sigma}(t)$ is nonnegative definite for all $t\geq 0$ and all possible measurement waveforms $y_{[0,\infty)}$. Hence, we have the following result.

Theorem 4.1: For the worst-case identification problem described by (3a)–(7), we have the following for each fixed $\epsilon > 0$.

1) The optimal performance level γ^* is given by

$$\begin{aligned}\gamma^* = \inf\{\gamma>0:\ & \text{the solution to (25), } \Xi_\gamma(t), \\ & \text{is nonnegative definite for all } t\geq 0 \text{ and for} \\ & \text{all possible measurement waveforms } y_{[0,\infty)}\}. \end{aligned} \tag{26}$$

2) For each $\gamma>\gamma^*$, the matrix $\Xi_\gamma(t)$ is positive definite for all $t\geq 0$ and all possible measurement waveforms $y_{[0,\infty)}$, and an identifier that achieves the performance level γ is given by

$$\delta_\gamma(t,\, y_{[0,t]},\, u_{[0,t]}) = [I_{r\times r}\quad \mathbf{0}_{r\times n}]\hat\xi \tag{27}$$

$$\dot{\hat\xi} = \overline{A}\hat\xi + \frac{1}{\epsilon^2}\begin{bmatrix}\Xi_\gamma^{-1} & -\Xi_\gamma^{-1}\Phi' \\ -\Phi\Xi_\gamma^{-1} & \Sigma_3^{-1}+\Phi\Xi_\gamma^{-1}\Phi'\end{bmatrix}\overline{C}'(y-\overline{C}\hat\xi)$$

$$\hat\xi(0) = \overline{\xi}_0. \tag{28}$$

Furthermore, if in addition the following persistency of excitation condition:

$$\lim_{t\to\infty}\lambda_{\min}\left(\int_0^t \Phi'e_1e_1'\Phi\, ds\right) = \infty \tag{29}$$

holds (where $\lambda_{\min}(X)$ is the minimum eigenvalue of the symmetric matrix X) and the disturbance quadruple $(x_0,\, \theta,\, w_{[0,\infty)},\, v_{[0,\infty)})$ belongs to $\mathcal{L}^2$, then the parameter estimates converge to the true value, i.e., $\lim_{t\to\infty}\delta_\gamma(t;\, y_{[0,t]},\, u_{[0,t]}) = \theta$.

Proof: We first note that the inequality

$$\inf_{\hat{\xi}_{[0,\infty)}} \sup_{\xi_0, w_{[0,\infty)}, v_{[0,\infty)}} J_\gamma(\hat{\xi}; \xi_0, w_{[0,\infty)}, v_{[0,\infty)}) \le 0$$

is equivalent to the one

$$\inf_{\hat{\xi}_{[0,t]}} \sup_{\xi_0, w_{[0,t]}, v_{[0,t]}} J_\gamma^t(\hat{\xi}; \xi_0, w_{[0,t]}, v_{[0,t]}) \le 0 \quad \text{for all} \quad t \ge 0$$

where

$$J_\gamma^t(\hat{\xi}; \xi_0, w_{[0,t]}, v_{[0,t]}) := \int_0^t \{|\xi - \hat{\xi}|_Q^2 - \gamma^2[|w(t)|^2 + |v(t)|^2]\}\, dt - \gamma^2 |\xi_0 - \overline{\xi}_0|_{Q_0}^2.$$

For each fixed $t \ge 0$, the second inf–sup problem can be rewritten as follows:

$$\begin{aligned}\inf_{\hat{\xi}_{[0,t]}} \sup_{\xi_0, w_{[0,t]}, v_{[0,t]}} J_\gamma^t &= \inf_{\hat{\xi}_{[0,t]}} \sup_{y_{[0,t]}, \xi} \sup_{\xi_0, w_{[0,t]}, v_{[0,t]} | y_{[0,t]}, u_{[0,t]}, \xi(t) = \xi} J_\gamma^t \\ &= \inf_{\hat{\xi}_{[0,t]}} \sup_{y_{[0,t]}, \xi} W_\gamma(t, \xi, \hat{\xi}_{[0,t]}, w_{[0,t]}, v_{[0,t]}).\end{aligned}$$

The second equality holds since the estimate $\hat{\xi}$ depends only on the measurement waveform $y_{[0,t]}$.

Now, for a fixed $\gamma > 0$, it is necessary to have $\overline{\Sigma}$ nonnegative definite for the existence of an identifier that achieves the performance level γ, since if $\overline{\Sigma}$ has any negative eigenvalues then W_γ can be made arbitrarily large positive by an appropriate choice of the uncertainty. Because of the structure (20) for $\overline{\Sigma}$ and the positive-definiteness of its subblock matrix Σ_3, it is further necessary to have Ξ_γ be nonnegative definite.

Now, fix a γ that is strictly larger that the right-hand side (RHS) of (26). Note that

$$\Xi_\gamma(t) = \Xi_{\gamma_1}(t) + \left(\frac{1}{\gamma_1^2} - \frac{1}{\gamma^2}\right) \int_0^t Q\, ds$$

for any γ_1 less than γ and larger than γ^*. This, coupled with the fact that $\Xi_\gamma(0) > 0$, implies that $\Xi_\gamma(t) > 0$. This further implies that $\overline{\Sigma}(t; \gamma) > 0$. The identifiers (27) and (28) are well defined for this value of γ. Substituting this identifier into the system, i.e., picking $\hat{\xi} = \overline{\xi}$, leads to the following inequality:

$$\sup_{y_{[0,t]}, \xi} W_\gamma(t, \xi, \hat{\xi}_{[0,t]}, w_{[0,t]}, v_{[0,t]}) \le 0$$

for any $t \ge 0$. Hence, the identifier achieves the performance level γ.

So far, we have shown statement (1) and that the identifier (27) and (28) achieves the desired performance level $\gamma > \gamma^*$. To prove the convergence of the parameter estimates, we note the following equality:

$$\Xi_\gamma(t) = \frac{\gamma_1^2}{\gamma^2} \Xi_{\gamma_1}(t) + \left(1 - \frac{\gamma_1^2}{\gamma^2}\right) \int_0^t \Phi' e_1 e_1' \Phi\, ds$$

for any $\gamma_1 \in (\gamma^*, \gamma)$, $t \ge 0$ and any measurement waveform $y_{[0,\infty)}$. This implies that under the persistency of excitation condition (29)

$$\lim_{t \to \infty} \lambda_{\min}[\Xi_\gamma(t)] = \infty.$$

We now note the following identity from (15) and (16):

$$\begin{aligned}|\xi_0 - \overline{\xi}_0|_{Q_0}^2 &= |\xi(t) - \hat{\xi}(t)|_{\overline{\Sigma}(t;\gamma)}^2 \\ &\quad + \int_0^t \left[\frac{1}{\gamma^2} |\xi - \hat{\xi}|_Q^2 + |w - \overline{D}'\overline{\Sigma}(\xi - \hat{\xi})|^2 \right. \\ &\quad \left. - |w|^2 - |v|^2 + \left| v + \frac{1}{\epsilon} \overline{C}(\xi - \hat{\xi}) \right|^2 \right] ds.\end{aligned}$$

When the quadruple $(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)})$ belongs to $\mathcal{L}^2$, the above equality implies that $|\xi(t) - \hat{\xi}(t)|_{\overline{\Sigma}(t)}^2$ is uniformly bounded for $t \in [0, \infty)$ which further implies that $|\theta - \delta_\gamma(t)|_{\Xi_\gamma(t)}^2$ is uniformly bounded for $t \in [0, \infty)$. This, coupled with the unboundedness of $\Xi_\gamma(t)$, leads to the convergence of the parameter estimates $\delta_\gamma(t)$. □

Remark 4.1: Although Theorem 4.1 provides an implicit characterization for γ^*, this optimal performance level is difficult to evaluate for a general cost matrix Q. To check whether a given $\gamma > 0$ is larger than γ^* or not, one needs to show that the matrix function $\Xi_\gamma(t; y_{[0,t]}, u_{[0,t]})$ is positive definite for all possible measurement waveforms $y_{[0,t]}$ and $t \ge 0$. This validation procedure is almost impossible to carry out in general. There are some special choices of Q, however, for which γ^* can be determined explicitly. One such choice is

$$Q(t, y_{[0,t]}, u_{[0,t]}) = \frac{1}{\epsilon^2} \Phi' e_1 e_1' \Phi$$

in which case $\gamma^* = 1$. Hence, this stands out as a *natural choice* for the error weighting matrix.

Another choice for the cost matrix Q of special interest is

$$Q(t, y_{[0,t]}, u_{[0,t]}) = \Xi Q_P \Xi$$

where Q_P is a positive definite constant matrix. This choice of Q weights the identification error heavily when the confidence level in the parameter estimates (Ξ) is high, but weights it lightly when the confidence level is low. For this choice, the matrix Ξ is uniformly bounded from above when the underlying system is stable which means that the identification gain does not converge to zero. Thus, whenever there is sufficient excitation in the system, we are guaranteed an exponential rate of convergence for the parameter estimates without covariance resetting. To compute the optimal performance level that is associated with this choice of the cost matrix, we observe that the worst-case covariance matrix Ξ^{-1} satisfies the following dynamics:

$$\frac{d}{dt}(\Xi^{-1}) = -\Xi^{-1} \frac{1}{\epsilon^2} \Phi' e_1 e_1' \Phi \Xi^{-1} + \frac{1}{\gamma^2} Q_P.$$

Hence, the matrix Ξ^{-1} increases at most linearly in time when the system does not have excitation. Consequently, the matrix Ξ exists on $[0, \infty)$ for any value of $\gamma > 0$. The optimal performance level for this choice of Q is therefore zero.

The two choices above for the cost matrix Q lead to precisely computable optimal performance levels for the worst-case identification problem. Obviously, any linear combination of these two choices will also lead to a well-defined and computable optimal performance level. Through simulations, we have observed that indeed such a combination of the two choices above leads to an identifier that performs the best. ⋄

Simplification of the Identifier (27) and (28): The identifier structure (27) and (28) can be simplified considerably if the matrix P_0 is chosen in such a way that makes Σ_3 time invariant. We note that the matrix $A_2 - (1/\epsilon^2)\Sigma_3^{-1}e_1e_1'$ is of the form A_f introduced by (9) which is the system matrix for the Kreisselmeier observer. In the time-invariant case, the evaluation of the matrix function $\Psi(t)$ can be reduced to the evaluation of a $2n$-dimensional vector differential equation. By straightforward algebraic manipulations, we obtain the following simplified form of the worst-case identifier (27) and (28):

$$A_2\Pi + \Pi A_2' + I - \frac{1}{\epsilon^2}\,\Pi e_1 e_1 \Pi = 0 \tag{30}$$

$$A_{fo} := A_2 - \frac{1}{\epsilon^2}\,\Pi e_1 e_1' \tag{31}$$

$$\dot{\eta} = A_{fo}\eta - e_1 y; \quad \eta(0) = \mathbf{0}_{n\times 1} \tag{32}$$

$$\dot{\lambda} = A_{fo}\lambda - e_1 u; \quad \lambda(0) = \mathbf{0}_{n\times 1} \tag{33}$$

$$\begin{aligned}\Phi = {}&[\eta \quad A_{fo}^{-1}\eta \quad \cdots \quad A_{fo}^{-n+1}\eta]A_{11}\\ &+ [\lambda \quad A_{fo}^{-1}\lambda \quad \cdots \quad A_{fo}^{-n+1}\lambda]A_{12}\end{aligned} \tag{34}$$

$$\begin{aligned}\frac{d}{dt}(\Xi^{-1}) &= -\Xi^{-1}\left(\frac{1}{\epsilon^2}\,\Phi' e_1 e_1'\Phi - \frac{1}{\gamma^2}\,Q\right)\Xi^{-1}\\ \Xi^{-1}(0) &= Q_0^{-1}\end{aligned} \tag{35}$$

$$\delta(t,\, y_{[0,t]},\, u_{[0,t]}) = \hat{\theta}(t) \tag{36}$$

$$\dot{\hat{\theta}} = -\frac{1}{\epsilon^2}\,\Xi^{-1}\Phi' e_1(y - e_1'\hat{x}); \quad \hat{\theta}(0) = \overline{\theta}_0 \tag{37}$$

$$\begin{aligned}\dot{\hat{x}} = {}& A_2\hat{x} + (yA_{11} + uA_{12})\hat{\theta}\\ &+ \frac{1}{\epsilon^2}(\Pi + \Phi\Xi^{-1}\Phi')e_1(y - e_1'\hat{x}); \quad \hat{x}(0) = \overline{x}_0.\end{aligned} \tag{38}$$

The observation made above is put into precise terms in the following corollary.

Corollary 4.1: Consider the worst-case identification problem described by (3a)–(7), and let the initial condition P_0 be chosen as Π^{-1}, where Π is the solution to the ARE (30). Then:

1) The optimal performance level γ^* is given by

$$\begin{aligned}\gamma^* = \inf\{\gamma > 0:\ &\text{the solution to (35), } \Xi_\gamma^{-1}(t),\\ &\text{is positive definite for all } t \geq 0 \text{ and all}\\ &\text{possible measurement waveforms } y_{[0,\infty)}\}.\end{aligned} \tag{39}$$

2) For each $\gamma > \gamma^*$, the identifier defined by (30)–(38) achieves the desired performance level γ. Furthermore, if the persistency of excitation condition (29) is satisfied and the uncertainty quadruple $(x_0,\, \theta,\, w_{[0,\infty)},\, v_{[0,\infty)})$ belongs to $\mathcal{L}^2$, then the parameter estimates converge to the true value, i.e., $\lim_{t\to\infty}\hat{\theta}(t) = \theta$. ⋄

A comparison of the identifier (30)–(38) with that based on the Kreisselmeier observer shows that the minimax identifier (30)–(38) admits a Kreisselmeier observer structure. The parameter vector f for the Kreisselmeier observer is fixed at an "optimal" value $(1/\epsilon^2)\Pi e_1$ which is further independent of the desired performance level $\gamma > 0$. Hence, as a byproduct of this study on the worst-case identification problem, we obtain a set of "optimal" parameter values for the Kreisselmeier observer. On the other hand, the identifier (30)–(38) is substantially different from Kreisselmeier observer-based identifiers in the sense that it contains additional filter dynamics for the state variable x of the unknown plant which is coupled with the dynamics of $\hat{\theta}$ in a nontrivial way. A detailed comparison of the performances for these two identification schemes on an illustrative example is included in Section VI. Note that the filters (32) [respectively, (33)] for y (respectively, u), serve the same purpose as (10) [respectively, (11)]. Here, we have designed them to filter $e_1 y$ (respectively, $e_1 u$) instead of $e_n y$ (respectively, $e_n u$) to avoid the multiplication of the matrix A_{fo} in the reconstruction of the matrix Φ. Another remark is that when the parameter $\epsilon > 0$ is small, which corresponds to the case of weak measurement noise, the matrix A_{fo} includes terms of $O(1/\epsilon)$ whose implication is discussed next.

Reduced-Order Identifier for $0 < \epsilon \ll 1$: Consider the special case when the intensity of the measurement noise is very low which is formally captured by letting $\epsilon \ll 1$. As ϵ decreases to zero, the computation of the identifier (30)–(38), as well as the on-line computation of the parameter estimates, becomes numerically ill-conditioned. To alleviate this numerical stiffness problem, we will pursue a singular perturbation analysis to obtain an "approximate" identifier for (30)–(38) for sufficiently small values of ϵ.

For ease of reference to the reader, we introduce below the complete set of dynamics for the approximate (reduced-order) identifier, followed by the arguments that lead to it. First, we define a $1 \times (n-1)$-dimensional matrix A_{21} and an $(n-1) \times (n-1)$-dimensional matrix A_{22} by the following partitioning of the matrix A_2:

$$A_2 =: \begin{bmatrix} 0 & A_{21} \\ \mathbf{0}_{(n-1)\times 1} & A_{22} \end{bmatrix}.$$

Using these two matrices, the approximate identifier can be written in the following form:

$$A_{22}\Pi_s + \Pi_s A_{22}' - \Pi_s A_{21}' A_{21}\Pi_s + I_{n-1} = 0 \tag{40}$$

$$A_{fs} := A_{22} - \Pi_s A_{21}' A_{21} \tag{41}$$

$$\dot{\eta}_s = A_{fs}\eta_s + \Pi_s A_{21}' y; \quad \eta_s(0) = 0 \tag{42}$$

$$\dot{\lambda}_s = A_{fs}\lambda_s + \Pi_s A_{21}' u; \quad \lambda_s(0) = 0 \tag{43}$$

$$\begin{aligned}\Phi_s = {}&[\eta_s \quad A_{fs}^{-1}\eta_s \quad \cdots \quad A_{fs}^{-n+1}\eta_s]A_{11}\\ &+ [\lambda_s \quad A_{fs}^{-1}\lambda_s \quad \cdots \quad A_{fs}^{-n+1}\lambda_s]A_{12}\end{aligned} \tag{44}$$

$$\begin{aligned}\frac{d}{dt}(\Xi_s^{-1}) = -\Xi_s^{-1}\Big[&(\Phi_s' A_{21}' - A_1' e_1)\\ &\cdot (A_{21}\Phi_s - e_1' A_1) - \frac{1}{\gamma^2}\,Q\Big]\,\Xi_s^{-1}\end{aligned} \tag{45}$$

$$\delta_s(t, y_{[0,t]}, u_{[0,t]}) = \hat{\theta}_s \tag{46}$$

$$\dot{\hat{\theta}}_s = -\Xi_s^{-1}(\Phi_s' A_{21}' - A_1' e_1)\frac{1}{\epsilon}(y - e_1'\hat{x}_s); \quad \hat{\theta}_s(0) = \overline{\theta}_0 \tag{47}$$

$$\dot{\hat{x}}_s = A_1\hat{\theta}_s + A_2\hat{x}_s + \left[\frac{1}{\Pi_s A_{21}' + \Phi_s\Xi_s^{-1}(\Phi_s' A_{21}' - A_1' e_1)}\right] \cdot \frac{1}{\epsilon}(y - e_1'\hat{x}_s); \quad \hat{x}_s(0) = \overline{x}_0. \tag{48}$$

The ARE (30) is of the standard singularly perturbed form with respect to the parameter ϵ. Thus, its solution Π can be shown to admit the following approximation [11]:

$$\Pi = \begin{bmatrix} \epsilon + O(\epsilon^2) & \epsilon A_{21}\Pi_s + O(\epsilon^2) \\ \epsilon\Pi_s A_{21}' + O(\epsilon^2) & \Pi_s + O(\epsilon) \end{bmatrix}$$

where Π_s is the solution to ARE (40). Given the above equality, the matrix A_{fo} can be expressed as

$$A_{fo} = \begin{bmatrix} \frac{1}{\epsilon} + O(1) & A_{21} \\ -\frac{1}{\epsilon}\Pi_s A_{21}' + O(1) & A_{22} \end{bmatrix}.$$

This specific structure allows a time-scale decomposition of the filters (32) and (33). For (32), first introduce the state transformation $\eta = (\epsilon\eta_f, \eta_s')'$; then the filter is described in the standard singularly perturbed form with the state variables $(\eta_f, \eta_s')'$. The slow dynamics for (32) [respectively, (33)] are exactly (42) [respectively, (43)]. Some further straightforward algebraic manipulations on the identifier (30)–(38) lead to the approximate identifier (40)–(48), after neglecting fast filter variables and higher-order terms of ϵ.

The optimality of the approximate identifier (40)–(48) is established in the following theorem.

Theorem 4.2: Consider the worst-case identification problem described by (3a)–(7). Let the initial condition P_0 be chosen as Π^{-1}, where Π is the unique positive definite solution to the ARE (30). The reduced-order identifier δ_s given by (40)–(48) achieves the desired performance level $\gamma > 0$, if the following GRDE admits a nonnegative definite solution $\hat{\Sigma}(t; \gamma, y_{[0,t]}, u_{[0,t]})$ on $t \geq 0$ for any measurement waveform $y_{[0,t]}$

$$\dot{\hat{\Sigma}} = -\hat{\Sigma}\overline{A} - \overline{A}'\hat{\Sigma} - \frac{1}{\gamma^2}\overline{Q} - \hat{\Sigma}(\overline{D}\,\overline{D}' + \overline{H}\,\overline{H}')\hat{\Sigma} - \frac{1}{\epsilon}\hat{\Sigma}\overline{H}\,\overline{C} - \frac{1}{\epsilon}\overline{C}'\overline{H}'\hat{\Sigma}; \quad \hat{\Sigma}(0) = \overline{Q}_0 \tag{49}$$

where

$$\overline{H} := \begin{bmatrix} \Xi_s^{-1}(\Phi_s' A_{21}' - A_1' e_1) \\ -1 \\ -\Pi_s A_{21}' + \Phi_s\Xi_s^{-1}(A_1' e_1 - \Phi_s' A_{21}') \end{bmatrix}. \tag{50}$$

Proof: Introduce $\hat{\xi}_s := (\hat{\theta}_s', \hat{x}_s')'$ and $\tilde{\xi}_s = \xi - \hat{\xi}_s$. It is then straightforward to obtain the following dynamic equation for $\tilde{\xi}_s$:

$$\dot{\tilde{\xi}}_s = \overline{A}\tilde{\xi}_s + \overline{D}w + \overline{H}\frac{1}{\epsilon}(y - \overline{C}\hat{\xi}_s).$$

Under the reduced-order identifier (40)–(48), the cost function (7) can be rewritten as

$$\begin{aligned} J_\gamma^s(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}) &:= J_\gamma(\delta_s, x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}) \\ &= \int_0^\infty [|\tilde{\xi}|_Q^2 - \gamma^2(|w|^2 + |v|^2)]\, dt - \gamma^2|\tilde{\xi}(0)|_{Q_0}^2. \end{aligned}$$

By an argument similar to that used in the proof of Theorem 4.1, proving the validity of the bound

$$\sup_{x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}} J_\gamma^s(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}) \leq 0$$

is equivalent to showing that

$$W^s(t, \tilde{\xi}, y_{[0,t]}, u_{[0,t]}) := \sup_{x_0, \theta, w_{[0,t]} | y_{[0,t]}, u_{[0,t]}, \tilde{\xi}(t) = \tilde{\xi}} \cdot \int_0^t \left[|\tilde{\xi}|_Q^2 - \gamma^2\left(|w|^2 + \frac{1}{\epsilon^2}|y - \overline{C}\xi + \overline{C}\tilde{\xi}|^2\right)\right] ds - \gamma^2|\tilde{\xi}(0)|_{Q_0}^2 \leq 0$$

for all $t \geq 0$ and any measurement waveform $y_{[0,t]}$.

By a "completion of squares" argument, the function $W^s(t, \tilde{\xi}, y_{[0,t]}, u_{[0,t]})$ can be shown to be less than or equal to $-\gamma^2|\tilde{\xi}(t)|_{\hat{\Sigma}(t)}^2$. This completes the proof of the theorem. □

Checking the nonnegative definiteness of $\hat{\Sigma}$ in advance for all possible measurement waveforms may not be possible. We observe, however, the following useful relationship between $\overline{\Sigma}^{-1}\overline{C}'$ and $\overline{H}$, where $\overline{\Sigma}$ is the solution of the GRDE (17) with $P_0 = \Pi$:

$$\overline{\Sigma}^{-1}\overline{C}' = \epsilon\overline{H} + O(\epsilon^2).$$

Hence, roughly speaking, the matrix function $\hat{\Sigma}$ approximates $\overline{\Sigma}$ for sufficiently small values of ϵ. This, in fact, holds for a wide class of measurement waveforms, as to be determined shortly.

Motivated by the results of [5], we introduce the following set of measurement waveforms and corresponding disturbance set to formally characterize the closeness of performances for the full-order identifier (30)–(38) and the approximate (reduced-order) identifier (40)–(48).

$\mathcal{LBP}_{EPW}^A(L, M, \overline{\rho}, T_m, K)$: This is the set of all waveforms $y_{[0,\infty)}$ that satisfy the following four conditions for some positive constants L, M, $\overline{\rho}$, T_m, and integer K.

1) There exist $K < \infty$ time instances $0 < t_1 < \cdots < t_K < \infty$, such that

$$\min_{k \in \{1, \cdots, K\}} t_k - t_{k-1} \geq T_m.$$

2) $\|A_{11}[y(t') - y(t)] + A_{12}[u(t') - u(t)]\|_2 \leq L|t' - t|$, $\forall t', t \in [0, t_1]$, or $\forall t', t \in [t_k, t_{k+1}]$, $k = 1, \cdots, K-1$, or $\forall t', t \in [t_K, \infty)$.
3) $\|A_1\|_2 \leq M$ for all $t \geq 0$.
4) $\Xi(t) \geq (t+1)\overline{\rho}I$, for all $t \geq 0$.

$\mathcal{W}[\mathcal{LBP}^A_{EPW}(L, M, \overline{\rho}, T_m, K)]$: This is the set of all disturbance quadruples $(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)})$ that lead to a measurement waveform $y_{[0,\infty)}$ that belongs to the set $\mathcal{LBP}^A_{EPW}(L, M, \overline{\rho}, T_m, K)$.

In terms of this notation, we can now state and prove the following result.

Corollary 4.2: Consider the SISO system (3a) and (3b) with the game cost function (7) indexed by γ. Let the initial cost matrix P_0 be chosen as Π^{-1}, where Π is the positive definite solution to the ARE (30). For any $\gamma > \gamma^*$, where γ^* is defined by (39) and any disturbance within the set $\mathcal{W}[\mathcal{LBP}^A_{EPW}(L, M, \overline{\rho}, T_m, K)]$, the reduced-order identifier δ_s given by (40)–(48) achieves the disturbance attenuation level γ for sufficiently small $\epsilon > 0$, i.e.,

$$\sup_{(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}) \in \mathcal{W}[\mathcal{LBP}^A_{EPW}(L, M, \overline{\rho}, T_m, K)]} \cdot J_\gamma(\delta_s; x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}) = 0. \tag{51}$$

Furthermore, if $(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}) \in \mathcal{W}[\mathcal{LBP}^A_{EPW}(L, M, \overline{\rho}, T_m, K)] \bigcap \mathcal{L}^2$, then the parameter estimates $\hat{\theta}_s(t)$ converge to the true value θ as $t \to \infty$ for sufficiently small values of ϵ.

Proof: Since there are at most K discontinuities in the measurement waveform $y_{[0,t]}$, it is sufficient to prove the theorem for the case when $K = 0$.

To study the solution to GRDE (49) more closely, partition the matrix $\hat{\Sigma}$ as

$$\hat{\Sigma} = \begin{bmatrix} \hat{\Sigma}_1 & \hat{\Sigma}_2 \\ \hat{\Sigma}_2' & \hat{\Sigma}_3 \end{bmatrix}$$

where $\hat{\Sigma}_1$ is an $r \times r$-dimensional symmetric matrix. Introduce further the notation

$$\hat{\Sigma}_3 =: \begin{bmatrix} \frac{1}{\epsilon}\hat{\Sigma}_{31} & \hat{\Sigma}_{32} \\ \hat{\Sigma}_{32}' & \hat{\Sigma}_{33} \end{bmatrix} \quad \hat{\Sigma}_3^{-1}\hat{\Sigma}_2' =: \begin{bmatrix} \epsilon\hat{\Sigma}_{21} \\ \hat{\Sigma}_{22} \end{bmatrix}$$
$$\hat{\Xi} := \hat{\Sigma}_1 - \hat{\Sigma}_2\hat{\Sigma}_3^{-1}\hat{\Sigma}_2'$$

where $\hat{\Sigma}_{31}$ is a scalar and $\hat{\Sigma}_{21}$ is a $1 \times r$-dimensional matrix.

Using a "Lyapunov function" approach similar to that used in the proof of [5, Th. 5.5], one can establish validity of the following approximations:

$$\begin{aligned} \hat{\Sigma}_{31} &= 1 + O(\epsilon^2); \quad \hat{\Sigma}_{32} = -A_{21} + O(\epsilon) \\ \hat{\Sigma}_{33} &= \Pi_s^{-1} + O(\epsilon); \quad \hat{\Sigma}_{21} = A_{21}\Phi_s - e_1'A_1 + O(\epsilon) \\ \hat{\Sigma}_{22} &= \Phi_s + O(\epsilon); \quad \hat{\Xi}(t) = \Xi(t) + (t+1)O(\epsilon) \end{aligned}$$

on the entire time interval $[0, \infty)$.

Then, the existence and nonnegative-definiteness of the matrix $\hat{\Sigma}$ follow on $[0, \infty)$ for sufficiently small ϵ which further implies that the reduced-order identifier achieves the disturbance attenuation level γ on the set of disturbances $\mathcal{W}[\mathcal{LBP}^A_{EPW}(L, M, \overline{\rho}, T_m, K)]$. Since $\lim_{t\to\infty} \lambda_{\min}[\hat{\Xi}(t)] = \infty$ and

$$|\tilde{\xi}(0)|^2_{Q_0} = |\tilde{\xi}(t)|^2_{\hat{\Sigma}(t)} + \int_0^t \left[\frac{1}{\gamma^2}|\tilde{\xi}|^2_{\overline{Q}} + |w - \overline{D}\hat{\Sigma}\tilde{\xi}|^2 + \left|v - \frac{1}{\epsilon\overline{C}\tilde{\xi}} - \overline{H}'\hat{\Sigma}\tilde{\xi}\right|^2 - |w|^2 - |v|^2\right] ds$$

the parameter estimates $\hat{\theta}_s(t)$ must converge to θ as $t \to \infty$.

This completes the proof of the corollary. □

Remark 4.2: Although the reduced-order identifier (40)–(48) is only two-integrator simpler in structure than the identifier (30)–(38), it avoids an interconnection between a fast dynamics and a nonlinear dynamics. Thus, the reduced-order identifier is expected to be numerically well conditioned for small values of ϵ. This observation is corroborated by the example presented in Section VI, where the total simulation time for a system using the reduced-order identifier (40)–(48) was found to be at least one-third shorter than that needed for the identifier (30)–(38).

The reduced-order worst-case identifier requires $[(n+m+1)(n+m)/2] + 4n + m - 2$ integrators in its structure, while the conventional identifier requires $[(n+m+1)(n+m)/2] + 3n + m$ integrators, with the difference being $n - 2$; hence in this sense the former is inferior to the latter. On the other hand, the reduced-order identifier exhibits much better performance against nonzero initial conditions, a wide range of unknown parameter values and nonstochastic disturbance inputs, as to be demonstrated in the context of a simulation example in Section VI. In addition, the reduced-order worst-case identifier ensures satisfaction of a disturbance attenuation bound. There is clearly a tradeoff between robust performance and computational complexity, and since the difference in complexity between the two identifiers is only linear in n, the reduced-order worst-case identifier stands out as the winner in this comparison. ◇

V. Extension to the MIMO Case

We now turn to study the worst-case identification problem for MIMO linear systems. The plant is known to have p inputs, m outputs, and a set of fixed observability indexes $\{n_1, \cdots, n_m\}$. Hence, it is $n := \sum_{i=1}^m n_i$ dimensional. We assume that the plant is controllable and observable and that the outputs are linearly independent, i.e., there is no m-dimensional nonzero vector k such that $k'y$ is identically zero for any $\mathcal{L}^2$ control input signal. By the results of [3], this class of plants can be described by the following state space model:

$$\dot{x} = A_2x + Fy + Bu \tag{52}$$
$$y = Cx + Gy \tag{53}$$

where

$$\begin{aligned} A_2 &= \text{block diagonal } \{A_2^{(1)}, \cdots, A_2^{(m)}\} \\ A_2^{(i)} &= \begin{bmatrix} \mathbf{0}_{(n_i-1)\times 1} & I_{n_i-1} \\ 0 & \mathbf{0}_{1\times(n_i-1)} \end{bmatrix} \quad i = 1, \cdots, m \\ F &= [f^{(1)} \ \cdots \ f^{(m)}]; \quad B = [b^{(1)} \ \cdots \ b^{(p)}] \\ C &= \text{block diagonal } \{C^{(1)}, \cdots, C^{(m)}\} \\ C^{(i)} &= [1 \quad \mathbf{0}_{(n_i-1)\times 1}]; \quad G = [g^{(1)} \ \cdots \ g^{(m)}] \end{aligned}$$

and the matrix G is further strictly lower triangular.

Since our interest is in obtaining a parameter identifier that is robust with respect to exogenous noise and disturbances which do not necessarily admit stochastic descriptions, we let the plant dynamics be perturbed by additive noise w and the measurement y be "contaminated" by measurement noise Ev, where the matrix E is taken to be

$$E := \text{diagonal } \{\epsilon_1, \cdots, \epsilon_m\}$$

and ϵ_i, $i = 1, \cdots, m$ are positive scalars. Hence, we consider the following general model:

$$\dot{x} = A_2 x + \left[\sum_{i=1}^{m} y_i A_{11}^{(i)} + \sum_{i=1}^{p} u_i A_{12}\right] \theta_1 + w; \quad x(0) = x_0 \tag{54}$$

$$y = Cx + \sum_{i=1}^{m} y_i G^{(i)} \theta_2 + Ev \tag{55}$$

where θ_1 is an r_1-dimensional unknown constant parameter; θ_2 is an $r_2 := r - r_1$-dimensional unknown constant parameter; $A_{11}^{(i)}$, $i = 1, \cdots, m$ and $A_{12}^{(i)}$, $i = 1, \cdots, p$ are known constant matrices of dimensions $n \times r_1$; A_2, C, and E are as defined above; w is the n-dimensional system noise; and v is an m-dimensional measurement noise. The matrices $G^{(i)}$, $i = 1, \cdots, m$ further satisfy the following property due to the strict lower triangular structure of G:

$$G^{(i)} := \begin{bmatrix} G_{(1)}^{(i)} \\ \vdots \\ G_{(m)}^{(i)} \end{bmatrix} \quad i = 1, \cdots, m$$

$$G_{(j)}^{(i)} = \mathbf{0}_{1 \times r_2}; \quad i \geq j, \quad i = 1, \cdots, m, \quad j = 1, \cdots, m$$

where the subblock matrices $G_{(j)}^{(i)}$ are of dimensions $1 \times r_2$, $i = 1, \cdots, m$, $j = 1, \cdots, m$.

Denote the unknown parameters $(\theta_1', \theta_2')'$ by θ. The objective is again to determine an estimate $\hat{\theta}$ for θ using a policy (4) that minimizes the performance index (5). The optimum performance level is again denoted by γ^* as in (6).

As in the SISO case, we transform the problem to one where the general framework of affine quadratic minimax controller design is applicable. This is done by associating the natural dynamics $\dot{\theta} = 0$ with the system (54) and (55) and by following an analysis similar to that in the SISO case. This leads to the counterpart of Theorem 4.1 which is stated below as Theorem 5.1, after introducing the expressions for the full-order identifier

$$\begin{aligned} \dot{\Sigma}_3 &= -A_2'\Sigma_3 - \Sigma_3 A_2 - \Sigma_3 \Sigma_3 + C'(EE')^{-1}C \\ \Sigma_3(0) &= P_0 \end{aligned} \tag{56}$$

$$A_{fo} := A_2 - \Sigma_3^{-1} C'(EE')^{-1} C \tag{57}$$

$$\begin{aligned} \dot{\Phi} &= A_{fo}\Phi + \sum_{i=1}^{m} [-A_{11}^{(i)} \quad \Sigma_3^{-1}C'(EE')^{-1}G^{(i)}] y_i \\ &\quad + \sum_{i=1}^{p} [-A_{12}^{(i)} \quad \mathbf{0}_{n\times r_2}] u_i; \quad \Phi(0) = \mathbf{0}_{n \times r} \end{aligned} \tag{58}$$

$$\begin{aligned} \dot{\Xi} &= (\Phi'C' - \overline{G}')(EE')^{-1}(\Phi C - \overline{G}) - \frac{1}{\gamma^2} Q \\ \Xi(0) &= Q_0 \end{aligned} \tag{59}$$

$$\overline{G} := \left[\mathbf{0}_{m \times r_1} \quad \sum_{i=1}^{m} y_i G^{(i)}\right] \tag{60}$$

$$\delta(t, y_{[0,t]}, u_{[0,t]}) = \hat{\theta}(t) \tag{61}$$

$$\begin{aligned} \dot{\hat{\theta}} &= \Xi^{-1}(\overline{G}' - \Phi'C')(EE')^{-1}(y - \overline{G}\hat{\theta} - C\hat{x}) \\ \hat{\theta}(0) &= \overline{\theta}_0 \end{aligned} \tag{62}$$

$$\begin{aligned} \dot{\hat{x}} &= A_2\hat{x} + \overline{A}_1\hat{\theta} + [\Phi\Xi^{-1}(\Phi'C' - \overline{G}') + \Sigma_3^{-1}C'] \\ &\quad \cdot (EE')^{-1}(y - \overline{G}\hat{\theta} - C\hat{x}); \quad \hat{x}(0) = \overline{x}_0 \end{aligned} \tag{63}$$

$$\overline{A}_1 := \left[\sum_{i=1}^{m} y_i A_{11}^{(i)} + \sum_{i=1}^{p} u_i A_{12}^{(i)} \quad \mathbf{0}_{n \times r_2}\right]. \tag{64}$$

Theorem 5.1: Consider the worst-case identification problem for the MIMO linear system (54) and (55) with the cost function (5), and let the optimal performance level γ^* be defined as in (6). Then, we have the following.

1) The optimal performance level γ^* is given by

$$\begin{aligned} \gamma^* = \inf \{\gamma > 0:\ &\text{the solution to (59), } \Xi_\gamma(t), \\ &\text{is nonnegative definite for all } t \geq 0 \text{ and all} \\ &\text{possible measurement waveforms } y_{[0,\infty)}\}. \end{aligned} \tag{65}$$

2) For each $\gamma > \gamma^*$, the matrix $\Xi_\gamma(t)$ is positive definite for all $t \geq 0$ and all possible measurement waveforms $y_{[0,\infty)}$, and the identifier (56)–(64) achieves the desired performance level γ. Furthermore, if the following persistency of excitation condition:

$$\lim_{t\to\infty} \lambda_{\min}\left[\int_0^t (\Phi'C' - \overline{G}')(EE')^{-1}(C\Phi - \overline{G})\, ds\right] = \infty \tag{66}$$

holds, and the disturbance quadruple $(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)})$ belongs to $\mathcal{L}^2$, then the parameter estimates converge to the true parameter value, i.e., $\lim_{t\to\infty} \hat{\theta}(t) = \theta$.

Proof: The theorem can be proved by following steps similar to those used in the proof of Theorem 4.1. □

Simplification of the Identifier (56)–(64): The identifier structure (56)–(64) can be simplified by a proper choice of the initial weighting matrix P_0, to result in a time-invariant Σ_3, and by further utilizing the block diagonal structure of the system matrices. Toward this end, we introduce the following notation:

$$A_{11}^{(i)} =: \begin{bmatrix} A_{11(1)}^{(i)} \\ \vdots \\ A_{11(m)}^{(i)} \end{bmatrix} \quad i = 1, \cdots, m$$

$$A_{12}^{(i)} =: \begin{bmatrix} A_{12(1)}^{(i)} \\ \vdots \\ A_{12(m)}^{(i)} \end{bmatrix} \quad i = 1, \cdots, m$$

where the subblock matrices $A_{11(j)}^{(i)}$ and $A_{12(j)}^{(i)}$ are of dimensions $n_j \times r_1$, $i = 1, \cdots, m$, $j = 1, \cdots m$.

In terms of this notation, and using some additional straightforward algebraic manipulations, we arrive at the following simplified form of the identifier (56)–(64):

$$\begin{aligned} &A_2^{(i)}\Pi^{(i)} + \Pi^{(i)}A_2^{(i)'} + I_{n_i} - ^{2})_i^2)\Pi^{(i)}C^{(i)'}C^{(i)}\Pi^{(i)} = 0 \\ &i = 1, \cdots, m \end{aligned} \tag{67}$$

$$A_{fo}^{(i)} := A_2^{(i)} - \frac{1}{\epsilon_i^2}\Pi^{(i)}C^{(i)'}C^{(i)}; \quad i = 1, \cdots, m \tag{68}$$

$$\frac{d}{dt}\eta_{(i)}^{(j)} = A_{fo}^{(j)}\eta_{(i)}^{(j)} - C^{(j)'}y_i; \quad \eta_{(i)}^{(j)}(0) = 0$$
$$i = 1, \cdots, m \quad j = 1, \cdots, m \tag{69}$$

$$\frac{d}{dt}\lambda_{(i)}^{(j)} = A_{fo}^{(j)}\lambda_{(i)}^{(j)} - C^{(j)'}u_i; \quad \lambda_{(i)}^{(j)}(0) = 0$$
$$i = 1, \cdots, p \quad j = 1, \cdots, m \tag{70}$$

$$\Phi^{(i)} = \sum_{j=1}^{m} [\eta_{(j)}^{(i)} \quad A_{fo}^{(i)^{-1}}\eta_{(j)}^{(i)} \quad \cdots \quad A_{fo}^{(i)^{-n_i+1}}\eta_{(j)}^{(i)}] \cdot \left[A_{11(i)}^{(j)} \quad -\frac{1}{\epsilon_i^2}\Sigma_3^{(i)^{-1}}C^{(i)'}G_{(i)}^{(j)}\right] + \sum_{j=1}^{p}[\lambda_{(j)}^{(i)} \quad A_{fo}^{(i)^{-1}}\lambda_{(j)}^{(i)} \quad \cdots \quad A_{fo}^{(i)^{-n_i+1}}\lambda_{(j)}^{(i)}] \cdot [A_{12(i)}^{(j)} \quad \mathbf{0}_{n_i\times r_2}]; \quad i = 1, \cdots, m \tag{71}$$

$$\frac{d}{dt}(\Xi^{-1}) = -\Xi^{-1}\left\{\sum_{i=1}^{m}\frac{1}{\epsilon_i^2}[\Phi^{(i)'}C^{(i)'} - \overline{G}_{(i)}'] \cdot [C^{(i)}\Phi^{(i)} - \overline{G}_{(i)}] - \frac{1}{\gamma^2}Q\right\}\Xi^{-1}$$
$$\Xi^{-1}(0) = Q_0^{-1} \tag{72}$$

$$\overline{G}_{(i)} := \left[\mathbf{0}_{1\times r_1} \quad \sum_{j=1}^{m} G_{(i)}^{(j)}y_j\right]; \quad i = 1, \cdots, m \tag{73}$$

$$\delta(t, y_{[0,t]}, u_{[0,t]}) = \hat{\theta}(t) \tag{74}$$

$$\dot{\hat{\theta}} = -\Xi^{-1}\sum_{i=1}^{m}\frac{1}{\epsilon_i^2}[\Phi^{(i)'}C^{(i)'} - \overline{G}_{(i)}'] \cdot \left[y_i - \sum_{i=1}^{m}\overline{G}_{(i)}\hat{\theta} - C^{(i)}\hat{x}\right]; \quad \hat{\theta}(0) = \overline{\theta}_0 \tag{75}$$

$$\dot{\hat{x}} = \overline{A}_1\hat{\theta} + A_2\hat{x} + \begin{bmatrix}\Phi^{(1)} \\ \vdots \\ \Phi^{(m)}\end{bmatrix}\Xi^{-1}\sum_{i=1}^{m}\frac{1}{\epsilon_i^2} \cdot [\Phi^{(i)'}C^{(i)'} - \overline{G}_{(i)}'][y_i - \overline{G}_{(i)}\hat{\theta} - C^{(i)}\hat{x}] + \begin{bmatrix}\frac{1}{\epsilon_1^2}\Pi^{(1)}C^{(1)'}[y_1 - \overline{G}_{(1)}\hat{\theta} - C^{(1)}\hat{x}] \\ \vdots \\ \frac{1}{\epsilon_m^2}\Pi^{(m)}C^{(m)'}[y_m - \overline{G}_{(m)}\hat{\theta} - C^{(m)}\hat{x}]\end{bmatrix}$$
$$\hat{x}(0) = \overline{x}_0. \tag{76}$$

This leads to the following counterpart of Corollary 4.1.

Corollary 5.1: Consider the worst-case identification problem for the MIMO linear system (54) and (55) with the cost function (5), and let the optimal performance level γ^* be defined as in (6). Assume that the initial weighting matrix P_0 is chosen as

$$P_0 = \text{block diagonal}\{\Pi^{(1)^{-1}}, \cdots, \Pi^{(m)^{-1}}\}$$

where $\Pi^{(i)}$, $i = 1, \cdots, m$ are the unique positive-definite solutions to the ARE's (67). Then:

1) The optimal performance level γ^* is given by

$$\gamma^* = \inf\{\gamma > 0: \text{ the solution to (72), } \Xi^{-1}(t), \text{ is positive definite for all } t \geq 0 \text{ and all possible measurement waveforms } y_{[0,\infty)}\}. \tag{77}$$

2) For each $\gamma > \gamma^*$, the identifier (67)–(76) achieves the desired performance level γ. Furthermore, if the following persistency of excitation condition:

$$\lim_{t\to\infty}\lambda_{\min}\left\{\int_0^t\sum_{i=1}^{m}\frac{1}{\epsilon_i^2}[\Phi^{(i)'}C^{(i)'} - \overline{G}_{(i)}'] \cdot [C^{(i)}\Phi^{(i)} - \overline{G}_{(i)}]\,ds\right\} = \infty \tag{78}$$

holds, and the disturbance quadruple $(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)})$ belongs to $\mathcal{L}^2$, then the parameter estimates converge to the true parameter value, i.e., $\lim_{t\to\infty}\hat{\theta}(t) = \theta$. $\diamond$

Approximate Identifier for $0 < \epsilon \ll 1$: When the intensity of the measurement noise is weak, which can be formally captured by letting $\epsilon := \max\{\epsilon_1, \cdots, \epsilon_m\} \ll 1$, the identifier (67)–(76) becomes numerically ill-conditioned for computation. To alleviate this numerical stiffness problem, we again seek an approximate identifier for sufficiently small values of ϵ which is numerically better conditioned than the identifier (67)–(76). Because of the $O(1/\epsilon_i^2)$ terms on the RHS of (71), it is numerically stable to evaluate the filtration of $A_{11(i)}^{(j)}y_j$ and $(1/\epsilon_i^2)\Sigma_3^{(i)'}C^{(i)'}G_{(i)}^{(j)}$, $i = 1, \cdots, m$, $j = 1, \cdots, m$ separately. Thus, the approximate filter may not contain fewer integrators than the identifier (67)–(76).

Let us define a $1 \times (n_i - 1)$-dimensional matrix $A_{21}^{(i)}$ and an $(n_i - 1) \times (n_i - 1)$-dimensional matrix $A_{22}^{(i)}$ by the following partitioning of the matrix $A_2^{(i)}$, $i = 1, \cdots, m$:

$$A_2^{(i)} =: \begin{bmatrix} 0 & A_{21}^{(i)} \\ \mathbf{0}_{(n-1)\times 1} & A_{22}^{(i)}\end{bmatrix}.$$

Using these matrices, the approximate identifier can be written in the following form:

$$A_{22}^{(i)}\Pi_s^{(i)} + \Pi_s^{(i)}A_{22}^{(i)'} + I_{n_i-1} - \Pi_s^{(i)}A_{21}^{(i)'}A_{21}^{(i)}\Pi_s^{(i)} = 0$$
$$i = 1, \cdots, m \tag{79}$$

$$A_{fs}^{(i)} := A_{22}^{(i)} - \Pi_s^{(i)}A_{21}^{(i)'}A_{21}^{(i)}; \quad i = 1, \cdots, m \tag{80}$$

$$\frac{d}{dt}\eta_{s(i)}^{(j)} = A_{fs}^{(j)}\eta_{s(i)}^{(j)} + \Pi_s^{(i)}A_{21}^{(i)'}y_i$$
$$\eta_{s(i)}^{(j)}(0) = 0 \quad i = 1, \cdots, m \quad j = 1, \cdots, m \tag{81}$$

$$\frac{d}{dt}\omega_{(i)}^{(j)} = A_2^{(j)}\omega_{(i)}^{(j)} + \begin{bmatrix}1 \\ \Pi_s^{(i)}A_{21}^{(i)'}\end{bmatrix}\frac{1}{\epsilon_i}[y_i - C^{(i)}\omega_{(i)}^{(j)}]$$
$$\eta_{s(i)}^{(j)}(0) = 0 \quad i = 1, \cdots, m-1 \quad j = i+1, \cdots, m \tag{82}$$

$$\frac{d}{dt}\lambda_{s(i)}^{(j)} = A_{fs}^{(j)}\lambda_{s(i)}^{(j)} + \Pi_s^{(i)}A_{21}^{(i)'}u_i$$

$$\lambda_{s(i)}^{(j)}(0) = 0 \quad i = 1, \cdots, p \quad j = 1, \cdots, m \tag{83}$$

$$\begin{aligned}\Phi_s^{(i)} = &\sum_{j=1}^{m} [\eta_{s(j)}^{(i)} \;\; A_{fs}^{(i)^{-1}} \eta_{s(j)}^{(i)} \;\; \cdots \;\; A_{fs}^{(i)^{-n_i+1}} \eta_{s(j)}^{(i)}] A_{11(i)}^{(j)} \\ &+ \sum_{j=1}^{p} [\lambda_{s(j)}^{(i)} \;\; A_{fs}^{(i)^{-1}} \lambda_{s(j)}^{(i)} \;\; \cdots \;\; A_{fs}^{(i)^{-n_i+1}} \lambda_{s(j)}^{(i)}] \\ &\cdot A_{12(i)}^{(j)}; \quad i = 1, \cdots, m\end{aligned} \tag{84}$$

$$\Phi_G^{(i)} = \sum_{j=1}^{m} \omega_{(j)}^{(i)} G_{(i)}^{(j)}; \quad i = 1, \cdots, m \tag{85}$$

$$\begin{aligned}\frac{d}{dt}(\Xi^{-1}) &= -\Xi^{-1} \left[\sum_{i=1}^{m} \overline{\Phi}^{(i)'} \overline{\Phi}^{(i)} - \frac{1}{\gamma^2} Q \right] \Xi^{-1} \\ \Xi^{-1}(0) &= Q_0^{-1}\end{aligned} \tag{86}$$

$$\overline{\Phi}^{(i)} := \begin{bmatrix} \Phi_s^{(i)'} A_{21}^{(i)'} \\ -C^{(i)'} \left[\sum_{j=1}^{m} A_{11(i)}^{(j)} y_j + \sum_{j=1}^{p} A_{12(i)}^{(j)} u_j \right] \\ \frac{1}{\epsilon_i} \left[\Phi_G^{(i)'} C^{(i)'} - \sum_{j=1}^{m} G_{(i)}^{(j)'} y_j \right]. \end{bmatrix}' \qquad i = 1, \cdots, m \tag{87}$$

$$\delta_s(t, y_{[0,t]}, u_{[0,t]}) = \hat{\theta}_s(t) \tag{88}$$

$$\begin{aligned}\dot{\hat{\theta}}_s &= -\Xi^{-1} \sum_{i=1}^{m} \overline{\Phi}^{(i)'} \frac{1}{\epsilon_i} \left[y_i - \sum_{i=1}^{m} \overline{G}_{(i)} \hat{\theta}_s - C^{(i)} \hat{x}_s \right] \\ \hat{\theta}_s(0) &= \overline{\theta}_0\end{aligned} \tag{89}$$

$$\begin{aligned}\dot{\hat{x}}_s &= \overline{A}_1 \hat{\theta}_s + A_2 \hat{x}_s + \begin{bmatrix} \Psi^{(1)} \\ \vdots \\ \Psi^{(m)} \end{bmatrix} \Xi^{-1} \sum_{i=1}^{m} \overline{\Phi}^{(i)'} \frac{1}{\epsilon_i} \\ &\quad \cdot [y_i - \overline{G}_{(i)} \hat{\theta}_s - C^{(i)} \hat{x}_s] \\ &\quad + \begin{bmatrix} \begin{bmatrix} 1 \\ \Pi_s^{(1)} A_{21}^{(1)'} \end{bmatrix} \frac{1}{\epsilon_1} [y_1 - \overline{G}_{(1)} \hat{\theta}_s - C^{(1)} \hat{x}_s] \\ \vdots \\ \begin{bmatrix} 1 \\ \Pi_s^{(m)} A_{21}^{(m)'} \end{bmatrix} \frac{1}{\epsilon_m} [y_m - \overline{G}_{(m)} \hat{\theta}_s - C^{(m)} \hat{x}_s] \end{bmatrix} \\ \hat{x}_s(0) &= \overline{x}_0\end{aligned} \tag{90}$$

$$\Psi^{(i)} = \left[\begin{bmatrix} \mathbf{0}_{1 \times r_1} \\ \Phi_s^{(i)} \end{bmatrix} \quad \Phi_G^{(i)} \right]; \quad i = 1, \cdots, m. \tag{91}$$

To state the counterpart of Corollary 4.2, we introduce the following set of measurement waveforms.

$\mathcal{LBP}_{EPW}^{A}(L, M, \overline{\rho}, T_m, K)$*:* This is the set of all waveforms $y_{[0,\infty)}$ that satisfy the following four conditions for some positive constants L, M, $\overline{\rho}$, T_m, and integer K.

1) There exist $K < \infty$ time instances $0 < t_1 < \cdots < t_K < \infty$, such that
$$\min_{k \in \{1, \cdots, K\}} t_k - t_{k-1} \geq T_m.$$
2) $\| \sum_{i=1}^{m} A_{11}^{(i)} [y_i(t') - y_i(t)] + \sum_{i=1}^{p} A_{12}^{(i)} [u_i(t') - u_i(t)] \|_2 \leq L|t' - t|$, and $\| \sum_{i=1}^{m} G^{(i)} [y_i(t') - y_i(t)] \|_2 \leq L|t' - t| \;\; \forall t', t \in [0, t_1]$, or $\forall t', t \in [t_k, t_{k+1}]$, $k = 1, \cdots, K-1$, or $\forall t', t \in [t_K, \infty)$.
3) $\| \sum_{i=1}^{m} A_{11}^{(i)} y_i(t) + \sum_{i=1}^{p} A_{12}^{(i)} u_i(t) \|_2 \leq M$ and $\| \sum_{i=1}^{m} G^{(i)} y_i(t) \|_2 \leq M$ for all $t \geq 0$.
4) $\Xi(t) \geq (t+1)\overline{\rho} I$, for all $t \geq 0$.

$\mathcal{W}[\mathcal{LBP}_{EPW}^{A}(L, M, \overline{\rho}, T_m, K)]$*:* This is the set of all disturbance quadruples $(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)})$ that lead to a measurement waveform $y_{[0,\infty)}$ that belongs to the set $\mathcal{LBP}_{EPW}^{A}(L, M, \overline{\rho}, T_m, K)$.

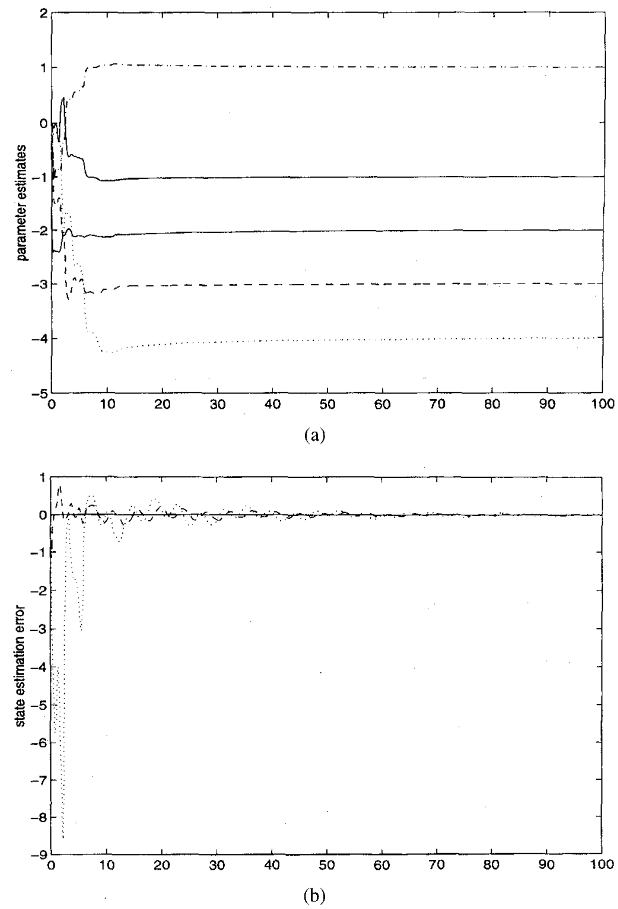


Fig. 1. Response of the full-order identifier without any disturbances and with no discrepancy between initial states and their estimates: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

Corollary 5.2: Consider the worst-case identification problem for the MIMO linear system (54) and (55) with the cost function (5), and let the optimal performance level γ^* be defined as in (6). Assume that the initial weighting matrix P_0 is chosen as
$$P_0 = \text{block diagonal}\,\{\Pi^{(1)^{-1}}, \cdots, \Pi^{(m)^{-1}}\}$$
where $\Pi^{(i)}$, $i = 1, \cdots, m$ are the unique positive-definite solutions of the ARE's (67). For any $\gamma > \gamma^*$, where γ^* is defined by (77) and any disturbance within the set $\mathcal{W}[\mathcal{LBP}_{EPW}^{A}(L, M, \overline{\rho}, T_m, K)]$, the approximate identifier δ_s given by (79)–(91) achieves the disturbance attenuation

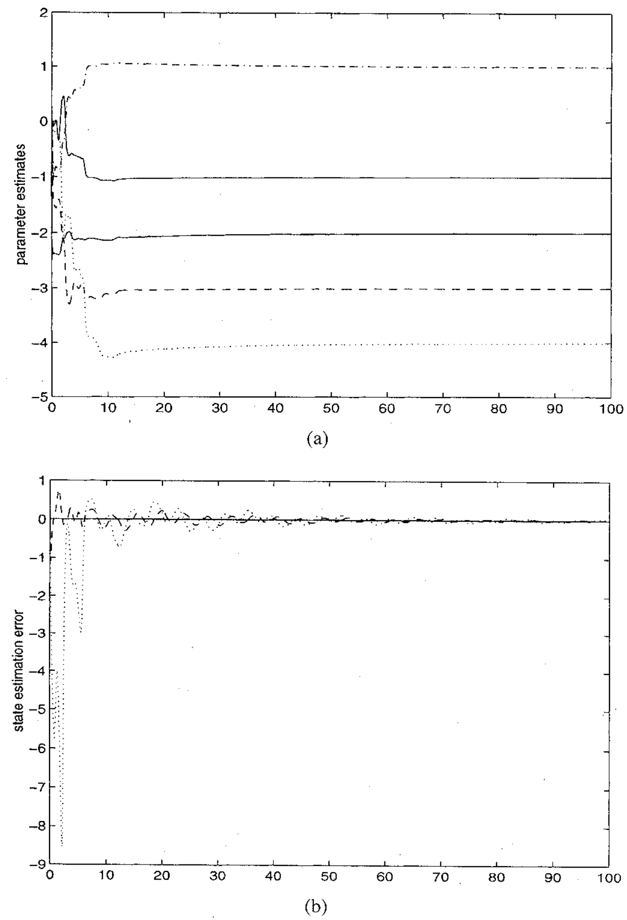


Fig. 2. Response of the reduced-order identifier without any disturbances and with no discrepancy between initial states and their estimates: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

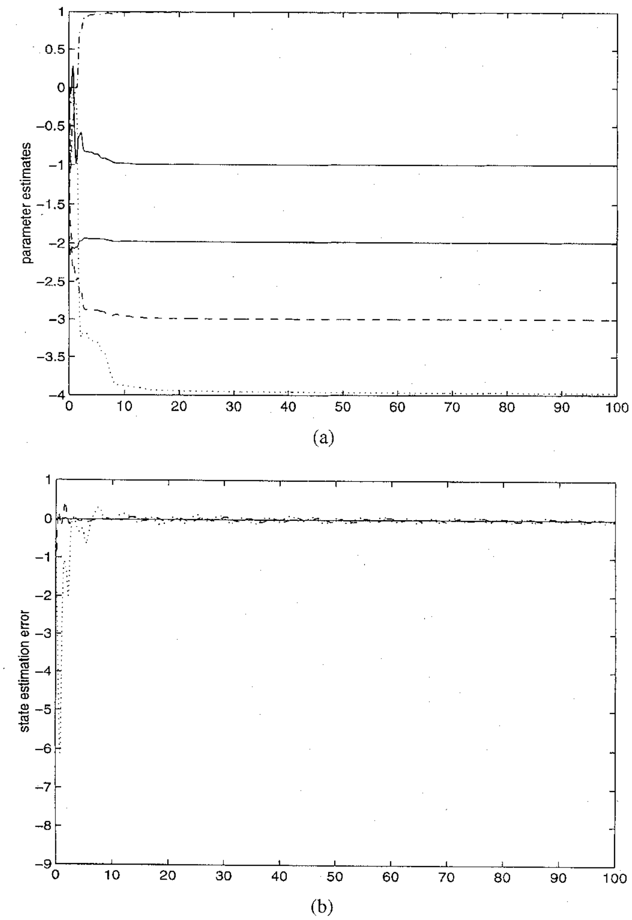


Fig. 3. Response of the conventional identifier without any disturbances and with no discrepancy between initial states and their estimates: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

level γ for sufficiently small $\epsilon > 0$, i.e.,

$$\sup_{(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}) \in \mathcal{W}[\mathcal{LBP}^A_{EPW}(L, M, \overline{p}, T_m, K)]} \cdot J_\gamma(\delta_s;\, x_0,\, \theta,\, w_{[0,\infty)},\, v_{[0,\infty)}) = 0. \tag{92}$$

Furthermore, if $(x_0, \theta, w_{[0,\infty)}, v_{[0,\infty)}) \in \mathcal{W}[\mathcal{LBP}^A_{EPW}(L, M, \overline{p}, T_m, K)] \bigcap \mathcal{L}^2$, then the parameter estimates $\hat{\theta}_s(t)$ converge to the true parameter value θ as $t \to \infty$ for sufficiently small ϵ. ◇

VI. An Example

To illustrate the results obtained in the previous sections, and particularly Section IV, we will now consider a third-order dynamic system with five unknown parameters

$$H(s) = \frac{b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s^1 + a_0}. \tag{93}$$

Let

$$\theta = [-a_2 \quad -a_1 \quad -a_0 \quad b_1 \quad b_0]'$$

where the true value of the parameter vector θ is fixed at

$$\theta = [-2 \quad -3 \quad -4 \quad 1 \quad -1]'.$$

In the form of state-space representation (3a) and (3b), the system can be equivalently written as

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x + (yA_{11} + uA_{12})\theta + w \tag{94}$$

$$y = [1 \quad 0 \quad 0]x + \epsilon v \tag{95}$$

where

$$A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A_{11} = [I_3 \quad \mathbf{0}_{3\times 2}]; \quad A_{12} = \begin{bmatrix} \mathbf{0}_{1\times 3} & \mathbf{0}_{1\times 2} \\ \mathbf{0}_{2\times 3} & I_2 \end{bmatrix}.$$

The cost function associated with this system is given by (5) with

$$\overline{\theta}_0 = \mathbf{0}_{5\times 1}; \quad \overline{x}_0 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

$$Q_0 = 0.01 I_5$$

and P_0 being the solution to ARE (30). The design parameter ϵ is taken to be 0.01, and the incremental weighting matrix

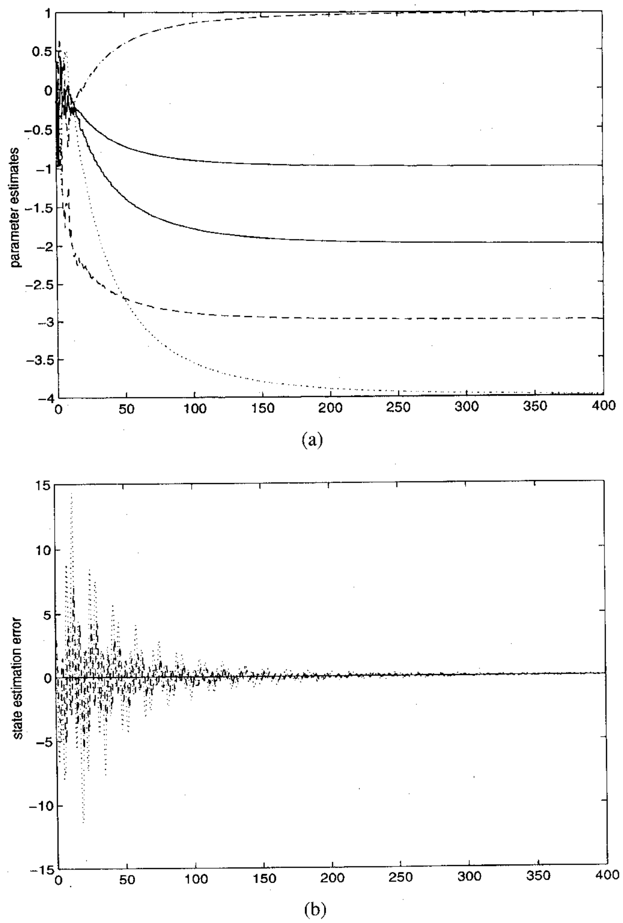


Fig. 4. Response of the full-order identifier with exponentially decaying disturbances: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

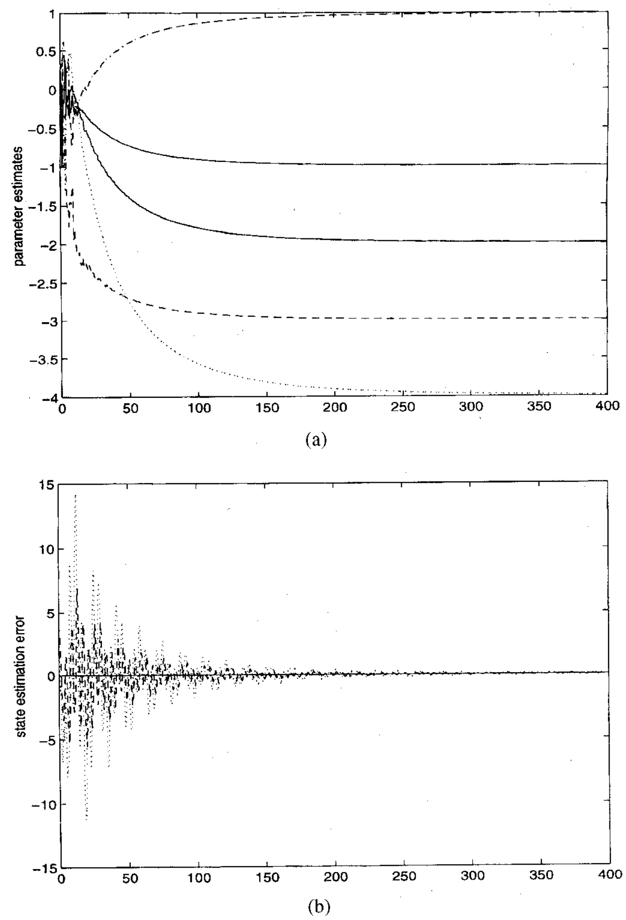


Fig. 5. Response of the reduced-order identifier with exponentially decaying disturbances: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

function $Q(t, y_{[0,t]}, u_{[0,t]})$ is chosen to be $(1/\epsilon^2)\Phi' e_1 e_1' \Phi$, as recommended in Remark 4.1 for the full-order identifier (30)–(38) which implies that $\gamma^* = 1$. For the reduced-order identifier (40)–(48), this matrix function is chosen to be $(\Phi_s' A_{21}' - A_1' e_1)(A_{21}\Phi_s - e_1' A_1)$.[1]

The desired performance level is fixed at $\gamma = 1.2$ which should be achievable by both the full-order and reduced-order identifiers due to the choice of Q.

To compare the performances of these two identifiers (that is, the full-order and reduced-order ones) with that of a conventional one, we further design an identifier using the Kreisselmeier observer of Section III with the parameter vector $f = (-15, -75, -125)'$, or equivalently, a prefilter with transfer function

$$\frac{1}{(s+5)^3}.$$

For this Kreisselmeier observer, the initial covariance matrix was set to be $10^5 I_5$, and for the comparison to be fair we have not included any covariance resetting in the LS algorithm.

We now present a simulation-based comparison of the three identifiers: full-order identifier (cf. Theorem 4.1), reduced-order identifier (cf. Theorem 4.2), and conventional identifier (cf. Section III). The input signal $u(t)$ was chosen to be

$$u(t) = 4\sin(0.4t) + 2\sin(1.5t) + 2\sin(4t)$$

for all simulations. The measurement noise v was set to be identically zero in all the simulations, consistent with our real objective, as stated in Section II. The robustness property of the full-order and reduced-order identifiers holds true even when there is measurement noise present. However, the parameter estimates may be biased when v is taken as a general $\mathcal{L}^\infty$ signal.

First, we simulated the system with these three identifiers in the absence of any disturbance input and with the initial condition of the system set to be $x_0 = (3, 0, 0)'$ which is the same as the initial estimate for it. The results are presented in Figs. 1–3.

We observe that the parameter estimates exhibit fast convergence to their true values for all three identifiers. The initial convergence rate is faster for the conventional identifier, but it falls somewhat behind the full-order and reduced-order identifiers in the long run.

Next, we simulated the system with the same three identifiers against sizable initial deviations in the unmeasured state

[1] This choice is made to avoid using additional integrators to generate Q. It is asymptotically equal to $(1/\epsilon^2)\Phi' e_1 e_1' \Phi$ as $\epsilon \to 0$.

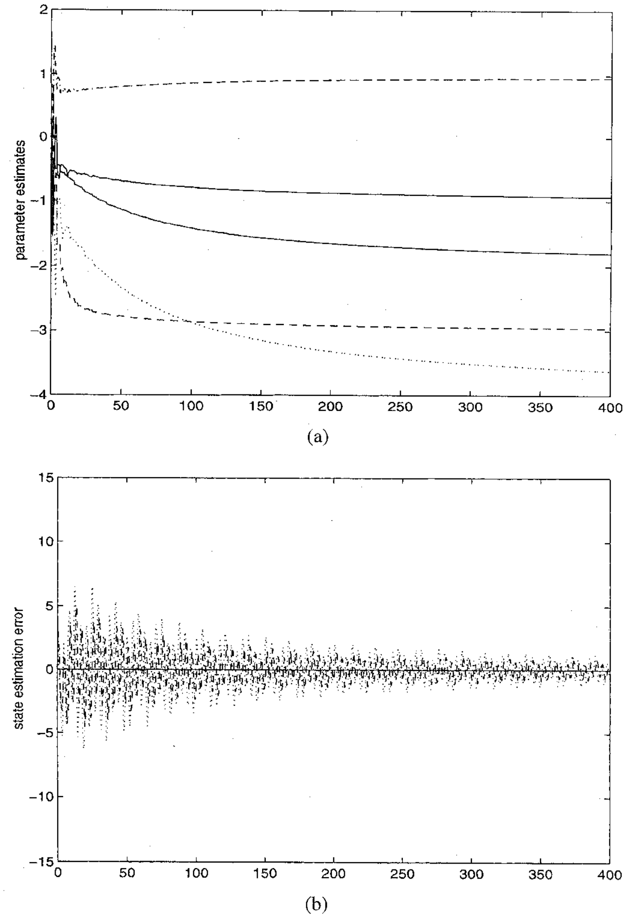


Fig. 6. Response of the conventional identifier with exponentially decaying disturbances: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

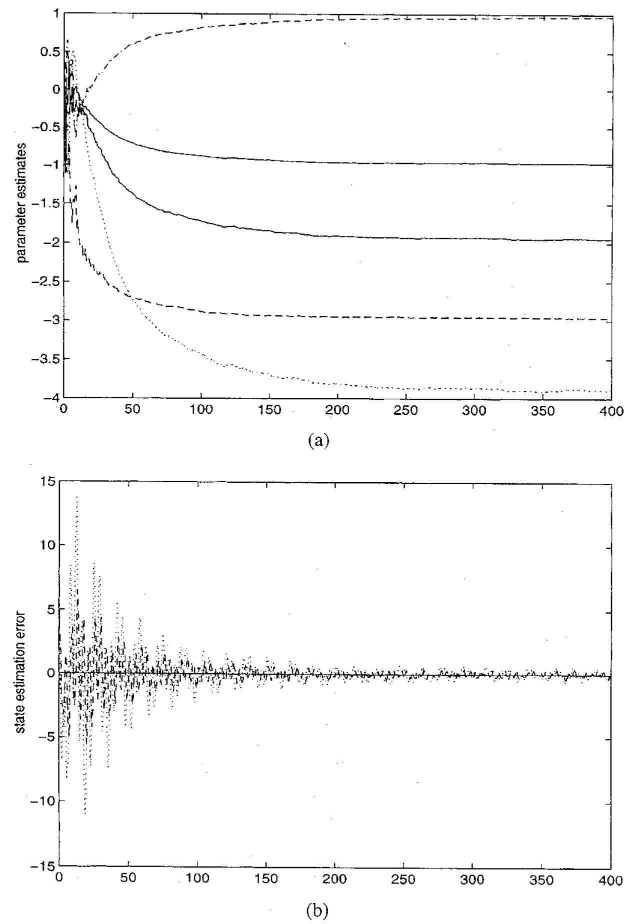


Fig. 7. Response of the full-order identifier with band-limited white noise disturbances: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

variables; namely, the initial condition of the system was fixed at $x_0 = (3, 4, 8)'$ which is substantially different from the initial state estimate $\overline{x}_0 = (3, 0, 0)'$. This initial bias in the estimates of the unmeasured states can equivalently be viewed as injecting exponentially decaying disturbances into the system. The simulation results are depicted in Figs. 4–6. The theory developed in the paper and the existing theory on the Kreisselmeier observer predicts that the parameter estimates should converge to their true values asymptotically for all three identifiers. From the graphs, we see that the convergence rate is much faster for the full-order and reduced-order identifiers than for the conventional identifier. These results clearly demonstrate the superior robustness property of the worst-case parameter identifier over the Kreisselmeier observer design.

While the theory developed in this paper only proves that the estimates of the unknown parameters converge to their true values under $\mathcal{L}^2$ disturbances with persistent excitation, we have to emphasize that this is only a sufficient condition for parameter convergence, and it is expected that similar results hold even for more general disturbances. Indeed, it has been observed through simulations that the parameter estimates converge to their true values under white Gaussian noise. It is plausible that a rigorous proof can be devised for the convergence of parameter estimates under some standard stochastic assumptions [12], [13], in view of the well-established connection between zero-sum differential games and the risk-sensitive stochastic optimal control problem [14], but this has not been carried out in this paper as the main concern here has been the optimal attenuation of worst-case disturbance inputs. As one example of a non-$\mathcal{L}^2$ disturbance, we simulated the system with the components of the disturbance w taken as independent band-limited white noises, each with variance 0.01 and sample rate 1, and with the initial condition set at $x_0 = (3, 4, 8)'$. The numerical results are depicted in Figs. 7–9.

We observe that the parameter estimates for the full-order and the reduced-order identifiers again converge to their true values asymptotically. In this case also, the performances of the worst-case identifiers are superior to that of the conventional identifier.

Finally, we present a set of simulation results to demonstrate the fact that the performance of the Kreisselmeier observer depends very much on the choice of the parameter vector f. In the previous simulations, to obtain improved performance from the conventional identifier, we had chosen the poles of

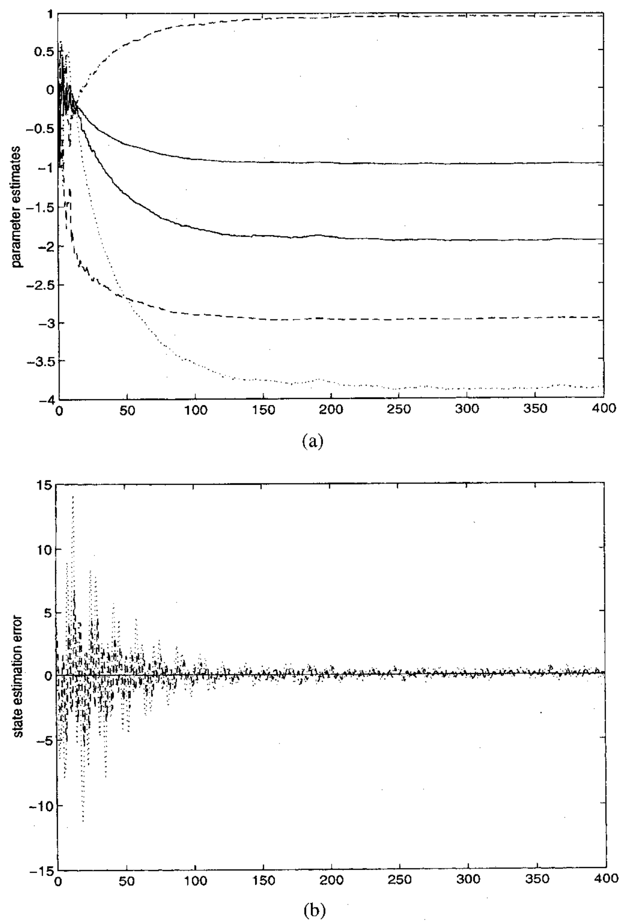


Fig. 8. Response of the reduced-order identifier with band-limited white noise disturbances: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

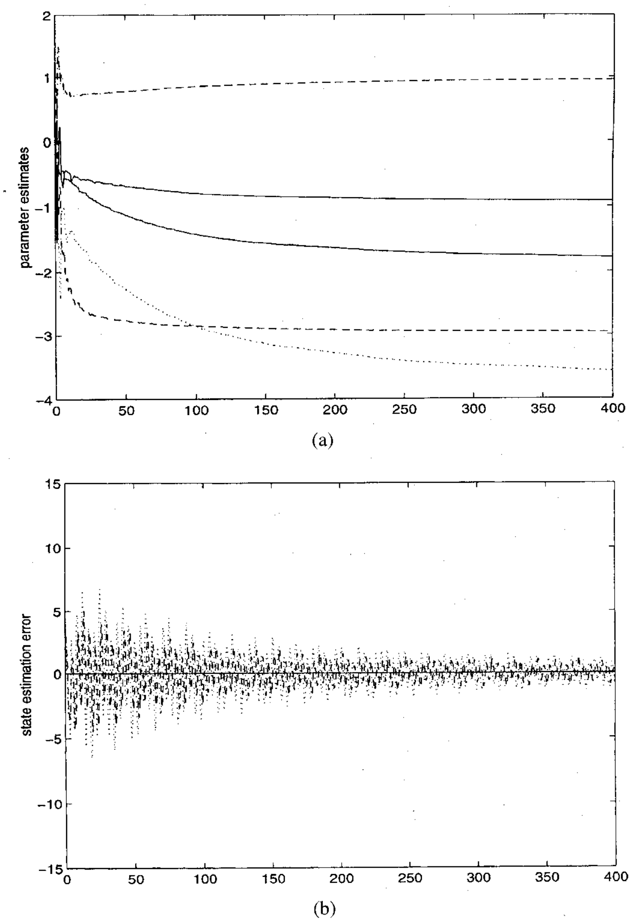


Fig. 9. Response of conventional identifier with band-limited white noise disturbances: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

the prefilter which equivalently determine the vector f, to be at -5, -5, and -5, based on the fact that the true system admits a pole at -1.65. Let us now pretend that we are not aware of this *a priori* information on the underlying system (a realistic scenario) and design the Kreisselmeier observer with the parameter vector f set at $f = (-3, -3, -1)'$, or equivalently, a prefilter with transfer function

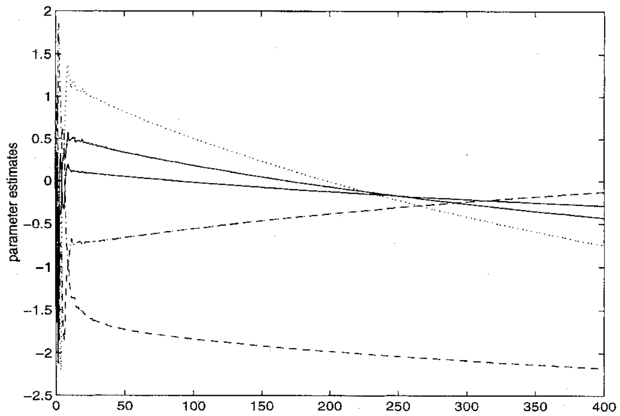
$$\frac{1}{(s+1)^3}.$$

For this Kreisselmeier observer, the initial covariance matrix and the initial state are set at exactly the same values as before, that is $10^5 I_5$ and $(3, 4, 8)'$, respectively, and the disturbance inputs are taken to be identically zero. The results obtained are presented in Fig. 10 which should be compared against Figs. 4 and 5. Clearly, the conventional identifier is much inferior to either the full-order or the reduced-order worst-case identifier. This clearly demonstrates the inapplicability of the conventional identifier for a wide range of parameter values without sufficient *a priori* information on the underlying system. Therefore, for the purpose of robust on-line parameter identification, the worst-case identifiers are clearly a much superior choice.

When we compare the performances of the two, however, we see almost no difference between the full-order identifier and the reduced-order identifier for both parameter estimation and state estimation in all of the simulation runs above. This is in line with the results of Corollary 4.2. The only difference is that the simulation of the full-order identifier requires much more computational power than that of the reduced-order identifier. In the last set of simulations, the simulation for the full-order identifier took 409 005 635 flops which is 70% more than that of the reduced-order identifier, which was 240 348 937 flops.[2] In this respect, the reduced-order identifier is clearly superior, as to be expected.

Hence, these simulation results clearly indicate that the full- and reduced-order worst-case identifiers outperform the conventional identifier, and the reduced-order identifier is numerically better conditioned than the full-order identifier when ϵ is small. We have also simulated the response of the worst-case identifiers with unmodeled dynamics in addition to the uncertainty arising from mismatch of initial conditions and white noise disturbance inputs. The plots (not included here) have shown that the parameter estimates are robust with

[2]These simulations were implemented in Matlab Simulink. Both flop counts were generated using the simulation algorithm GEAR.



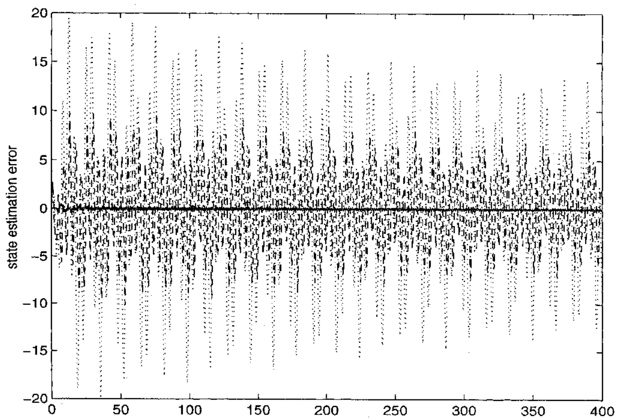


Fig. 10. Response of the conventional identifier with no disturbances, but with some discrepancy between true and estimated values of the initial state, in the absence of any *a priori* information on the system: (a) parameter estimates $\hat{\theta}$ and (b) state estimation errors $x - \hat{x}$.

respect to the unmodeled dynamics and converge to constant values that lie in the interval $(-4, 1)$. This is consistent with and corroborates the statements of Theorems 4.1 and 4.2.

VII. Conclusion

In this paper, we have constructed worst-case full-order and approximate identifiers for both SISO and MIMO linear systems. By representing the system in output injection form in state space, we have brought the worst-case identification problem with partial state measurements into one where the *cost-to-come function* method can be applied as in [5]. In the SISO case, the worst-case identifier contains a Kreisselmeier observer as part of its structure; here the observer parameters are given some optimal values which are obtained from the solution of an algebraic Riccati equation. Outside this, the identifier presented in this paper is significantly different from the conventional LS identifier in that it contains additional dynamics for the state estimates. As the noise intensity in the measurement decreases, the full-order identifier becomes numerically ill-conditioned. To alleviate this difficulty, we have obtained an approximate (reduced-order) identifier which is simpler in structure by two integrators and is numerically much better conditioned for implementation. These results have been generalized to the MIMO case and stated without detailed proofs. An extensive simulation study has been included to demonstrate the effectiveness of the identification scheme presented.

The results of this paper can be extended to nonlinear systems with partial state measurements, provided that they are in the following output injection form:

$$\dot{x} = A_2 x + A_1(y, u)\theta$$
$$y = Cx$$

where A_2 and C are known constant matrices such that the pair (A_2, C) is observable and A_1 is a known nonlinear function of y and u. The main difficulty in this case is to obtain a coordinate-free characterization of nonlinear systems that can be transformed into this output injection form; once this is done, then the extension is immediate as in the case of a similar extension discussed in [5] for the full-state. A future direction of research would be to study the performance of these worst-case identifiers when used in the control loop for the design of robust adaptive controllers for uncertain linear and nonlinear systems.

Acknowledgment

Fruitful discussions with Dr. G. Didinsky are gratefully acknowledged.

References

[1] K. J. Åström, *Introduction to Stochastic Control.* New York: Academic, 1970.

[2] G. Kreisselmeier, "Adaptive observers with exponential rate of convergence," *IEEE Trans. Automat. Contr.,* vol. AC-22, pp. 2–8, 1977.

[3] A. S. Morse and F. M. Pait, "MIMO design models and internal regulators for cyclicly switched parameter-adaptive control systems," *IEEE Trans. Automat. Contr.,* vol. 39, pp. 1809–1818, Sept. 1994.

[4] M. de Mathelin and M. Bodson, "Multivariable adaptive control-identifiable parameterizations and parameter covergence," *IEEE Trans. Automat. Contr.,* vol. 39, pp. 1612–1616, Aug. 1994.

[5] G. Didinsky, Z. Pan, and T. Başar, "Parameter identification for uncertain plants using H^∞ methods," *Automatica,* vol. 31, pp. 1227–1250, Sept. 1995.

[6] G. Didinsky, "Design of minimax controllers for nonlinear systems using cost-to-come methods," Ph.D. dissertation, Univ. Illinois, Urbana, IL, Aug. 1994.

[7] G. Didinsky and T. Başar, "Design of minimax controllers for linear systems with nonzero initial states under specified information structures," *Int. J. Robust Nonlinear Contr.,* vol. 2, pp. 1–30, 1992.

[8] G. Didinsky, T. Başar, and P. Bernhard, "Structural properties of minimax controllers for a class of differential games arising in nonlinear H^∞-control," *Syst. Contr. Lett.,* vol. 21, pp. 433–441, Dec. 1993.

[9] ——, "Structural properties of minimax controllers for a class of differential games arising in nonlinear H^∞-control and filtering," in *Proc. IEEE 32nd Conf. Decision Contr.,* Dec. 1993, pp. 184–189.

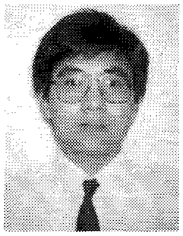
[10] T. Başar and P. Bernhard, *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach,* 2nd ed. Boston, MA: Birkhäuser, 1995.

[11] Z. Pan and T. Başar, "H^∞-optimal control for singularly perturbed systems—Part I: Perfect state measurements," *Automatica,* vol. 29, pp. 401–423, Mar. 1993.

[12] T. L. Lai and C. Z. Wei, "Least squares estimates in stochastic regression with applications to identification and control of dynamic systems," *Ann. Math. Stats.,* vol. 10, pp. 154–166, 1982.

[13] L. Guo and H. Chen, "The Åström–Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers," *IEEE Trans. Automat. Contr.,* vol. 36, pp. 802–812, July 1991.

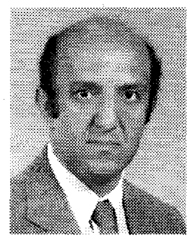
[14] M. R. James, J. Baras, and R. J. Elliott, "Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems," *IEEE Trans. Automat. Contr.,* vol. 39, pp. 780–792, Apr. 1994.



Zigang Pan (S'92–M'96) received the B.S. degree in automatic control from Shanghai Jiao Tong University in 1990, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana-Champaign, in 1992 and 1996, respectively.

During the period of January 1996–August 1996, he has been a Research Engineer at the Center for Control Engineering and Computation of the University of California, Santa Barbara. In September 1996, he will join the Polytechnic University, Brooklyn, NY, as an Assistant Professor in the Department of Electrical Engineering. His current research interests include perturbation theory for robust control systems, robust parameter identification, and robust adaptive control with applications to communication networks and robotics.

Dr. Pan received the 1992 Henry Ford II Scholarship Award, the 1993 M. E. Van Valkenburg Fellowship Award, and the 1994 Robert T. Chien Memorial Award, all at the University of Illinois. He is also the coauthor of a paper that won the 1995 George Axelby Outstanding Paper Award. He is a member of SIAM and the Phi Kappa Phi Honor Society.



Tamer Başar (S'71–M'73–SM'79–F'83) received the B.S.E.E. degree from Robert College in 1969 and the M.S., M.Phil., and Ph.D. degrees in engineering and applied science from Yale University, New Haven, CT, in 1972.

After being at Harvard University, Marmara Research Institute, and Bogaziçi University, he joined the University of Illinois at Urbana-Champaign in 1981, where he is currently a Professor of Electrical and Computer Engineering. He has spent sabbatical years at Twente University of Technology, the Netherlands, 1978–1979, and INRIA, France 1987–1988, and 1994–1995. He has authored or coauthored more than 150 journal articles and book chapters, and also numerous conference publications, in the general areas of optimal, robust and nonlinear control, dynamic games, stochastic control, estimation theory, stochastic processes, information theory, and mathematical economics. He is coauthor of the texts *Dynamic Noncooperative Game Theory* (New York: Academic, 1982; 2nd ed., 1995) and *H^∞-Optimal Control and Related Minimax Design Problems* (New York: Birkhäuser, 1991; 2nd ed., 1995).

Dr. Başar carries memberships in several scientific organizations, among which are SIAM, SEDC, and ISDG. Currently, he is the Deputy Editor-in-Chief of the IFAC journal *Automatica,* the Managing Editor of the *Annals* of the International Society of Dynamic Games, and Associate Editor of various journals, among which are the *Journal of Economic Dynamics and Control,* and *Systems & Control Letters.* Among some of the recent awards and recognitions he has received are: the Medal of Science of Turkey (1993), Distinguished Member Award of the IEEE Control Systems Society (1993), and the George S. Axelby Outstanding Paper Award of the same Society (1995) for a 1994 paper he has coauthored.