Joint 48th IEEE Conference on Decision and Control and
28th Chinese Control Conference
Shanghai, P.R. China, December 16-18, 2009



# A Convex Optimization Approach to Synthesizing Bounded Complexity $\ell^\infty$ Filters

Franco Blanchini    Mario Sznaier.

*Abstract*— **This paper considers 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 in this paper 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 paper shows that this can be accomplished by using Linear Time Invariant (LTI) filters of order no larger than $r$. Further, synthesizing these filters reduces to a combination of convex optimization and line search.**

## I. INTRODUCTION

Classical stochastic estimation methods are not well suited for situations where it is of interest to obtain hard bounds on estimation errors or where the only information available on exogenous disturbances is a bound on a suitable norm (or, alternatively, a set-membership characterization). These cases can be handled by resorting to a deterministic, unknown-but-bounded approach where the goal is to design an estimator that minimizes, in a suitable sense, the worst case estimation error due to exogenous inputs only known to belong to a given set. Initial work in this area dates back to the early 70's [12], [3], where it was shown that in the case of $\ell^2$ bounded exogenous disturbances, the set of states consistent with the experimental observations is an ellipsoid whose center and covariance matrix can be recursively obtained via a Kalman–filter like estimator. Unfortunately, this is no longer the case for point-wise in time (e.g. $\ell^\infty$ like) constraints on the disturbance. In this case, even constraining the disturbances to belong to an ellipsoid at each point in time does not lead to easily characterizable consistency sets for the states, although these sets can be conservatively overbounded by an ellipsoid.

Worst case estimation in the presence of $\ell^\infty$ bounded disturbances was studied in [8], [10], [18] (see also the survey [9]). 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. In the case of stable systems, given $\epsilon > 0$, $\epsilon$–suboptimal approximations can be found by simply dropping all measurements older than an a-priori pre–computable horizon $N(\epsilon)$. Still, guaranteeing a small approximation error requires large values of $N$ (see [18] for details.) Moreover, the filter is non–recursive, in the sense that current estimates are obtained by solving an LP problem that involves all available information, rather than by propagating past estimates.

The use of nonlinear recursive filters was proposed in [16], where the idea is 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 and its sub-optimality properties hard to ascertain.

A semi–recursive algorithm was proposed in [19]. In the case of known initial conditions, the optimal $\ell^\infty$ estimation problem is reduced to an $\ell^1$ model matching problem [2], [6], [5] that can be solved (with arbitrary precision) by using the techniques in [5]. The case of unknown initial conditions is handled by first pre–computing an horizon $N$ after which the estimation error due to these initial conditions falls below a pre-specified error level $\epsilon$. The complete, semi–recursive estimator is obtained by using a non–recursive pointwise optimal estimator similar to that in [18] for the first N-1 time steps, switching afterwards to the recursive $\ell^1$ estimator. Since this estimator is based on solving a 2–block $\ell^1$ model matching problem its complexity (and hence that of the overall estimator) cannot be bounded a priori.

An alternative approach involves set–valued observers [14], [15], where pointwise optimal estimators are obtained by recursively applying the Fourier–Motzkin algorithm to construct a polyhedral set guaranteed to contain the states of the plant. An $\ell^\infty$ point–wise optimal estimator is then obtained from these sets, by simply using as estimate of the unknown output $z$ the center $z_c$ of the set of all output values compatible with the present set estimate of the state. However, propagation of these estimates is not recursive, e.g. $z_c(k+1)$ cannot be directly constructed from the past estimates $z_c(k-i)$. Moreover, in principle the complexity of the estimator (measured in terms of the number of hyperplanes defining the set observer) is not bounded a-priori and increases with time.

Motivated by the high complexity entailed in the approaches above, the goal of this paper is to synthesize *fixed*

Franco Blanchini has been partially supported by the Italian Ministry for University and Research and by the Italian National Council of Research (C.N.R.). M. Sznaier has been partially supported by NSF grants ECS–0648054, IIS–0713003, and AFOSR grant FA9550–09–1–0253.

Corresponding author, Dept. of Mathematics and Computer Science, University of Udine, Italy. (blanchini@uniud.it)

Dept. of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA (msznaier@coe.neu.edu)

order recursive filters for systems subject to $\ell^\infty$ bounded disturbances, with guaranteed worst case estimation error. Our main results show that this problem can be reduced to a linear programming problem, provided that the initial condition is confined to a suitable set. For initial conditions outside this set, the estimation error converges, in finite time, to the design value.

## II. PRELIMINARIES

For ease of reference, the notation used in the paper is summarized below:

- $\|y\|_\infty$ — $\infty$ norm of the vector $y \in R^n$: $\|y\|_\infty \doteq \max_i |y|_i$.
- $\|M\|_1$ — $\infty \to \infty$ induced norm of matrix $M \in R^{n\times m}$: $\|M\|_1 \doteq max_i \sum_j |M_{ij}|$
- $\ell^1_n, \ell^\infty_n$ — extended Banach spaces of vector valued real sequences $\{y\}_0^\infty \in R^n$ equipped with the norms $\|y\|_{\ell^1} \doteq \sum_{i=0}^\infty \|y_i\|_\infty$ and $\|y\|_{\ell^\infty} \doteq \sup_i \|y_i\|_\infty$, respectively.
- $\mathcal{B}\ell^1$, $\mathcal{B}\ell^\infty$ — unit balls in $\ell^1$, $\ell^\infty$.
- $\mathcal{B}\ell^\infty(\mu)$ — scaled unit ball in $\ell^\infty_n$. Given $\mu \doteq [\mu_1 \dots \mu_n]$
$$\mathcal{B}\ell^\infty(\mu) \doteq \{e \in \ell^\infty_n : e_i/\mu_i \in \mathcal{B}\ell^\infty\}$$
- $\|G\|_{\ell^\infty \to \ell^\infty}$ — $\ell^\infty$ to $\ell^\infty$ induced norm of the operator $G : \ell^\infty \to \ell^\infty$, e.g. $\|G\|_{\ell^\infty\to\ell^\infty} \doteq \sup_{y\neq 0} \frac{\|Gy\|_{\ell^\infty}}{\|y\|_{\ell^\infty}}$
- $y(\lambda)$ — $\lambda$–transform of a sequence $\{y_k\}_0^\infty$ $y(\lambda) \doteq \sum_{i=0}^\infty y_k \lambda^k$

In the sequel, scalar ARMA models of the form

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

will be associated with their corresponding $\lambda$–transform representation[1]:

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

The notion of equalized performance, introduced in [4] (see also [11]) will play a key role in obtaining bounded complexity filters.

*Definition 2.1:* Consider an LTI plant described by a model of the form (1). Given $r \geq n$, the plant achieves an equalized $r$–performance level $\mu$ if, whenever the input and output sequences $\{v\}, \{y\}$ satisfy $|v(t)| \leq 1$ and $|y(t)| \leq \mu$ for all $t = k, k-1, \dots, k-r+1$, 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 [11]) achieve (finite) equalized performance. However, any stable plant achieves finite equalized $r$–performance for some large enough $r$. Further, if a SISO plant achieves $r$–performance $\mu$ for some finite $r$, the it achieves $r'$–performance $\mu$ for any $r' > r$.

[1]This representation can be obtained from the the standard $z$–representation by simply setting $\lambda = 1/z$.

Next, we recall, for ease of reference, some properties concerning the relationship between equalized performance and the $\ell^\infty$ 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 FIR plants.
2) $\mu(r) \downarrow \|G\|_{\ell^\infty\to\ell^\infty}$.

### A. *Why equalized filtering?*

As already shown in [13], [16] recursive set valued observers based upon the idea of propagating a set known to contain the (unknown) state of the plant have high complexity. To avoid this difficulty, in this paper, 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 $||v||_{\ell^\infty} \leq 1, \|w\|_{\ell^\infty} \leq 1$. Further, we are interested in synthesizing the tightest hyperrectangle satisfying this property. The main result of this paper 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.

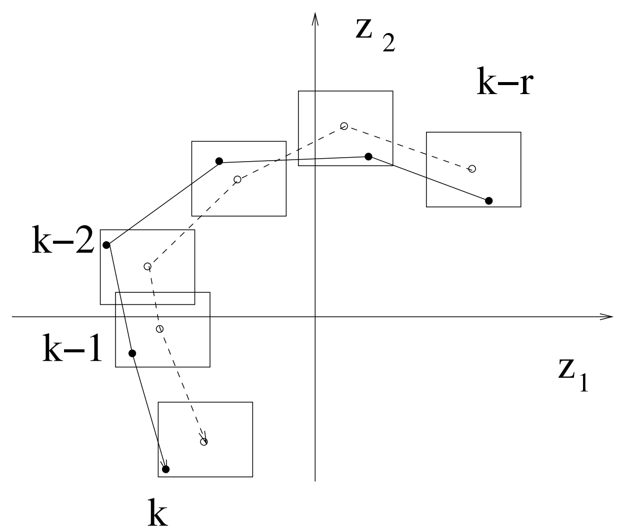


Fig. 1. The equalized filtering idea: actual (black dots) versus estimated (circles) trajectories

## III. PROBLEM SETUP AND PRELIMINARY RESULTS

Consider an LTI plant subject to $\ell^\infty$ bounded disturbances, with state space realization:

$$x_{k+1} = Ax_k + Bv_k \quad (3)$$
$$z_k = Hx_k \quad (4)$$
$$y_k = Cx_k + Dw_k \quad (5)$$

or, equivalently,with $\lambda$–transform representation

$$z(\lambda) = \frac{M(\lambda)}{d(\lambda)} v(\lambda) \quad (6)$$

$$y(\lambda) = \frac{N(\lambda)}{d(\lambda)} v(\lambda) + Dw(\lambda) \quad (7)$$

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 $d(\lambda) = det(I - \lambda A)$. For the time being, we will assume that $z$ is a scalar, but this assumption will be relaxed later. Our goal is to design a filter of the form:

$$\hat{z}(\lambda) = \frac{B(\lambda)}{a(\lambda)} y(\lambda) \quad (8)$$

such that the estimation error

$$e(\lambda) = z(\lambda) - \hat{z}(\lambda) \quad (9)$$

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:

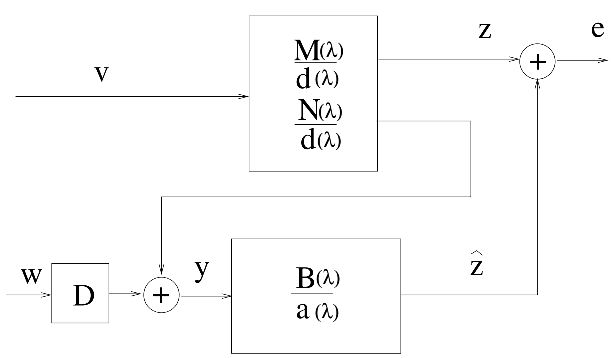


Fig. 2. The filtering scheme.

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

$$\xi_{k+1} = P\xi_k + Ly_k \quad (10)$$

$$\hat{x}_k = Q\xi_k + Ry_k \quad (11)$$

$$\hat{z}_k = H\hat{x}_k \quad (12)$$

is a generalized state observer for system (3)–(5) if $P$ is a stable matrix and $\hat{x}_k - x_k \to 0$ as $k \to \infty$, when $w(k) \equiv 0$ and $v(k) \equiv 0$.

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

*Lemma 3.1:* The system (10)–(12) is a generalized observer for (3)–(5) iff $P$ is stable and there exists a full rank matrix $T$ such that

$$TA - LC = PT, \quad (13)$$

$$QT + RC = I, \quad (14)$$

*Proof:* See [7], [17]. ■

*Remark 3.1:* The standard Luenberger observer corresponds to the choice $T = I$ and $R = 0$. 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 (8) is a generalized state observer for system (3)–(5), then the polynomial matrices $M(\lambda)$ (of dimension $1 \times p$), $N(\lambda)$ (of dimension $q \times p$), $B(\lambda)$ (of dimension $1 \times d$) and the polynomials $a(\lambda)$ and $d(\lambda)$ satisfy the following condition:

$$M(\lambda)a(\lambda) - B(\lambda)N(\lambda) = C(\lambda)d(\lambda) \quad (15)$$

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

*Proof:* From equations (10)–(14) it follows that:

$$\begin{aligned} [Tx - \xi]_{k+1} &= TAx_k - P\xi_k - L(Cx_k + Dw_k) + TBv_k \\ &= P[Tx_k - \xi_k] + TBv_k - LDw_k \\ x_k - \hat{x}_k &= x_k - Q\xi_k - RCx_k - RDw_k \\ &= Q[Tx_k - \xi_k] - RDw_k \end{aligned}$$

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:

$$\begin{aligned} \begin{bmatrix} \eta_{k+1} \\ \theta_{k+1} \end{bmatrix} &= \begin{bmatrix} A & 0 \\ 0 & P \end{bmatrix} \begin{bmatrix} \eta_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} B & 0 \\ TB & -LD \end{bmatrix} \begin{bmatrix} v_k \\ w_k \end{bmatrix} \\ e_k &= \begin{bmatrix} 0 & HQ \end{bmatrix} \begin{bmatrix} \eta_k \\ \theta_k \end{bmatrix} + \begin{bmatrix} 0 & -HRD \end{bmatrix} \begin{bmatrix} v_k \\ w_k \end{bmatrix} \end{aligned}$$

Thus $\eta$ is unobservable from $e$. Hence, the modes of $A$ are canceled in the transfer function $T_{e,\eta}$. From (6)–(9) it follows that:

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

Since $d(\lambda) = det(I - A\lambda)$, the cancelation of the modes of $A$ in $T_{e\eta}$ implies that $M(\lambda)a(\lambda) - B(\lambda)N(\lambda)$ has $d(\lambda)$ as a factor, precisely what (15) states. ■

Since we are interested in generalized observer like filters, in the sequel we will limit our attention to polynomial matrices satisfying (15) for some $C(\lambda)$. It is trivial to show that in this case the estimation error is governed by the equation

$$e(\lambda) = \frac{C(\lambda)}{a(\lambda)} v(\lambda) + \frac{B(\lambda)}{a(\lambda)} Dw(\lambda) \quad (16)$$

## IV. Equalized performance filtering

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

*Problem 4.1:* Given an integer $r \geq n$ and $\mu > 0$ find a filter of the form (8) of order $r$ satisfying the constraint (15) and such that $a(\lambda)$ is stable (i.e. all its poles are outside the unit circle) and

$$\begin{array}{c} |e_k| \leq \mu, k = 0, 1, 2, \ldots, r-1 \Rightarrow |e_t| \leq \mu \\ \text{for all } t \text{ and all sequences } v, w \in \mathcal{B}\ell^\infty \end{array} \quad (17)$$

Note that the problem above does not explicitly make any assumptions on $x_o$, the initial conditions of the plant. 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 condition will be satisfied after a finite number of steps.

*Theorem 4.1:* An $r^{th}$ order filter of the form (6)–(7) with:

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

solves Problem 4.1 above if and only if

$$\begin{array}{rcl} \mu\|[a_1\ a_2 \ldots a_r]\|_1 & + & \|[C_0\ C_1 \ldots C_r]\|_1 \\ & + & \|B_0 D\ \ldots B_r D\|_1 \leq \mu \end{array} \tag{18}$$

*Proof:* From (16) the ARMA model relating the signals $e, v, w$ is given by

$$e_k = -\sum_{i=1}^{r} a_i e_{k-i} + \sum_{i=0}^{r} C_i v_{k-i} + \sum_{i=0}^{r} B_i D w_{k-i} \tag{19}$$

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

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

Therefore the condition is sufficient. To prove necessity, start by rewriting (19) as

$$\begin{array}{rcl} |e(k)| & = & |[\mu a_1 \ldots \mu a_r\ C_0 \ldots C_r\ B_0 D \ldots B_r D]x| \\ & \doteq & |\Xi x| \end{array}$$

where $x \doteq \left[e_{k-1}/\mu \ldots e_{k-r}/\mu \quad v_k \ldots v_{k-r} w_k \ldots \quad w_{k-r}\right]^T$. 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,

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

which proves necessity. To conclude the proof, we need to establish that condition (18) implies stability of the filter. This follows immediately from the fact that it implies

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

■

In the sequel we will refer to filters satisfying (18) as $r$–equalized filters, with performance $\mu$ or, whenever clear from the context, simply as equalized filters.

## V. Optimal fixed–order synthesis

It is not difficult to see that equalized filter synthesis reduces to a combination of convex optimization and bisection. To this effect, note that, for fixed $\mu$, (18) is convex with respect to $a_k$, $B_k$ and $C_k$, and (15) is a linear constraint in such variables. Hence finding a pair $(a, B)$ so that 15 and (18) are satisfied reduces to a convex feasibility problem. Finally, the optimal $\mu$ can be found via bisection. These observations are summarized in the following algorithm.

*Algorithm 5.1:* 0.- Select $\mu > 0$, tolerances $\epsilon$ and $\delta$, and set $\mu^- = 0$.

1.- Solve the feasibility problem (18) subject to (15). If it is unfeasible, set $\mu = 2\mu$ and go to step 1, else set $\mu^+ = \mu$.
2.- Solve the feasibility problem for $\mu = (\mu^+ + \mu^-)/2$.
3.- If it is feasible, set $\mu^+ = \mu$ else set $\mu^- = \mu$.
4.- If $\mu^+ - \mu^- < \delta$ then STOP, else go to step 3.

### A. The multi–output case

In the previous sections we considered the case where $z$, the quantity to be estimated, is a scalar. The main result of this section shows that these results can be extended to the multiple outputs case, $z \in R^s$ by simply considering an array of single–output filters, each of which estimates one of the components of $z$. To this effect, we begin by extending the definition of equalized filtering performance to the multi–output case.

*Definition 5.1:* The filter (8) with error $\hat{z} - z = e \in R^s$ is said to achieve a vector equalized performance level $\mu \doteq \left[\mu_1, \mu_2, \ldots \mu_s\right]$ if it is stable and:

$$\begin{array}{l} e_{k-j} \in \mathcal{B}\ell^\infty(\mu), j = 1, 2, \ldots, r \Rightarrow e_k \in \mathcal{B}\ell^\infty(\mu); \\ \text{for all sequences } v, w \in \mathcal{B}\ell^\infty \end{array} \tag{20}$$

The next result shows that vector equalized performance is equivalent to component–wise scalar equalized performance.

*Theorem 5.1:* A filter $F\colon y \in \ell_n^\infty \to \hat{z} \in \ell_s^\infty$ achieves a vector equalized performance level $\mu$ iff each component $F_i\colon y \in \ell_n^\infty \to \hat{z}^i \in \ell^\infty$ achieves scalar equalized performance (in the sense of (17)) $\mu_i$, where $\hat{z}^i$ denotes the $i^{\text{th}}$ component of $\hat{z}$.

*Proof:* Clearly, if each component $F_i$ achieves an equalized performance level $\mu_i$, the overall filter $F$ obtained by stacking each component satisfies the conditions in Definition 5.1. Conversely, assume that the filter $F$ satisfies (20). Note that the multiple–output version of the filter (19) can be written in terms of its $h$ component as follows

$$e_k^h = -\sum_{i=1}^{r} a_i^h e_{k-i}^h + \sum_{i=0}^{r} C_i^h v_{k-i} + \sum_{i=0}^{r} B_i^h D w_{k-i} \tag{21}$$

and that the error terms $e_{k-i}^h$, $i = 1, 2, \ldots, r$ can be initialized independently in each "partial filter". Assume now that for a given $i$ the corresponding mapping $F_i$ does not satisfy (17). Then, it is easily seen that initializing all the other variables $e^j$, $j \neq h$, to $e^j_{k-i} = 0$, $i = 1, 2, \ldots, r$ leads to violation of (20). ■

Hence, optimal MIMO filters can be synthesized by simply applying Algorithm 5.1 componentwise.

## VI. Filter initialization

In this section we consider the problem of filter initialization. The main result shows that, given $r$ measurements, $\mathbf{y} \doteq \left[y_o, y_1, \ldots, y_{r-1}\right]$ there exists a set $\mathcal{X}_o(\mathbf{y}, \mu)$ and a filter initial condition $\xi_o$ such that if the unknown initial condition of the system $x_o \in \mathcal{X}_o(\mathbf{y}, \mu)$, then the estimation error satisfies $e_k \in \mathcal{B}\ell^\infty(\mu)$ for all $k$. For initial conditions $x_o \notin \mathcal{X}_o$, then $e_k \in \mathcal{B}\ell^\infty(\mu)$ for $k > r$.

Proceeding as in [3], [14], [15], [8], given a sequence of measurements $\mathbf{y}$ define recursively the following sequence of sets:

$$\begin{aligned} \tilde{\mathcal{X}}_k &= A\mathcal{X}_{k-1} + B\mathcal{B}\ell^\infty && (22) \\ \mathcal{X}_k &= \tilde{\mathcal{X}}_k \bigcap \{x : Cx - y_k \in D\mathcal{B}\ell^\infty\} && (23) \\ \mathcal{Z}_k &= H\mathcal{X}_k && (24) \end{aligned}$$

where $\tilde{\mathcal{X}}_o$ is a set known to contain the initial condition (if no information is available then $\tilde{\mathcal{X}}_o = R^n$). Set $\tilde{\mathcal{X}}_k$ is the set of states that can be reached from $\mathcal{X}_{k-1}$, while $\mathcal{X}_k$ is the subset compatible with the measurements, e.g., the best current estimate of the set of all states consistent with the a priori information and the experimental measurements. Finally, $\mathcal{Z}_k$ is the corresponding set of possible values of $z_k$. The techniques to compute these sets are well established (see for instance [3], [14] and references therein). Assuming that $\mathcal{Z}_k$ becomes bounded after some time[2], the filter can be initialized (after $r$ measurements) as follows (see Fig. 3).

1) Compute the sets $\mathcal{Z}_{k-i}$, $i = 1, 2, \ldots, r$.
2) For $j = 1, \ldots, s$, let:

$$\begin{aligned} z^{j,+}_{k-i} &\doteq \max_{\xi \in \mathcal{Z}_{k-i}} \{\xi^j\}, \\ z^{j,-}_{k-i} &\doteq \min_{\xi \in \mathcal{Z}_{k-i}} \{\xi^j\}, \\ \mu^j_{k-i} &\doteq \tfrac{1}{2}|z^{j,+}_{k-i} - z^{j,-}_{k-i}|, \\ z^{j,c}_{k-i} &\doteq \tfrac{z^{j,+}_{k-i} + z^{j,-}_{k-i}}{2}, \end{aligned} \qquad (25)$$

3) Let $\mu^{init,j} = \max_{k-r \leq t \leq k-1} \{\mu^j_t\}$ and choose as the first $r$ filter estimates $\hat{z}^j_t = z^{j,c}_t$, $t = k-r, \ldots, k-1$.

Note that if $\mathcal{X}_o$ is convex, then $z^+, z^-$, and $z^{c}$[3] above can be found by simply solving a convex optimization problem. Further, if $\mathcal{X}_o$ is a polytope, this problems reduces to LP. Since $\| \left[a_1 \ldots a_r\right] \|_1 < 1$ by construction, it can be easily shown that if (17) holds for some $\tilde{\mu}$, then it also holds for all $\mu \geq \tilde{\mu}$. It follows that if $\mu^{init,j} \leq \mu^j_o$, the optimal equalized performance level in (18), then the filter (8), with the initialization above, achieves optimal equalized performance level $\mu^j_o$ for all initial conditions $x \in \mathcal{X}$. On the other hand, as we show next, if $\mu^{init,j} > \mu^j_o$, then the worst case $\ell^\infty$ estimation error is bounded above by $\mu^{init,j}$ and converges, in a finite number of steps, to $\mu^j_o$.

[2]this condition holds for all $t$ if $\mathcal{X}_0$ is compact, and for $t \geq n$ if $(A, C)$ is observable

[3]The estimate $z_c$ is precisely the central estimator introduced in [13].

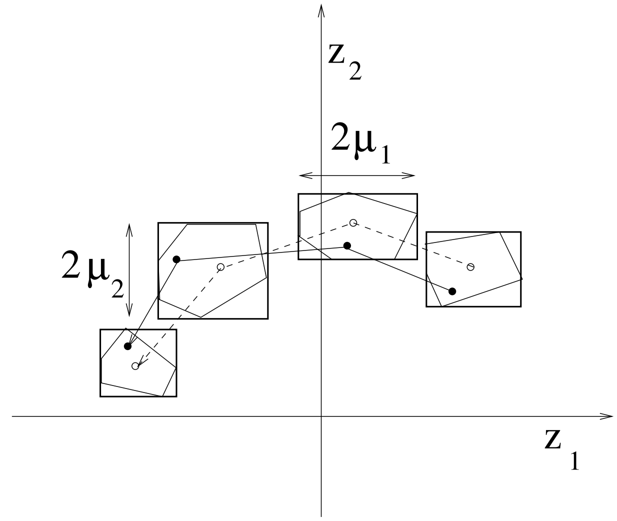


Fig. 3. The filter initialization

*Theorem 6.1:* Consider a filter of the form (8). Given $\mu > 0$ satisfying (18), 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:* Follows from showing that the sequence $\psi_k \doteq \max_{i=1,\ldots,r} \; |e_{k-i}|$ is non increasing in $\psi_k > \mu$ and contains a strictly decreasing subsequence. (Details, omitted for space reasons, can be obtained by contacting the authors). ■

## VII. Connection with existing results

Next, we briefly comment on the connection with existing approaches. The initialization procedure described above is equivalent to using the set valued observers proposed in [13], [14] for the first $r$ steps, switching afterwards to the filter (8). In this sense, the proposed algorithm resembles the approach in [19], where pointwise optimal estimators are used until the $\ell^\infty$–induced filter reduces the error due to the unknown initial conditions below a given tolerance $\epsilon$, switching then to the latter filter. However, the approach proposed here differs in several aspects, in addition to its ability to fix, a–priori, the complexity of the filter. Specifically, (i) the set–valued filters are used for a fixed horizon (equal to the order $r$ of the equalized filter), as opposed to a problem–dependent horizon, and (ii) the filter (8) is switched on after the estimation error sequence has been driven to a hyperrectangle of size $\mu$, the optimal worst–case estimation error, rather than below the tolerance $\epsilon$ (typically $\mu \gg \epsilon$).

## VIII. Illustrative examples

***Example 1:*** Consider the following second order plant:

$$\frac{m(\lambda)}{d(\lambda)} = \frac{\lambda^2}{1} \quad \frac{n(\lambda)}{d(\lambda)} = \frac{(1-0.5\lambda)(1-2\lambda)}{1}$$

and assume that the process and measurement noise satisfy $|v_k| \leq \gamma$ and $|w_k| \leq \beta$, respectively. In the sequel, we analyze optimal filter behavior as a function of these parameters. For $0 < \gamma < 2$ the optimal equalized estimate is $\hat{z} = 0$, (e.g. zero filter), with $\mu_{opt} = \gamma$.

For $\beta = 1$, $\gamma = 2$, the problem admits multiple solutions, amongst them $\frac{b(\lambda)}{a(\lambda)} = 0$ and

$$\frac{a(\lambda)}{b(\lambda)} = \frac{-0.1517 - 0.4870\lambda - 0.1609\lambda^2 - 0.0464\lambda^3}{1.0000 - 0.0656\lambda - 0.0450\lambda^2 - 0.0464\lambda^3}$$

Finally, the case $\beta = 1$, and $\gamma > 2$ seems to yield, independently of $\gamma$, 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 $\gamma$:

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

with $\kappa \approx 0.28$.

***Example 2:*** Next, we consider the case of a plant with poles on the stability boundary[4]:

$$\frac{m(\lambda)}{d(\lambda)} = \frac{\lambda^2}{1 - \lambda^2} \quad \frac{n(\lambda)}{d(\lambda)} = \frac{(1 - 0.5\lambda)(1 - 2\lambda)}{1 - \lambda^2}$$

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 $\mu_{opt} = 5.1$

## IX. CONCLUSION AND DISCUSSION

Most of the existing work on filtering in the presence of unknown–but–bounded noise is based on constructing first the consistency set for the states of the plant (e.g. the set of states compatible with both a-priori assumptions and experimental measurements). Unfortunately, this approach leads to filters whose complexity can be arbitrarily large, and potentially grows online. Overbounding these sets (using for instance ellipsoids or the approach in [12], [16]), produces conservative filters with hard–to–ascertain optimality properties. Alternatively, receding horizon based approaches to filtering (see for instance [1]) require solving online non–trivial optimization problems.

To avoid these difficulties, in this paper we propose a different approach, based on the idea of *equalized* performance, first introduced in [4] in the context of suboptimal $\ell^1$ controller design. The main idea is to, rather than attempting to find bounded complexity sets that contain the consistency set, work directly with r-length estimation error sequences, confining them to the tightest possible hyperrectangle. As shown in the paper, this can be achieved with an $r^{th}$ order LTI filter, whose coefficients can be found via convex optimization. Further, as opposed to the control case, multiple outputs can be readily handled by simply considering a collection of scalar filters.

These results were illustrated with some simple examples. 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 estimation error equation governing the filtering error does not exhibit this feature.

Research is currently underway seeking to extend the results presented in this paper to switched, piecewise linear systems.

[4]Note that, due to these poles, this case cannot be handled by the approach in [19].

## REFERENCES


[1] A. Alessandri, M. Baglietto, and G. Battistelli. "Receding-horizon estimation for discrete-time linear system", IEEE Trans. Automat Control, Vol. 46, no. 3, pp. 473–478, 2003.

[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] Bertsekas, D. P. and Rhodes, I. B.,"Recursive state estimation for a set-membership description of uncertainty", IEEE Trans. Automat Control, Vol. 16, pp. 117–128, 1971.

[4] Blanchini, F. and Sznaier, M., "A convex optimization approach to fixed–order controller design for disturbance rejection for SISO systems", IEEE Transactions on Automatic Control, Vol. AC-45, no. 4, pp. 784-789, April 2000.

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

[6] M. A. Dahleh and J. B. Pearson, "$l^1$–optimal feedback controllers for MIMO discrete–time system", IEEE Trans. Automat Control, Vol, 32, no. 4, pp. 314–322, 1987.

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

[8] Milanese, M. and Tempo, R., "Optimal algorithms theory for robust estimation and prediction", IEEE Trans. Automat Control, Vol. 30, no. 5, pp. 730–738, 1985.

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

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

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

[12] Schweppe, F. C., "Recursive state estimation: Unknown but bounded errors and system inputs," IEEE Trans. Automat. Contr., Vol. 13, pp. 22–28, 1968.

[13] Shamma, J. S. and Tu, E. "Optimality of Set-Valued Observers for Linear Systems," *Proc. 34*$^{\text{th}}$ *IEEE CDC,* New Orleans, LA, Dec. 1995, pp. 2081–2086.

[14] J. S. Shamma and K. Y. Tu, "Approximate set-valued observers for nonlinear systems", IEEE Trans. Automat Control, Vol. 42, no. 5, pp. 648–658, 1997.

[15] Shamma, J. and Tu, KY, "Set–valued observer and Optimal Disturbance Rejection", IEEE Trans. Automat. Control, Vol. 44, no.2, pp. 253–164, Feb. 1999.

[16] Stoorvogel, A. A., "$\ell^1$ state estimation for linear systems using nonlinear observers,", *Proc. 35*$^{\text{th}}$ *IEEE CDC,* Kobe, Japan, pp. 2407–2411, Dec. 1996.

[17] Tsui, C., "A new design approach to unknown input observers," IEEE Trans. Automat. Control, Vol. 41, pp. 464–468, 1996.

[18] Tempo, R., "Robust estimation and filtering in the presence of bounded noise," IEEE Trans. Aut. Control, Vol. 33, no. 9, pp. 864–867, 1988.

[19] Voulgaris, P., "On optimal $\ell^\infty$ to $\ell^\infty$ filtering", Automatica, Vol. 31, pp. 489–495, 1995.