A Convex Optimization Approach to Synthesizing Bounded Complexity ℓ^∞ Filters

Franco Blanchini and Mario Sznaier

***Abstract*—We consider the worst-case estimation problem in the presence of unknown but bounded noise. Contrary to stochastic approaches, the goal here is to confine the estimation error within a bounded set. Previous work dealing with the problem has shown that the complexity of estimators based upon the idea of constructing the state consistency set (e.g., the set of all states consistent with the *a priori* information and experimental data) cannot be bounded *a priori*, and can, in principle, continuously increase with time. To avoid this difficulty we propose a class of bounded complexity filters, based upon the idea of confining r—length error sequences (rather than states) to hyperrectangles. The main result of the technical note shows that this can be accomplished by using linear time invariant filters of order no larger than r. Further, synthesizing these filters reduces to a combination of convex optimization and line search.**

***Index Terms*—Linear time invariant (LTI), ℓ^∞ filtering, worst-case estimation.**

I. INTRODUCTION

In several estimation problems, the only information about the noise is a pointwise bound on the norm. Typical examples include quantization, sampling, measurement errors and, more in general, all cases in which a stochastic characterization of the noise is not available. In these circumstances the so called unknown-but-bounded approach, which aims at minimizing the worst case estimation error could be preferred to stochastic estimation methods [1].

Initial work in this area dates back to the early 70's [3], [13]. It was immediately apparent that characterizing the set of states compatible with measurements is in general hard. An ellipsoidal approximation was therefore proposed, which is, in general, quite conservative.

Subsequent work on worst estimation in the presence of ℓ^∞ bounded disturbances was studied in [9], [11], [18] (see also the survey [10]). The main result of these papers shows that pointwise optimal estimators can be obtained as the product of a subset of past measurements and a (time varying) gain. Both the gain and the set of relevant measurements result from solving a linear programming optimization problem. However, this optimization problem involves all past measurements. Thus, the complexity of these estimators grows with time.

The use of nonlinear recursive filters was proposed in [17], where the idea was to bound the set of possible states consistent with the output observations by a set whose center is propagated recursively and whose shape can be found by solving (at each instant) an optimization problem. Still, the complexity of the resulting observer is potentially high. A semi-recursive algorithm was proposed in [21]. In the case of known initial conditions, the optimal ℓ^∞ estimation problem was reduced to an ℓ^1 model matching problem [2], [5], [6] that can be solved (with arbitrary precision) by using the techniques in [5]. The case of unknown initial conditions was handled by using initially a non-recursive pointwise optimal estimator similar to that in [18], switching afterwards to the recursive ℓ^1 estimator. Since this estimator is based on solving a 2—block ℓ^1 model matching problem its complexity (and hence that of the overall estimator) cannot be bounded *a priori*.

Pointwise optimal estimators were obtained in [15], [16], by constructing a polyhedral set guaranteed to contain the states of the plant. However, the resulting algorithm is non-recursive and, again, the complexity of the estimator is not bounded *a priori* and potentially increases with time.

Motivated by the high complexity entailed in the approaches above, our goal is to synthesize *fixed* order recursive filters for systems subject to ℓ^∞ bounded disturbances, with guaranteed worst case estimation error. The main idea underlying the proposed approach is to, rather than attempting to confine the state of the system to a given set, to simply confine the estimation error to hyperrectangles. Intuitively, this amounts to willingly dropping information in return for obtaining bounded complexity filters. Our main results show that the problem of synthesizing bounded complexity filters that confine the error to the tightest possible hyperrectangle, for a set of suitable initial conditions, can be reduced to a combination linear programming/line search. For initial conditions outside this set, the estimation error converges, in finite time, to the design value. Finally, we briefly discuss an extension to the case of switching plants.

II. PRELIMINARIES

In the sequel, we denote by $\|y\|_\infty$ the ∞—norm of $y \in R^n$: $\|y\|_\infty \doteq \max_i |y|_i$. $\|M\|_1$ is the $\infty \to \infty$ induced norm of matrix $M \in R^{n\times m}$: $\|M\|_1 \doteq max_i \sum_j |M_{ij}|$. We denote by ℓ_n^1 and ℓ_n^∞ the extended Banach spaces of vector valued real sequences, $\{y\}_0^\infty \in R^n$, respectively equipped with the norms $\|y\|_{\ell^1} \doteq \sum_{i=0}^\infty \|y_i\|_\infty$, and $\|y\|_{\ell^\infty} \doteq \sup_i \|y_i\|_\infty$. $\mathcal{B}\ell^1$ and $\mathcal{B}\ell^\infty$ are the unit balls in ℓ^1, ℓ^∞. $\|G\|_{\ell^\infty\to\ell^\infty}$ is the ℓ^∞ to ℓ^∞ induced norm of the operator $G : \ell^\infty \to \ell^\infty$, e.g., $\|G\|_{\ell^\infty\to\ell^\infty} \doteq \sup_{y\neq 0} \|Gy\|_{\ell^\infty}/\|y\|_{\ell^\infty}$. $Y(\lambda)$ denotes the λ—transform[1] of a sequence $\{y_k\}_0^\infty$:

$$Y(\lambda) \doteq \sum_{i=0}^{\infty} y_k \lambda^k.$$

Given a scalar ARMA model of the form

$$y(k) = -\sum_{i=1}^{n} a_i y(k-i) + \sum_{i=0}^{m} b_i v(k-i), \quad n \geq m \tag{1}$$

its λ—transform representation is

$$y(\lambda) = \frac{\sum_{i=0}^{n} b_i\lambda^i}{\sum_{i=0}^{m} a_i\lambda^i} v(\lambda). \tag{2}$$

The notion of equalized performance, introduced in [4] (see also [12]) will play a key role in obtaining bounded complexity filters.

[1]this corresponds to setting $\lambda = 1/z$ in the usual z-transform representation.

Manuscript received June 09, 2010; revised March 18, 2011; accepted June 13, 2011. Date of publication July 25, 2011; date of current version December 29, 2011. This work was supported in part by National Science Foundation (NSF) grants IIS-0713003, ECCS-0901433, AFOSR grant FA9550-09-0253, and the Alert DHS Center of Excellence under Award 2008-ST-061-ED0001. Recommended by Associate Editor T. Zhou.

F. Blanchini is with the Department of Mathematics and Computer Science, University of Udine, Udine 33100, Italy (e-mail: blanchini@uniud.it).

M. Sznaier is with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115 USA (e-mail: msznaier@coe.neu.edu).

Color versions of one or more of the figures in this technical note are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2011.2162893

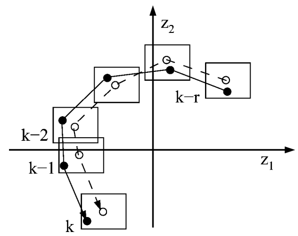

Fig. 1. Equalized filtering idea: black disks and white circles denote the true and estimated output trajectory, respectively.

Definition 2.1: Consider a linear time invariant (LTI) plant described by a model of the form (1). Given $r \geq n$, the plant is said to achieve an equalized r—performance level μ if, whenever the input and output sequences $\{v\}, \{y\}$ satisfy $|v(t)| \leq 1$ and $|y(t)| \leq \mu$ for all $t = k, k-1, \ldots, k-r+1$, then $|y(k+1)| \leq \mu$ (thus $|y(k+i)| \leq \mu$, for $i > 0$). In particular the case $r = n$ will be simply referred to as equalized performance.

As shown in [4], only superstable plants (in the sense of [12]) achieve (finite) equalized performance. However, any stable plant achieves equalized r—performance for some large enough r. Further, if a SISO plant achieves r—performance μ for some finite r, then it achieves r'—performance μ for any $r' > r$.

Next, we recall, for ease of reference, some properties concerning the relationship between equalized performance and the ℓ^∞ induced norm.

Lemma 2.1 ([4]): Given a stable, LTI SISO plant $y(\lambda) = G(\lambda)v(\lambda)$, as in (2) with finite r-equalized performance $\mu(r_o)$ for some $r_o \geq n$, the following holds:

1) $\|G\|_{\ell^\infty \to \ell^\infty} \leq \mu(r_o)$, with the equality holding for finite impulse response (FIR) plants.
2) $\mu(r) \downarrow \|G\|_{\ell^\infty \to \ell^\infty}$, as $r \longrightarrow \infty$.

A. Why Equalized Filtering?

Recursive set-valued observers [14], [17] are based upon the idea of propagating a set known to contain the (unknown) state of the plant, which may have high complexity. To avoid this difficulty, in this technical note, rather than attempting to confine the state, we will work directly with the estimation error and attempt to design a filter such that, if at some time instant t_o the past r values of the error are "captured" in an r-hyperrectangle, then this property will hold for all $t > t_o$ and all $\|v\|_{\ell^\infty} \leq 1, \|w\|_{\ell^\infty} \leq 1$ (see Fig. 1). Further, we are interested in synthesizing the tightest hyperrectangle satisfying this property. The main result of this technical note shows that this can be accomplished by reducing the problem to an equalized performance one. Moreover, contrary to the controller design case considered in [4], in the filtering case the results are easily extended to MIMO systems by simply considering a collection of component-wise filters.

III. Equalized Performance Filtering

Consider an LTI plant subject to ℓ^∞ bounded disturbances, with state space realization

$$\begin{aligned} x_{k+1} &= Ax_k + B_v v_k \\ z_k &= Hx_k \\ y_k &= C_y x_k + Dw_k \end{aligned} \tag{3}$$

or with λ—transform representation

$$z(\lambda) = \frac{M(\lambda)}{d(\lambda)} v(\lambda) \tag{4}$$

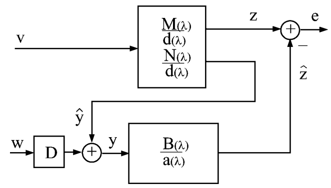

Fig. 2. Filtering scheme.

$$y(\lambda) = \frac{N(\lambda)}{d(\lambda)} v(\lambda) + Dw(\lambda) \tag{5}$$

where $z \in \mathbb{R}^s$, $y \in \mathbb{R}^q$, $v \in \mathbb{R}^p$ and $w \in \mathbb{R}^q$ represent the output to be estimated, the measurements available to the filter, process and measurement noise, respectively, and where $d(\lambda) = det(I - \lambda A)$. Note that we have assumed that the plant is strictly proper with respect to the input v. This assumption is made for notational simplicity and can be removed at the price of a more involved notation in the subsequent development. For the time being, we will also assume that z is a scalar $(s = 1)$, but this assumption will be removed later.

Our goal is to design a filter of the form

$$\hat{z}(\lambda) = \frac{B(\lambda)}{a(\lambda)} y(\lambda) \tag{6}$$

such that the estimation error

$$e(\lambda) = z(\lambda) - \hat{z}(\lambda) \tag{7}$$

is confined to an hyperrectangle. The complete filtering scheme is illustrated in Fig. 2.

In the sequel, we will limit our attention to filters that belong to the class of generalized Luenberger observers, defined as follows:

Definition 3.1 ([7]): A system of the form

$$\xi_{k+1} = P\xi_k + Ly_k \tag{8}$$

$$\hat{x}_k = Q\xi_k + Ry_k \tag{9}$$

$$\hat{z}_k = H\hat{x}_k \tag{10}$$

is a generalized Luenberger (state) observer for system (3) if P is a stable matrix and $\hat{x}_k - x_k \longrightarrow 0$ as $k \longrightarrow \infty$, when $w(k) \equiv 0$ and $v(k) \equiv 0$.

Next we recall a characterization of the class of the generalized observers.

Lemma 3.1: The system (8)–(10) is a generalized observer for (3) if and only if P is stable and there exists a full column rank matrix T such that

$$TA - LC_y = PT, \tag{11}$$

$$QT + RC_y = I. \tag{12}$$

Proof: See [7], [19]. ■

Remark 3.1: The standard n—order Luenberger observers corresponds to the choice $T = I$ and $R = 0$. Kalman filters fall in this category. Selecting a "tall" T matrix leads to a higher order observer, with additional degrees of freedom that can be used to optimize performance.

Next we show that restricting the filter to be a generalized observer imposes a constraint on its structure.

Lemma 3.2: If the filter (6) is a generalized state observer for system (3), then the polynomial matrices $M(\lambda)$ (of dimension $1 \times p$), $N(\lambda)$ (of dimension $q \times p$), $B(\lambda)$ (of dimension $1 \times q$) and the polynomials $a(\lambda)$ and $d(\lambda)$ satisfy the following condition:

$$M(\lambda)a(\lambda) - B(\lambda)N(\lambda) = C(\lambda)d(\lambda) \tag{13}$$

for some polynomial matrix $C(\lambda)$.

Proof: From (8)–(12) it follows that:

$$\begin{aligned}[Tx-\xi]_{k+1} &= TAx_k - P\xi_k - L(C_y x_k + Dw_k)\\ &\quad + TB_v v_k\\ &= P[Tx_k - \xi_k] + TB_v v_k - LDw_k\\ x_k - \hat{x}_k &= x_k - Q\xi_k - RC_y x_k - RDw_k\\ &= Q[Tx_k - \xi_k] - RDw_k. \end{aligned} \tag{14}$$

Consider now the change of variables $\eta = x$ and $\theta = [Tx - \xi]$. In term of these variables the state space representation of the combined plant-filter system is given by

$$\begin{bmatrix}\eta_{k+1}\\ \theta_{k+1}\end{bmatrix} = \begin{bmatrix} A & 0\\ 0 & P\end{bmatrix}\begin{bmatrix}\eta_k\\ \theta_k\end{bmatrix} + \begin{bmatrix} B_v & 0\\ TB_v & -LD\end{bmatrix}\begin{bmatrix} v_k\\ w_k\end{bmatrix}$$
$$e_k = [0 \quad HQ]\begin{bmatrix}\eta_k\\ \theta_k\end{bmatrix} + [0 \quad -HRD]\begin{bmatrix} v_k\\ w_k\end{bmatrix}.$$

Thus η is unobservable from e. Hence, the modes of A are canceled in the transfer function from v to e, $T_{e,v}$. From (4)–(7) it follows that:

$$\begin{aligned} e(\lambda) &= \left[\frac{M(\lambda)}{d(\lambda)} - \frac{B(\lambda)}{a(\lambda)}\frac{N(\lambda)}{d(\lambda)}\right] v(\lambda)\\ &\quad + \left[\frac{B(\lambda)}{a(\lambda)}\right] Dw(\lambda)\\ &= \left[\frac{M(\lambda)a(\lambda) - B(\lambda)N(\lambda)}{a(\lambda)d(\lambda)}\right] v(\lambda)\\ &\quad + \left[\frac{B(\lambda)}{a(\lambda)}\right] Dw(\lambda). \end{aligned}$$

Since $d(\lambda) = det(I - A\lambda)$, the cancelation of the modes of A in $T_{e,v}$ implies that $M(\lambda)a(\lambda) - B(\lambda)N(\lambda)$ has $d(\lambda)$ as a factor, precisely what (13) states. ■

In the sequel we will limit our attention to polynomial matrices satisfying (13) for some $C(\lambda)$, so that the estimation error is governed by

$$e(\lambda) = \frac{C(\lambda)}{a(\lambda)} v(\lambda) + \frac{B(\lambda)}{a(\lambda)} Dw(\lambda). \tag{15}$$

We are now in the position to formally state the equalized-performance filtering problem.

Problem 3.1: Given an integer $r \geq n$ and $\mu > 0$, find a filter of the form (6) of order r satisfying the constraint (13) and such that $a(\lambda)$ is stable (i.e., all its poles are outside the unit circle) and

$$\begin{aligned}|e_{k-j}| \leq \mu, j =, 1, 2, \ldots, r \Rightarrow |e_t|\\ \leq \mu \text{ for all } t \geq k \text{ and all sequences } v, w \in \mathcal{B}\ell^\infty. \end{aligned} \tag{16}$$

Note that the problem above does not explicitly make any assumptions on the initial conditions of the plant x_0. As we will show later, if the plant achieves an equalized performance level $\mu < \infty$, then there exist a set of initial conditions $\mathcal{X}_o(\mu)$ such that if $x_o \in \mathcal{X}_o(\mu)$ then $|e_k| \leq \mu$ for all k. For initial conditions outside this set, the property (16) will be satisfied after a finite number of steps.

Theorem 3.1: An r^{th} order filter of the form (6), subject to (13), with

$$\begin{aligned} a(\lambda) &= 1 + a_1\lambda + \ldots + a_r\lambda^r\\ B(\lambda) &= B_0 + B_1\lambda + \ldots + B_r\lambda^r\\ C(\lambda) &= C_0 + C_1\lambda + \ldots + C_r\lambda^r \end{aligned}$$

solves Problem 3.1 above if and only if

$$\begin{aligned}\mu\|[a_1\ a_2 \ldots a_r]\|_1 + \|[C_0\ C_1 \ldots C_r]\|_1\\ + \|B_0 D\ \ldots B_r D\|_1 \leq \mu. \end{aligned}\tag{17}$$

Proof: From (15) the ARMA model relating the signals e, v, w is given by

$$e_k = -\sum_{i=1}^{r} a_i e_{k-i} + \sum_{i=0}^{r} C_i v_{k-i} + \sum_{i=0}^{r} B_i D w_{k-i}. \tag{18}$$

Thus, if $|e_{k-i}| \leq \mu$ and $i = 1, 2, \ldots, r$, with $v, w \in \mathcal{B}\ell^\infty$ then

$$\begin{aligned} |e_k| &= |-\sum_{i=1}^{r} a_i e_{k-i} + \sum_{i=0}^{r} C_i v_{k-i} + \sum_{i=0}^{r} B_i D w_{k-i}|\\ &\leq \sum_{i=1}^{r} |a_i||e_{k-i}| + \left\| [C_0\ C_1 \ldots C_r]\begin{bmatrix} v_{k-1}\\ \vdots\\ v_{k-r}\end{bmatrix}\right\|_\infty\\ &\quad + \left\| [B_0 D\ B_1 D \ldots B_r D]\begin{bmatrix} w_{k-1}\\ \vdots\\ w_{k-r}\end{bmatrix}\right\|_\infty\\ &\leq \sum_{i=1}^{r} |a_i|\mu + \|[C_0\ C_1 \ldots C_r]\|_1\\ &\quad + \|[B_0 D\ B_1 D \ldots B_r]D\|_1 \leq \mu. \end{aligned}$$

Therefore the condition is sufficient. To prove necessity, start by rewriting (18) as

$$|e(k)| = |[\mu a_1 \ldots \mu a_r\ C_0 \ldots C_r\ B_0 D \ldots B_r D]\, x| \doteq |\Xi x|$$

where $x \doteq [e_{k-1}/\mu \ldots e_{k-r}/\mu \quad v_k \ldots v_{k-r} w_k \ldots \quad w_{k-r}]^T$. From the hypothesis it follows that x is an arbitrary element of $\mathcal{B}\ell^\infty$. Hence

$$\sup_{\|x\|_\infty \leq 1} |e_k| \leq \mu \iff \|\Xi\|_1 \leq \mu$$

or, equivalently

$$\begin{aligned}|e(k)| = |\mu a_1| + |\mu a_2| + \ldots |\mu a_r| + \|[C_0\ C_1 \ldots C_r]\|_1\\ + \|[B_0 D\ B_1 D \ldots B_r D]\|_1 \leq \mu\end{aligned}$$

which proves necessity. To conclude the proof, we show that condition (17) implies stability of the filter. This follows immediately from the fact that it implies:

$$\|[a_1\ a_2 \ldots a_r]\|_1 = \sum_{i=1}^{r} |a_i| = \rho < 1.$$

■

We will refer to r—equalized filters, namely satisfying (17) as equalized filters, with performance μ. To synthesize one of such filters write

$$\begin{aligned} d(\lambda) &= 1 + d_1\lambda + \ldots + d_l\lambda^l\\ M(\lambda) &= M_0 + M_1\lambda + \ldots + M_l\lambda^l\\ N(\lambda) &= N_0 + N_1\lambda + \ldots + N_l\lambda^l \end{aligned}$$

and rewrite (13) as the following linear constraints in the variables a_k, B_k and C_k:

$$\begin{aligned} &M_0a_0 - B_0N_0 = C_0d_0 \\ &M_1a_0 + M_0a_1 - B_1N_0 - B_0N_1 = C_1d_0 + C_0d_1 \\ &M_2a_0 + M_1a_1 + M_0a_2 - B_2N_0 - B_2N_1 - B_0N_2 \\ &\quad = C_2d_0 + C_1d_1 + C_0d_2 \\ &\vdots \\ &\quad M_la_r - B_rN_l = C_rd_l. \end{aligned}$$

Since for a fixed μ, (17) is convex in a_k, B_k and C_k, it follows that establishing feasibility of (13)–(17) reduces to a convex feasibility problem, involving the linear constraints above. The optimal filter (and its associated optimal filtering error μ_{opt}) can be found via bisection, by increasing/decreasing μ in case of infeasibility/feasibility.

Remark 3.2: In fact, the feasibility problem associated with the linear constraints and (17) can be reduced to a linear programming problem. Therefore synthesizing filters of large order to improve performance is not an issue. The order can be taken large compatibly with the technology available for implementation.

To address the multi-output case, we begin by extending the definition of equalized filtering performance. Given $\mu \doteq [\,\mu_1 \dots \mu_n\,]$, let $\mathcal{B}\ell^\infty(\mu)$ be the scaled unit ball in ℓ_n^∞: $\mathcal{B}\ell^\infty(\mu) \doteq \{e \in \ell_n^\infty : e_i/\mu_i \in \mathcal{B}\ell^\infty\}$.

Definition 3.2: The filter (6) with error $\hat{z} - z = e \in R^s$ is said to achieve a vector equalized performance level $\mu \doteq [\,\mu_1, \mu_2, \dots \mu_s\,]$ if it is stable and

$$e_{k-j} \in \mathcal{B}\ell^\infty(\mu), j = 1, \dots, r \Rightarrow e_k \in \mathcal{B}\ell^\infty(\mu), \quad \text{for all sequences } v, w \in \mathcal{B}\ell^\infty. \tag{19}$$

Next, we extend our previous results to the MIMO case.

Theorem 3.2: A filter $F: y \in \ell_n^\infty \to \hat{z} \in \ell_s^\infty$ achieves a vector equalized performance level μ iff each component $F_h: y \in \ell_n^\infty \to \hat{z}^h \in \ell^\infty$ achieves scalar equalized performance (in the sense of (16)) μ_h, where $\hat{z}^h$ denotes the h^{th} component of $\hat{z}$.

Proof: Clearly, if each component F_h achieves an equalized performance level μ_h, the overall filter F obtained by stacking each component satisfies the conditions in Definition 3.2. Conversely, assume that the filter F satisfies (19). Note that the multiple-output version of the filter (18) can be written in terms of its h component as follows:

$$e_k^h = -\sum_{i=1}^{r} a_i^h e_{k-i}^h + \sum_{i=0}^{r} C_i^h v_{k-i} + \sum_{i=0}^{r} B_i^h D w_{k-i} \tag{20}$$

and that the error terms e_{k-i}^h, $i = 1, 2, \dots, r$ can be initialized independently in each "partial filter". Assume now that for a given i the corresponding mapping F_h does not satisfy (16). Then, initializing all the other variables e^j, $j \neq h$, to $e_{k-i}^j = 0$, $i = 1, 2, \dots, r$ leads to violation of (19). ■

IV. Filter Initialization, Optimality and Robustness

The main result of this section shows that, given an initial set of r measurements, $\mathbf{y} \doteq [\,y_o, y_1, \dots, y_{r-1}\,]$ there exists a finite performance level μ and a filter initial condition ξ_o such that the estimation error satisfies $e_k \in \mathcal{B}\ell^\infty(\mu)$ for all k. Let $\mathcal{K}_o, T_y$ and T_z^j denote the r^{th} order Kalman observability matrix of the system (3) and the Toeplitz operators mapping v to y and the j^{th} component of z, respectively, e.g.

$$\mathcal{K}_o = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{bmatrix} T_y = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ CB & 0 & 0 & \dots & 0 \\ CAB & CB & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ CA^{r-1}B & CA^{r-2}B & \dots & CB & 0 \end{bmatrix}$$

$$T_z^j \doteq \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ H(j,:)B & 0 & 0 & \dots & 0 \\ H(j,:)AB & H(j,:)B & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ H(j,:)A^{r-1}B & H(j,:)A^{r-2}B & \dots & H(j,:)B & 0 \end{bmatrix}$$

where $M(i,:)$ denotes the ith row of the matrix M, then the filter can be initialized (after r measurements) as follows.

1) For $k = 0, \dots, r-1$ and $j = 1, \dots, s$ compute

$$\begin{aligned} z_k^{j,+} &\doteq \max_{x,v,w} H(j,:)A^{k-1}\mathbf{x} + T_z^j(k,:)\mathbf{v} \\ z_k^{j,-} &\doteq \min_{x,v,w} H(j,:)A^{k-1}\mathbf{x} + T_z^r(k,:)\mathbf{v} \\ &\text{subject to: } \mathbf{x} \in \tilde{\mathcal{X}}_o \\ &\mathbf{y} = T_y\mathbf{v} + \mathcal{D}\mathbf{w},\ \mathbf{w}, \mathbf{v} \in \mathcal{B}\ell_\infty \end{aligned} \tag{21}$$

where $\mathcal{D} \doteq \mathrm{diag}\{D^T\}$ and where $\tilde{\mathcal{X}}_o$ is a set known to contain the initial condition (if no information is available then $\tilde{\mathcal{X}}_o = R^n$).

2) Define

$$\begin{aligned} z_k^{j,c} &\doteq \frac{z_k^{j,+} + z_k^{j,-}}{2}, \\ \mu_k^j &\doteq \frac{1}{2}|z_k^{j,+} - z_k^{j,-}| \end{aligned} \tag{22}$$

3) Let $\mu^{init,j} = \max_{0 \le t \le r-1} \{\mu_t^j\}$ and choose a filter initial condition such that the first r filter estimates are $\hat{z}_t = z_t^{j,c},\ t = 0, \dots, r-1$. This is always feasible, since the order of the filter is precisely r.

Note that if $\mathcal{X}_o$ is convex, then the optimization problem (21) is convex. Further, if $\mathcal{X}_o$ is a polytope, this problems reduces to LP. Thus z^+, z^-, and z^c above can be found efficiently[2].

Since $\|\,[\,a_1 \dots a_r\,]\,\|_1 < 1$ by construction, it turns out that, if (17) holds for some $\tilde{\mu}$, then it also holds for all $\mu \geq \tilde{\mu}$, and so does (16). It follows that if $\mu^{init,j} \leq \mu_{opt}^j$ the optimal equalized performance level in (17), then the filter (6), with the initialization above, achieves optimal equalized performance level μ_{opt}^j for all $t \geq r$. On the other hand, as we show next, if $\mu^{init,j} > \mu_{opt}^j$, then the worst case ℓ^∞ estimation error is bounded above by $\mu^{init,j}$ and converges, in a finite number of steps, to μ_{opt}^j.

Theorem 4.1: Consider a filter of the form (6), with the initialization above. Then, for all $t, |e_t| \leq \mu^{init}$. Moreover, given $\mu \geq \mu_{opt}$ satisfying (17), for any plant and filter initial condition pairs $\{x_o, \xi_o\}$ there exists a finite time $T(x_o, \xi_o, \mu)$ such that for all $t > T, |e_t| \leq \mu$.

Proof: Consider the "Lyapunov like" function

$$\psi_k \doteq \max_{i=1,\dots,r} |e_{k-i}|.$$

[2]The estimate z_c can be thought off as a smoothing problem equivalent of the central estimator introduced in [14] or, equivalently, the pointwise optimal estimators proposed in [18].

From (18) we have that:

$$\begin{aligned}
|e_k| &\le \sum_{i=1}^{r} |a_i|\psi_k + \|[C_0 \dots C_r]\|_1 + \|[B_0 \dots B_r]D\|_1 \\
&= \sum_{i=1}^{r} |a_i|\mu + \|[C_0 \dots C_r]\|_1 + \|[B_0 \dots B_r]D\|_1 \\
&\quad + \sum_{i=1}^{r} |a_i|(\psi_k - \mu) \\
&< \mu + (\psi_k - \mu)\sum_{i=1}^{r} |a_i| = \mu + (\psi_k - \mu)\rho
\end{aligned}$$

with $\rho < 1$. Thus

$$|e_k| < \rho\psi_k + (1-\rho)\mu \le \max\{\mu, \psi_k\}.$$

Hence the sequence ψ_k is non increasing as long as $\psi_k > \mu$. Since by construction $|e_k| \le \mu^{init}, 0 \le t \le r-1$, it follows that $|e_t| \le \mu^{init}$ for all t. Further, due to the strict inequality, the subsequence ψ_{kr} (r is the parameter in Definition 2.1) is strictly decreasing (since each value generated after $|e_k|$ is strictly smaller than ψ_k). Indeed

$$\psi_{(k+1)r} < \max\{\rho\psi_{kr} + (1-\rho)\mu, \mu\}.$$

This equation implies that the subsequence ψ_{kr} and (thus ψ_k) converges to μ from above, and it can be easily seen that this convergence occurs in finite time, e.g., $\psi_t = \mu$ for t $>$ some T. The fact that $|e_t| \le \mu,\ t > T$ follows immediately follows from the definition of ψ_k. ■

The described initialization procedure is similar to that proposed in [21], with the difference that the initialization horizon is not problem-dependent.

Since the proposed filter is equalized-optimal among *all linear filters*, a natural question is whether it is optimal with respect to the class of worst-case filters. It can be shown, using concepts from information based-complexity [8], [20], that at least under special conditions, μ_{opt} is indeed the radius of information (see [8] for a definition) and thus the lowest possible worst-case error attainable by any filter. This issue is not investigated further here due to space constraints.

Next, we briefly consider the case of a system that switches among m plants with state space realizations given by: $A^{(k)}, B_v^{(k)}, H^{(k)}, C_y^{(k)}, D^{(k)}$. In this situation it is plausible to associate with each plant an equalized filter $\{B^{(k)}(\lambda), a^{(k)}(\lambda)\}$ and switch these filters using the same switching rule as the plant.

Proposition 4.1: A family of equalized switching filters of order r implemented as

$$\hat{z}(k) = -\sum_{i=1}^{r} a_i^{(k)}\hat{z}(k-i) + \sum_{i=0}^{m} b_i^{(k)} y(k-i);\ n \ge m \qquad (23)$$

is switching stable.

Proof: We have seen that the equalized filters of order r satisfies $\sum_{i=1}^{r}|a_i^{(k)}| = \rho_k < 1$ (then is superstable [12]). Therefore, for $y = 0$ we have

$$\begin{aligned}
|\hat{z}(k)| &\le \sum_{i=1}^{r} |a_i^{(k)}||\hat{z}(k-i)| \le \max_k\ \rho_k \max\{\hat{z}(k-i)|\} \\
&< \rho \max_{1\le i\le r}\{\hat{z}(k-i)|\}
\end{aligned}$$

where $\rho = \max\rho_k$ hence stability. ■

Note that although the result above guarantees stability of the switched filter, at the present time no bounds are available concerning its worst-case performance under arbitrary switching sequences. This issue is currently under investigation.

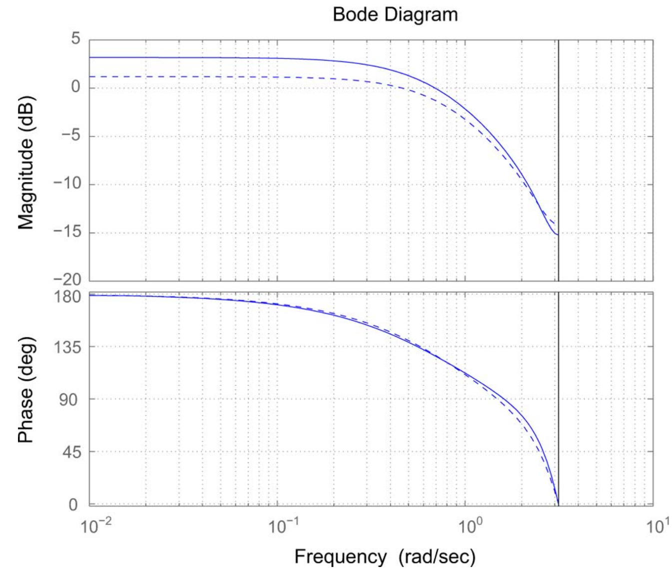

Fig. 3. The equalized filter (plain) versus the $\mathcal{H}_2$ filter (dashed) frequency response.

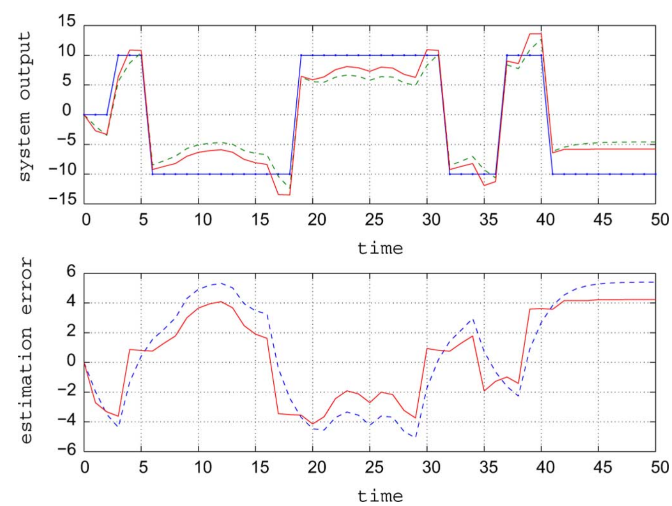

Fig. 4. Equalized versus $\mathcal{H}_2$ filters: Top figure: the output (thick), the $\mathcal{H}_2$ filter output (dashed), the equalized output (plain); Bottom figure: the $\mathcal{H}_2$ filter error (dashed), the equalized filter error (plain).

V. Illustrative Examples

Example 1: Consider the following second order plant:

$$\frac{M(\lambda)}{d(\lambda)} = \frac{\lambda^2}{1}\ \frac{N(\lambda)}{d(\lambda)} = \frac{(1-0.5\lambda)(1-2\lambda)}{1}$$

and assume $D = 1$ and that the process and measurement noise satisfy $|v_k| \le \gamma$ and $|w_k| \le 1$, respectively. For $0 < \gamma < 2$ the optimal equalized estimate is $\hat{z} = 0$, (e.g., zero filter), with $\mu_{opt} = \gamma$.

For $\gamma > 2$ seems to yield, independently of γ, the following filter:

$$\frac{a(\lambda)}{b(\lambda)} = \frac{-0.2463 - 0.6158\lambda - 0.2933\lambda^2 - 0.1173\lambda^3}{1.0000 - 0.1173\lambda^3}$$

with poles at 0.4895 and $-0.2448 \pm j0.4239$. The corresponding equalized cost is given by the following piecewise affine function of γ:

$$\mu_{opt}(\gamma) = \begin{cases} \gamma & \text{for } 0 < \gamma < 2 \\ 2 + \kappa(\gamma - 2) & \text{for } 2 < \gamma \end{cases}$$

with $\kappa \approx 0.28$.

We see that the optimal filter coefficients are not continuous with respect to the noise bounds and the system parameters. However, the optimal value is continuous. Moreover, if switching to a new filter due to parameter changes is necessary, robustness is assured by Proposition 4.1.

For comparison we considered the $\mathcal{H}_2$—optimal filter whose state-space realization is

$$A_{h_2} = \begin{bmatrix} 0.3700 & 0.0750 \\ -0.2430 & 0.6076 \end{bmatrix} \; B_{h_2} = \begin{bmatrix} -0.3700 \\ 0.2430 \end{bmatrix}$$
$$C_{h_2} = [1.1777 \quad -0.4443] \; D_{h_2} = [-0.1777].$$

We report in Fig. 3 the frequency response of the equalized and the $\mathcal{H}_2$—optimal filters and in Fig. 4 the simulations in the presence of random piecewise-constant noise. We see that the equalized filter has a slightly sharper cut-off effect. The noise is randomly generated by taking $v = \pm\gamma$ and $w = \pm\beta$ and randomly changing sign with probability $\pi = 1/10$ at each instant. We took $\gamma = 10$. It is apparent that the error produced by the equalized is always smaller than the $\mathcal{H}_2$ filter. Several experiments with randomly generated sequences show that for this example we have roughly a 30% improvement on the worst case error.

Example 2: Next, we consider the case of a plant with poles on the stability boundary[3]

$$\frac{M(\lambda)}{d(\lambda)} = \frac{\lambda^2}{1-\lambda^2} \; \frac{N(\lambda)}{d(\lambda)} = \frac{(1-0.5\lambda)(1-2\lambda)}{1-\lambda^2}.$$

In this case, the optimal equalized filter corresponding to $\beta = 1, \gamma = 8$ and $r = 3$ is given by

$$\frac{b(\lambda)}{a(\lambda)} = \frac{-0.3268 - 0.8171\lambda - 0.3891\lambda^2 - 0.1556\lambda^3}{1 - 0.1556\lambda^3}$$

and achieves an equalized performance level $\mu_{opt} = 5.1$

An intriguing fact borne out of these examples is that while in the context of control design the optimal equalized closed loop was almost always "near dead-beat" (e.g., "almost zero" closed-loop poles) the optimal equalized filter does not exhibit this feature.

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[2] N. E. Barabanov and O. N. Granichin, "Optimal controller for a linear plant with bounded noise," *Autom. Remote Control*, vol. 45, no. 5, pp. 578–584, 1984.

[3] D. P. Bertsekas and I. B. Rhodes, "Recursive state estimation for a set-membership description of uncertainty," *IEEE Trans. Autom. Control*, vol. AC-16, no. 2, pp. 117–128, Apr. 1971.

[4] F. Blanchini and M. Sznaier, "A convex optimization approach to fixed-order controller design for disturbance rejection for SISO systems," *IEEE Transs. Autom. Control*, vol. AC-45, no. 4, pp. 784–789, Apr. 2000.

[5] M. A. Dahleh and I. D. Diaz-Bobillo, *Control of Uncertain Systems: A Linear Programming Approach*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[6] M. A. Dahleh and J. B. Pearson, "l^1—optimal feedback controllers for MIMO discrete—time system," *IEEE Trans. Autom. Control*, vol. AC-32, no. 4, pp. 314–322, Apr. 1987.

[7] D. Luenberger, "An introduction to observers," *IEEE Trans. Autom. Control*, vol. AC-16, no. 6, pp. 596–602, 1971.

[8] C. A. Micchelli and T. J. Rivlin, *A Survey of Optimal Recovey, Optimal Recovery in Approximation Theory*, C. A Micchelli and T. J. Rivlin, Eds. New York: Plenum, 1977, pp. 1–54.

[9] M. Milanese and R. Tempo, "Optimal algorithms theory for robust estimation and prediction," *IEEE Trans. Autom. Control*, vol. AC-30, no. 8, pp. 730–738, Aug. 1985.

[10] M. Milanese and A. Vicino, "Optimal estimation theory for dynamic systems with set membership uncertainty," *An overview, Automatica*, vol. 27, pp. 997–1009, 1991.

[11] M. Milanese and G. Belforte, "Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: Linear families of models and estimators," *IEEE Trans. Autom. Control*, vol. AC-27, no. 2, pp. 408–414, Apr. 1982.

[12] B. T. Polyak and M. E. Halpern, "The use of a new optimization criterion for discrete-time feedback controller design," in *Proc. 38th IEEE CDC*, Tucson, AZ, Dec. 1999, pp. 894–899.

[13] F. C. Schweppe, "Recursive state estimation: Unknown but bounded errors and system inputs," *IEEE Trans. Autom. Control*, vol. AC-13, no. 1, pp. 22–28, Feb. 1968.

[14] J. S. Shamma and E. Tu, "Optimality of set-valued observers for linear systems," in *Proc. 34th IEEE CDC*, New Orleans, LA, Dec. 1995, pp. 2081–2086.

[15] J. S. Shamma and K. Y. Tu, "Approximate set-valued observers for nonlinear systems," *IEEE Trans. Autom. Control*, vol. 42, no. 5, pp. 648–658, May 1997.

[16] J. Shamma and K. Y. Tu, "Set-valued observer and optimal disturbance rejection," *IEEE Trans. Autom. Control*, vol. 44, no. 2, pp. 253–164, Feb. 1999.

[17] A. A. Stoorvogel, "ℓ^1 state estimation for linear systems using nonlinear observers," in *Proc. 35th IEEE CDC*, Kobe, Japan, Dec. 1996, pp. 2407–2411.

[18] R. Tempo, "Robust estimation and filtering in the presence of bounded noise," *IEEE Trans. Autom Control*, vol. AC-33, no. 9, pp. 864–867, Sep. 1988.

[19] C. Tsui, "A new design approach to unknown input observers," *IEEE Trans. Autom. Control*, vol. 41, no. 3, pp. 464–468, Mar. 1996.

[20] J. B. Traub, H. Wozniakowski, and G. W. Wasilkowski, *Information Based Complexity*. New York: Academic Press, 1988.

[21] P. Voulgaris, "On optimal ℓ^∞ to ℓ^∞ filtering," *Automatica*, vol. 31, pp. 489–495, 1995.

[3]Note that, due to these poles, this case cannot be handled by the approach in[21].

Cascade High Gain Predictors for a Class of Nonlinear Systems

Tarek Ahmed-Ali, Estelle Cherrier, and Françoise Lamnabhi-Lagarrigue

***Abstract*—This work presents a set of cascade high gain predictors to reconstruct the vector state of triangular nonlinear systems with delayed output. By using a Lyapunov-Krasvoskii approach, simple sufficient conditions ensuring the exponential convergence of the observation error towards zero are given. All predictors used in the cascade have the same structure. This feature will greatly improve the easiness of their implementation. This result is illustrated by some simulations.**

***Index Terms*—Cascade systems, high gain observer, time-delay systems.**

I. Introduction

In this technical note, the design of nonlinear observers for nonlinear systems with delayed output measurements is investigated. This problem appears in many control systems areas, such as networked control systems, where the data are transmitted through a communication

Manuscript received November 29, 2010; revised April 06, 2011; accepted June 29, 2011. Date of publication July 14, 2011; date of current version December 29, 2011. Recommended by Associate Editor X. Xia.

T. Ahmed-Ali and E. Cherrier are with GREYC UMR CNRS 6072, Caen Cedex 14050, France (e-mail: tarek.ahmed-ali@ensicaen.fr; estelle.cherrier@ensicaen.fr).

F. Lamnabhi-Lagarrigue is with CNRS, Laboratoire des Signaux et Systemes (INSIS – INS2I), European Embedded Control Institute (EECI) SUPELEC Gif-sur-Yvette 91192, France (e-mail: lamnabhi@lss.supelec.fr).

Digital Object Identifier 10.1109/TAC.2011.2161795