Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

H-Infinity Parameter Estimation for State-Space Models

Kiriakos Kiriakidis and Richard T. O'Brien, Jr.

Abstract— This paper investigates the nonlinear problem of parameter identification for systems with time-varying state-space models using the idea of an extended filter. To treat uncertainty in the data or the model, the paper follows the approach of H-infinity filtering. It shows that the parameter estimation error satisfies exactly albeit locally the prescribed gain criterion as long as the extended filter meets a modified—tighter—criterion. In the infinite horizon case, the paper also shows that the modified H-infinity criterion leads to an asymptotically stable state estimation error without assuming zero-state observability of the associated error equation.

I. INTRODUCTION

Over the last decade, the problem of estimation and filtering in an $\mathcal{H}_\infty$ setting associated with discrete-time models has been investigated in the literature [1], [2], [3], [4]. The particular problem of parameter identification using an $\mathcal{H}_\infty$ criterion, however, has received little attention. For a time-invariant system, a technique based on transforming its ARMA model to a state-space one has been reported [5]. Using the extended filter approach [6], an $\mathcal{H}_\infty$-based parameter identification algorithm—with guaranteed convergence and performance for time-invariant models—has been proposed [7].

The main issue with parameter estimation based on state-space models is that it results in the intractable problem of solving a Hamilton-Jacobi equation associated with a nonlinear state-space model. Linearization of the model is an alternative but it implies that the $\mathcal{H}_\infty$ criterion is no longer met exactly. Another is an issue intrinsic to the $\mathcal{L}_2$ gain framework that encompasses the $\mathcal{H}_\infty$ criterion. To prove that the state estimation error equation is asymptotically stable one must assume it to be zero-state observable from the parameter estimation error; refer to [8]. At the same time, zero-state observability is difficult (if not impossible) to check. Although the issue of asymptotic stability does not arise in the finite horizon case, it is essential in parameter identification where the filter's horizon is infinite.

First, the paper shows that one can guarantee that the $\mathcal{H}_\infty$ performance criterion is met albeit locally by designing the extended filter in order to meet a slightly modified $\mathcal{H}_\infty$ criterion. Second, the paper analyzes the internal stability of the filter and finds that the modified $\mathcal{H}_\infty$ criterion circumvents the zero-state observability assumption and implies asymptotic stability, at least locally.

The next section presents the modeling assumptions and the filter structure. Section III defines the filter's objective and provides a pointwise sufficient condition in order to meet it. Then, Section IV presents the extended filter's modified objective and design. The analysis of the extended filter's performance follows in Section V. Section VI investigates the internal stability of the (exact) augmented state error equation and Section VII concludes the paper.

II. MODEL ASSUMPTIONS

Consider the following set of linear time-varying models:

$$\begin{aligned} x_{l+1} &= A_l(\theta)\, x_l \\ y_l &= C_l(\theta)\, x_l + w_l \end{aligned} \tag{2.1}$$

where the sequence w_l is a disturbance. (To simplify presentation but without loss of generality, the disturbance term has been omitted from the state equation.)

Assumption 1: The parameter vector, θ, is constant and belongs to the set $D_\theta \subset \Re^{n_\theta}$.

Assumption 2: The state equation is stable for all $\theta \in D_\theta$.

Let $\|v\|^2_{2,[0,J]} = \sum_{l=0}^{l=J} v_l^T v_l$, where J can be finite or $J = \infty$, denote the $\mathcal{L}_2$ norm of a sequence v_l.

Assumption 3: The disturbance w is such that $\|w\|^2_{2,[0,J]} < \infty$.

Based on the Assumptions 1-3, the trajectories of the state equation lie within a set $D_x \subset \Re^{n_x}$. From Assumption 3, if $J = \infty$, then the disturbance $w_l \to 0$ as $l \to \infty$; thus, $\|w_l\| \le c_0$ for all $l \in [0, \infty)$.

To extract the unknown parameters using state-estimation methods, the parameter vector, θ, needs to be part of an augmented state vector as follows:

$$z_l = \begin{bmatrix} x_l \\ \theta_l \end{bmatrix}$$

Because the parameter vector is constant, its drift dynamics are

$$\theta_{l+1} = \theta_l$$

This work was supported by the Office of Naval Research Grant N0001403WR20292

Kiriakos Kiriakidis is with the Faculty of Systems Engineering, United States Naval Academy, Annapolis, MD 21402, USA kiriakid@usna.edu

Richard T. O'Brien, Jr. is with the Faculty of Systems Engineering, United States Naval Academy, Annapolis, MD 21402, USA riobrien@usna.edu

In turn, one obtains the augmented model below:

$$\begin{aligned} z_{l+1} &= f_l(z_l) \\ y_l &= h_l(z_l) + w_l \\ \theta_l &= L\, z_l \end{aligned} \qquad (2.2)$$

where the nonlinear mappings f and h, and the matrix L are given in the Appendix. As an estimator for the parameter vector or, equivalently, the output-of-interest, θ, one may assume the following filter structure, based on the model (2.2):

$$\begin{aligned} \hat{z}_{l+1} &= f_l(\hat{z}_l) + N_l[y_l - h_l(\hat{z}_l)] \\ \hat{\theta}_l &= L\hat{z}_l \end{aligned} \qquad (2.3)$$

where N_l is the filter's gain. A projection mechanism may be employed to guarantee that $\hat{x}_l \in D_x$ and $\hat{\theta}_l \in D_\theta$.

III. PERFORMANCE CRITERION

The design of the filter's gain is based on the dynamics of the error in the state estimate, $e_{z,l} := z_l - \hat{z}_l$, and the error in the output-of-interest, $e_{\theta,l} := \theta_l - \hat{\theta}_l$. The governing equations for the above errors follow from (2.2) and (2.3)

$$\begin{aligned} e_{z,l+1} &= f_l(z_l) - f_l(\hat{z}_l) - N_l[h_l(z_l) - h_l(\hat{z}_l) + w_l] \\ e_{\theta,l} &= Le_{z,l} \end{aligned} \qquad (3.1)$$

In the $\mathcal{H}_\infty$ setting, the objective of the filter is to make the $\mathcal{L}_2$ gain from the disturbance w to the estimation error e_θ less than some (time-invariant) tolerance γ. In this sense, a small disturbance will produce a small error in the output-of-interest. At the same time, a small tolerance will result in a robust filter. Using a time-varying tolerance, γ_l, has certain advantages, which have been investigated in [7]. With this in mind, one may express the filter's objective as follows:

$$\|e_\theta\|^2_{2,[0,J]} < \|\gamma w\|^2_{2,[0,J]} + e^T_{z,0}\, R\, e_{z,0} \qquad (3.2)$$

where J is the filter's horizon and the second term on the right-hand-side reflects initial condition error.

Lemma 4: Consider any trajectory of the error dynamics (3.1). Suppose there exists a p.d. function $V_l = V_l(e_{z,l})$ such that $V_0(e_{z,0}) = e^T_{z,0}\, R\, e_{z,0}$ and

$$\Delta V_l(e_{z,l}) < \gamma_l^2\, \|w_l\|_2^2 - \|e_{\theta,l}\|_2^2 \qquad (3.3)$$

where $\Delta V_l(e_{z,l}) := V_{l+1}(e_{z,l+1}) - V_l(e_{z,l})$. Then, inequality (3.2) holds.

Proof: From the hypothesis, letting the index l vary from 0 to $J-1$ and adding the resulting inequalities one gets

$$V_J(e_{z,J}) < \|\gamma w\|^2_{2,[0,J]} - \|e_\theta\|^2_{2,[0,J]} + V_0(e_{z,0}) \qquad (3.4)$$

which proves the proposition. ■

In the sequel, the design and analysis of the extended filter assume a time-varying tolerance, which the time-invariant tolerance is a special case of.

IV. THE EXTENDED FILTER

The problem of solving for the filter's gain, N_l, based on the exact error equation (3.1) can be circumvented by using the concept of the extended filter [6]. Applying the mean value theorem, one expresses the nonlinear mappings in (2.2) as follows (p. 867, [9])

$$\begin{aligned} f_l(z_l) &= f_l(\hat{z}_l) + F_l(\hat{z}_l)(z_l - \hat{z}_l) + g_{1,l}(\hat{z}_l, z_l - \hat{z}_l) \\ h_l(z_l) &= h_l(\hat{z}_l) + H_l(\hat{z}_l)(z_l - \hat{z}_l) + g_{2,l}(\hat{z}_l, z_l - \hat{z}_l) \end{aligned} \qquad (4.1)$$

where F and H, given in the Appendix, denote the Jacobian matrices of f and h, respectively.

From (4.1) and (3.1), one gets the error equation

$$\begin{aligned} e_{z,l+1} &= [F_l(\hat{z}_l) - N_l H_l(\hat{z}_l)]\, e_{z,l} - N_l w_l + g_{3,l}(\hat{z}_l, e_{z,l}) \\ e_{\theta,l} &= L\, e_{z,l} \end{aligned} \qquad (4.2)$$

The remainder terms satisfy growth bounds as follows (p. 147, [10]):

$$\|g_{i,l}(\hat{z}_l, e_{z,l})\|_2 \le \mu_i \|e_{z,l}\|_2^2,\ i = 1, 2, 3$$

for some positive μ_i.

From the modeling assumptions in Section II, it follows that $\|F_l(\hat{z}_l)\|$ and $\|H_l(\hat{z}_l)\|$, where $\|\cdot\|$ denotes a matrix norm, are bounded for all $l \in [0, \infty)$.

Example 5: Consider the following scalar model:

$$\begin{aligned} x_{l+1} &= \theta\, x_l \\ y_l &= x_l + w_l \end{aligned} \qquad (4.3)$$

where $\theta \in D_\theta = \{\theta \,|\, |\theta| \le 0.5\}$ and $x \in D_x = \{x \,|\, |x| \le 1.5\}$. The Jacobian matrix F evaluated along an estimated trajectory $\hat{z}$ is

$$F_l(\hat{z}_l) = \begin{bmatrix} \hat{\theta}_l & \hat{x}_l \\ 0 & 1 \end{bmatrix} \qquad (4.4)$$

Using the absolute matrix norm [11], one has

$$\|F_l(\hat{z}_l)\| < 2\ \max(|\hat{\theta}_l|, |\hat{x}_l|, 1) \le 3. \qquad (4.5)$$

In accordance with the concept of the extended filter [6], the design of (2.3) will be based on the following approximate error dynamics:

$$\begin{aligned} e_{z,l+1} &= F_{c,l}(\hat{z}_l) e_{z,l} - N_l w_l \\ e_{\theta,l} &= L\, e_{z,l} \end{aligned} \qquad (4.6)$$

where $F_{c,l}(\hat{z}_l) = F_l(\hat{z}_l) - N_l H_l(\hat{z}_l)$. Then, one will need to show that, at least for small $\|e_{z,l}\|_2$, the exact error equation (4.2) meets the objective (3.2). To accomplish this, consider the following modified objective for the extended filter:

$$\|e_\theta\|^2_{2,[0,J]} + \varepsilon\, \|e_z\|^2_{2,[0,J]} < \|\gamma w\|^2_{2,[0,J]} + e^T_{z,0}\, R\, e_{z,0} \qquad (4.7)$$

where the positive ε is to be determined. By analogy to Lemma 4, if

$$\Delta V_l(e_{z,l}) < \gamma_l^2\, \|w_l\|_2^2 - \|e_{\theta,l}\|_2^2 - \varepsilon\, \|e_{z,l}\|_2^2 \qquad (4.8)$$

then the objective (4.7) is met. Define the Hamiltonian function

$$\mathcal{H}(w_l, N_l) := V_{l+1}(e_{z,l+1}) - \gamma_l^2 \|w_l\|_2^2 + \|e_{\theta,l}\|_2^2 + \varepsilon \|e_{z,l}\|_2^2 \tag{4.9}$$

which has a saddle point $(w_l^\star, N_l^\star)$. In the sequel, the derivation of saddle point $(w_l^\star, N_l^\star)$ assumes that the underlying p.d. function is quadratic as follows:

$$V_l(e_{z,l}) = e_{z,l}^T \bar{Q}_l e_{z,l} \tag{4.10}$$

where $\bar{Q}_0 = R$. Maximizing the resulting Hamiltonian with respect to w_l and, then, minimizing with respect to N_l leads—after suppressing the $\star$ notation—to

$$\begin{aligned} \bar{Q}_{l+1}^{-1} &= F_l(\hat{z}_l) Q_{\varepsilon,l}^{-1} F_l^T(\hat{z}_l) \\ &\quad - N_l \left[I + H_l^T(\hat{z}_l) Q_{\varepsilon,l}^{-1} H_l(\hat{z}_l) \right] N_l^T \end{aligned} \tag{4.11}$$

where

$$Q_{\varepsilon,l} = \bar{Q}_l - \gamma_l^{-2} L^T L - \varepsilon I \tag{4.12}$$

The following is the optimal gain of the extended filter:

$$N_l = F_l(\hat{z}_l) Q_{\varepsilon,l}^{-1} H_l^T(\hat{z}_l) \left[I + H_l^T(\hat{z}_l) Q_{\varepsilon,l}^{-1} H_l(\hat{z}_l) \right]^{-1} \tag{4.13}$$

In the infinite horizon case, the existence of a filter (2.3) so that (4.6) satisfies (4.7) is related to the existence of a bounded nonnegative solution for the Riccati equation (4.11); refer to p. 138 in [12]. Therefore, for all $l \in [0, \infty)$, one has $\|\bar{Q}_l\| \le c_1$, $\|N_l\| \le c_2$, and $\|F_{c,l}(\hat{z}_l)\| \le c_3$.

The next section shows that, under certain conditions, the filter gain (4.13) forces the trajectories of the error dynamics (4.2) to meet the objective (3.2).

V. PERFORMANCE ANALYSIS OF THE EXTENDED FILTER

At this time, consider the exact error equation (4.2) written as follows:

$$\begin{aligned} e_{z,l+1} &= f_{e,l}(\hat{z}_l, e_{z,l}, w_l) + g_{3,l}(\hat{z}_l, e_{z,l}) \\ e_{\theta,l} &= L\, e_{z,l} \end{aligned} \tag{5.1}$$

where $f_{e,l}(\hat{z}_l, e_{z,l}, w_l) = F_{c,l}(\hat{z}_l) e_{z,l} - N_l w_l$. The result of this section is given for the infinite horizon case but it also holds when $J < \infty$. The following proposition shows that the error equation above satisfies the filter objective in the neighborhood of the origin and, furthermore, quantifies the neighborhood in terms of the design parameter ε.

Proposition 6: Suppose there exists a bounded nonnegative solution $\bar{Q}_l$ of the Riccati equation (4.11) for $l \in [0, \infty)$. Then, inequality (3.2), associated with the error dynamics (5.1), holds as long as

$$\varepsilon > (2c_0 c_1 c_2 + 2c_1 c_3 \|e_{z,l}\|_2 + c_1 \|e_{z,l}\|_2^2) \mu_3 \tag{5.2}$$

where $\|w_l\|_2 \le c_0$, $\|\bar{Q}_l\| \le c_1$, $\|N_l\| \le c_2$, and $\|F_{c,l}(\hat{z}_l)\| \le c_3$.

Proof: Using the p.d. function (4.10), one has

$$\begin{aligned} V_{l+1}(e_{z,l+1}) &= e_{z,l+1}^T \bar{Q}_{l+1} e_{z,l+1} \\ &= [f_{e,l} + g_{3,l}]^T \bar{Q}_{l+1} [f_{e,l} + g_{3,l}] \\ &= f_{e,l}^T \bar{Q}_{l+1} f_{e,l} + 2 f_{e,l}^T \bar{Q}_{l+1} g_{3,l} \\ &\quad + g_{3,l}^T \bar{Q}_{l+1} g_{3,l} \end{aligned} \tag{5.3}$$

Since $\bar{Q}_l$ is a solution of the Riccati equation, the approximate dynamics (4.6) satisfy the inequality (4.8); thus,

$$f_{e,l}^T \bar{Q}_{l+1} f_{e,l} - e_{z,l}^T \bar{Q}_l e_{z,l} < \gamma_l^2 \|w_l\|_2^2 - \|e_{\theta,l}\|_2^2 - \varepsilon \|e_{z,l}\|_2^2 \tag{5.4}$$

In regard to rest of the terms on the right-hand-side of (5.3), one observes that

$$\begin{aligned} 2 f_{e,l}^T \bar{Q}_{l+1} g_{3,l} &\le 2\|e_{z,l}\|_2 \|F_{c,l}\| \|\bar{Q}_{l+1}\| \|g_{3,l}\|_2 \\ &\quad + 2\|w_l\|_2 \|N_l\| \|\bar{Q}_{l+1}\| \|g_{3,l}\|_2 \\ &\le 2c_3 c_1 \mu_3 \|e_{z,l}\|_2^3 + 2c_1 c_0 c_2 \mu_3 \|e_{z,l}\|_2^2 \end{aligned} \tag{5.5}$$

as well as

$$\begin{aligned} g_{3,l}^T \bar{Q}_{l+1} g_{3,l} &\le \|\bar{Q}_{l+1}\| \|g_{3,l}\|_2^2 \\ &\le c_1 \mu_3 \|e_{z,l}\|_2^4 \end{aligned} \tag{5.6}$$

From (5.4), (5.5) and (5.6), the rate of change of $V_l(e_{z,l})$ along a trajectory of the exact error equation (5.1) is bounded as follows:

$$\begin{aligned} &\Delta V_l(e_{z,l}) \le \gamma_l^2 \|w_l\|_2^2 - \|e_{\theta,l}\|_2^2 \\ &- [\varepsilon - (2c_1 c_3 \|e_{z,l}\|_2 - 2c_0 c_1 c_2 - c_1 \|e_{z,l}\|_2^2) \mu_3] \|e_{z,l}\|_2^2 \end{aligned} \tag{5.7}$$

Then, from Lemma 4, the hypothesis (5.2) proves the proposition. ■

VI. STABILITY OF THE EXACT ERROR EQUATION

In the infinite horizon case, the inequality (3.2) guarantees that the filter is input-output stable, but provides no information on whether the filter is asymptotically stable. The following proposition examines the asymptotic stability of the exact error equation (4.2), whose unforced version has as follows:

$$\begin{aligned} e_{z,l+1} &= F_{c,l}(\hat{z}_l) e_{z,l} + g_{3,l}(\hat{z}_l, e_{z,l}) \\ e_{\theta,l} &= L\, e_{z,l} \end{aligned} \tag{6.1}$$

As mentioned earlier, inequality (3.3) does not imply asymptotic stability, for (6.1) may not be zero-state observable from e_θ.

Proposition 7: Suppose the Riccati equation (4.11) has a bounded nonnegative solution $\bar{Q}_l$ for $l \in [0, \infty)$ and ε is such that

$$\varepsilon > (2c_1 c_3 \|e_{z,l}\|_2 + c_1 \|e_{z,l}\|_2^2) \mu_3 \tag{6.2}$$

where $\|\bar{Q}_l\| \le c_1$ and $\|F_{c,l}(\hat{z}_l)\| \le c_3$. Then, the exact error equation (6.1) is a.s.

Proof: Use of (4.10) as a candidate Lyapunov function leads to

$$\begin{aligned} V_{l+1}(e_{z,l+1}) &= [F_{c,l}e_{z,l} + g_3]^T \bar{Q}_{l+1} [F_{c,l}e_{z,l} + g_3] \\ &= e_{z,l}^T F_{c,l}(\hat{z}_l)^T \bar{Q}_{l+1} F_{c,l} e_{z,l} \\ &\quad + 2e_{z,l}^T F_{c,l}(\hat{z}_l)^T \bar{Q}_{l+1} g_3 + g_3^T \bar{Q}_{l+1} g_3 \end{aligned}$$

The matrix $\bar{Q}_l$ satisfies the Riccati equation (4.11) and, in turn, in inequality (4.8) holds for the approximate error dynamics. After setting $w \equiv 0$ in (4.8), one gets

$$\begin{aligned} e_{z,l}^T F_{c,l}(\hat{z}_l)^T \bar{Q}_{l+1} F_{c,l}(\hat{z}_l) e_{z,l} - e_{z,l}^T \bar{Q}_l e_{z,l} &< -\|e_{\theta,l}\|_2^2 \\ &\quad -\varepsilon \|e_{z,l}\|_2^2 \end{aligned} \tag{6.3}$$

Therefore, the rate of change of candidate Lyapunov function along a trajectory of the exact error equation is bounded as follows:

$$\begin{aligned} &\Delta V_l(e_{z,l}) \leq -\|e_{\theta,l}\|_2^2 \\ &-(\varepsilon - 2c_1c_3\mu_3\|e_{z,l}\|_2 - c_1\mu_3\|e_{z,l}\|_2^2) \|e_{z,l}\|_2^2 \end{aligned} \tag{6.4}$$

Then, from the hypothesis, (4.10) is a Lyapunov function. ■

The next corollary follows readily from Propositions 6 and 7.

Corollary 8: Suppose the modified objective (5.2) is used for the design of (2.3) as an extended filter. Then, at least locally, the original filter objective (3.2) will be met and the error dynamics (4.2) will be a.s.

VII. CONCLUSION

The design of the extended filter using the modified criterion (4.7) may result in a more conservative parameter estimator through a necessarily larger time-varying tolerance. In return, it guarantees, at least locally, that the filter is internally stable and its objective is met. In future research, the authors will investigate the possibility of inducing less conservatism by introducing a p.d. matrix, i.e., more than one design parameters in lieu of ε in the left-hand-side of (4.7).

VIII. ACKNOWLEDGMENTS

The authors gratefully acknowledge the contribution of the reviewers' comments.

Appendix

The output-of-interest matrix is the following:

$$L := \begin{bmatrix} 0 & I \end{bmatrix}$$

The definitions of the nonlinear mappings f and g are given below:

$$\begin{aligned} f(z) &:= \begin{bmatrix} A(\theta)\,x \\ \theta \end{bmatrix} \\ h(z) &:= C(\theta)\,x \end{aligned}$$

The respective Jacobian matrices are as follows:

$$\begin{aligned} F(z) = \frac{\partial}{\partial z} f(z) &= \begin{bmatrix} A(\theta) & \frac{\partial}{\partial \theta}[A(\theta)\,x] \\ 0 & I \end{bmatrix} \\ H(z) = \frac{\partial}{\partial z} h(z) &= \begin{bmatrix} C(\theta) & \frac{\partial}{\partial \theta}[C(\theta)\,x] \end{bmatrix} \end{aligned}$$

References

[1] M. J. Grimble and A. E. Sayed, "Solution of the $\mathcal{H}_\infty$ optimal filtering problem for discrete-time systems," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 38, no. 7, pp. 1092–1104, 1990.

[2] U. Shaked and Y. Theodor, "$\mathcal{H}_\infty$-Optimal estimation: A tutorial," in Proceedings of the IEEE Conference on Decision and Control, (Tuscon, AZ), pp. 2278–2286, Dec. 1992.

[3] U. Shaked and N. Berman, "H_∞ nonlinear filtering of discrete-time processes," IEEE Transactions on Signal Processing, vol. 43, no. 9, pp. 2205–2209, 1995.

[4] X. Shen and L. Deng, "Game theory approach to discrete H_∞ filter design," IEEE Transactions on Signal Processing, vol. 45, no. 4, pp. 1092–1095, 1997.

[5] Y. Theodor, U. Shaked, and N. Berman, "Time-domain H_∞ identification," IEEE Transactions on Automatic Control, vol. 41, no. 7, pp. 1019–1023, 1996.

[6] A. H. Jazwinski, Stochastic process and filtering theory, vol. 64 of Mathematics in Science and Engineering. New York: Academic Press, 1970.

[7] J. R. O'Brien and K. Kiriakidis, "Convergence and performance in $\mathcal{H}_\infty$-based parameter identification," in IFAC Workshop on Robust Control Design, (Milan, Italy), June 2003.

[8] A. J. van der Schaft, "L_2–gain analysis of nonlinear systems and nonlinear state feedback H_∞ control," IEEE Transactions on Automatic Control, vol. 37, pp. 770–784, June 1992.

[9] T. K. Moon and W. C. Stirling, Mathematical Methods and Algorithms for Signal Processing. Printice Hall, 2000.

[10] H. K. Khalil, Nonlinear Systems. Upper Saddle River, NJ: Prentice Hall, 1996.

[11] P. Lancaster and M. Tismenetsky, The theory of matrices. San Diego, CA: Academic Press, second ed., 1985.

[12] A. Ichikawa and H. Katayama, Linear Time Varying Systems and Sampled-data Systems, vol. 265 of Lecture Notes in Control and Information Sciences. London: Springer-Verlag, 2001.