2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Bounded Complexity ℓ^∞ Filters for Switched Systems

Mario Sznaier Burak Yilmaz Franco Blanchini

Abstract— **This paper considers the worst–case estimation problem in the presence of unknown but bounded noise for piecewise linear switched systems. 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 approach leads to computationally tractable filters, that only require the on-line solution of a bounded complexity convex optimization problem. Moreover, as we show in the paper, these filters are (worst-case) optimal when operating in a simplified, restricted information scenario.**

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 [3], [11], where it was shown that in the case of ℓ^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. ℓ^∞ 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. A related line of work seeks to design Luenberger type observers for switched systems, based on the use of a single [2], [5] or multiple [10] Lyapunov functions. While the resulting observers are attractive because of their simplicity, applicability is limited to cases where Lyapunov functions that satisfy the required conditions (usually a set of Lyapunov type equations) exist.

Worst case estimation in the presence of ℓ^∞ bounded disturbances was studied in [6], [8], [16] (see also the survey [7]). 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 $\epsilon > 0$, ϵ–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 [16] 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 [15], 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. Nevertheless, the complexity of the resulting observer is potentially high and its sub-optimality properties hard to ascertain.

An alternative approach involves set–valued observers [13], [14], 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 ℓ^∞ 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 *bounded* complexity filters for systems subject to ℓ^∞ bounded disturbances, with guaranteed worst case estimation error. Intuitively, the main idea is to *purposely* drop information: rather than storing all past values of the output (as is required for optimal set–valued filters), the proposed filter only remembers the past r measurements and the fact that,

This work was supported in part by NSF grants IIS–0713003 and ECCS–0901433, AFOSR grant FA9550–09–1–0253, and the Alert DHS Center of Excellence under Award Number 2008-ST-061-ED0001.

M. Sznaier and B. Yilmaz are with the Dept. of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA email: msznaier@coe.neu.edu, yilmazbur@gmail.com. F. Blanchini is with the Dept. of Mathematics and Computer Science, University of Udine, Italy. email: blanchini@uniud.it

978-1-4577-0079-8/11/$26.00 ©2011 AACC

for the past r time instants, the estimation error is contained in a suitable hyperrectangle. The main result of the paper shows that this idea (a generalization to switched systems of the idea of *equalized* filtering presented in [4]), allows to obtain (worst-case) optimal filters by simply solving on–line a linear programming problem with $\mathcal{O}(r)$ number of constraints, where r is the memory of the filter.

II. PRELIMINARIES

A. Notation

The notation used in the paper is summarized below

$\mathbb{Z}^+$	set of positive integers
$\|y\|_\infty$	∞ norm of the vector $y \in R^n$: $\|y\|_\infty \doteq \max_i \|y\|_i$.
$\|M\|_1$	$\infty \to \infty$ induced norm of matrix $M \in R^{n\times m}$: $\|M\|_1 \doteq max_i \sum_j \|M_{ij}\|$
ℓ^∞_n	extended Banach space of vector valued real sequences $\{y\}_0^\infty \in R^n$ equipped with the norm $\|y\|_{\ell^\infty} \doteq \sup_i \|y_i\|_\infty$.
$\mathcal{B}\ell^\infty$	unit ball in ℓ^∞.
$\mathcal{B}\ell^\infty(\mu)$	scaled unit ball in ℓ^∞_n. Given $\mu \doteq [\mu_1 \ldots \mu_n]$

$$\mathcal{B}\ell^\infty(\mu) \doteq \{e \in \ell^\infty_n : e_i/\mu_i \in \mathcal{B}\ell^\infty\}$$

B. Background on Information Based Complexity

In this section we recall some key results from Information Based Complexity (IBC) required to establish (worst-case) optimality of the proposed filters. For simplicity, we consider the case of bounded operators in ℓ^∞. A general treatment can be found for instance in the book [17].

Let K denote a set in ℓ^∞ and consider two linear operators[1] $S_y, S_z\colon \ell^\infty \to \ell^\infty$. In this context, the estimation problem can be stated as, given an element $f_o \in K$, find an estimate $\hat{z}$ of $z \doteq S_z f_o$ using noisy experimental information $y = S_y f_o + \eta$, where the noise η is only known to belong to some bounded set $\mathcal{N} \subset \ell^\infty$. Note that in general S_y is not invertible, and thus it is not possible to recover f_o, even in the absence of noise. This is related to the concept of consistency set, defined as:

$$\mathcal{T}(y) \doteq \{f \in K\colon y = S_y f + \eta \text{ for some } \eta \in \mathcal{N}\} \qquad (1)$$

that is, the set of all possible elements in K that could have generated the observed data.

Given an estimation algorithm $\hat{z} = \mathcal{A}(y)$ (not necessarily linear), it is of interest to compute its worst case approximation error. For a given measurement y, the *local* error $\epsilon(y, \mathcal{A})$ of a given algorithm $\mathcal{A}$ is defined as the worst case distance between the true quantity z and its estimate $\mathcal{A}(y)$. Since all the elements f that could have generated y belong to $\mathcal{T}(y)$, it follows that:

$$\epsilon(\mathcal{A}, y) \doteq \sup_{f\in\mathcal{T}(y)} \|S_z f - \mathcal{A}(y)\|_\infty \qquad (2)$$

[1]In the IBC literature, S_y and S_z are usually referred to as the information and solution operators, respectively.

Similarly, the global error $\epsilon(\mathcal{A})$ of an algorithm $\mathcal{A}$ is defined as the worst possible case over all possible measurement sequences, that is:

$$\epsilon(\mathcal{A}) \doteq \sup_y \epsilon(\mathcal{A}, y) \qquad (3)$$

Definition 1: An algorithm $\mathcal{A}_o(.)$ is said to be globally optimal if $\epsilon(\mathcal{A}_o) = \inf_{\mathcal{A}} \epsilon(\mathcal{A})$.

The minimum global error $\epsilon(\mathcal{A}_o)$ is called the radius of information, $r(I)$. It provides a lower bound on achievable performance, since no estimation algorithm can have smaller global worst case error. Further, it is a standard fact in IBC that for linear, not necessarily time invariant systems, $r(I)$ can be explicitly computed. This result is quoted below for ease of reference.

Theorem 2.1: [9] Assume that the sets K and $\mathcal{N}$ are convex and balanced, and that the operators S_z and S_y are linear. Then

$$r(I) = \sup_{f\in\mathcal{T}(0)} \|S_z f\|_\infty$$

Note that this result shows that, for the purpose of estimation, the worst–case trajectory is the one that yields identically zero measurements, e.g. $y = 0$.

C. Problem Setup

In this paper we consider state–space switched linear plants of the form:

$$\begin{aligned} x_{k+1} &= A(\sigma_k)x_k + B(\sigma_k)v_k, \|v\|_{\ell^\infty} \le \eta_v &(4)\\ z_k &= H(\sigma_k)x_k &(5)\\ y_k &= C(\sigma_k)x_k + D(\sigma_k)w_k,\ \|w\|_{\ell^\infty} \le \eta_w &(6)\end{aligned}$$

where $\sigma_k \in \Sigma \subset \mathbb{Z}^+$ denotes the mode (or discrete state) variable, $z \in \mathbb{R}^s$, $y \in \mathbb{R}^q$, $v \in \mathbb{R}^p$ and $w \in \mathbb{R}^q$ represent the output to be estimated, the measurements available to the filter, and process and measurement noise, respectively. We will further assume that the a-priori information includes bounds η_v and η_w on the ℓ^∞ norm of the noise sequences. In the interest of clarity, we will assume for the time being that z is a scalar, but this assumption will be relaxed later. For notational simplicity, denote by $\Sigma_t^r \doteq \{\sigma_t, \sigma_{t-1}, \ldots, \sigma_{t-r}\}$ and $\mathcal{Y}_t^r \doteq \{y_t, y_{t-1}, \ldots, y_{t-r}\}$. With this notation, the goal of this paper can be simply stated as designing a causal, finite complexity filter $\hat{z}_t = \mathcal{F}(\mathcal{Y}_t^r, \Sigma_t^r)$ that minimizes the worst case estimation error $\|z_t - \hat{z}_t\|_{\ell^\infty}$.

III. MAIN RESULT

In this section we present the main result of the paper: a convex optimization based bounded complexity filter.

A. Reduced complexity scenario and worst case bounds

The intuition underlying this paper is to avoid the high (potentially infinite) complexity entailed in set valued observers by purposely dropping information. Specifically, rather than keeping the full information about the past, the proposed filter only "remembers" the immediate past r measurements y_{k-j}, and estimates $\hat{z}_{k-j}$, $j = 1, \ldots, r$, where r is a design parameter, and, that for the past r time instants, the quantity to be estimated was confined to a hyperrectangle of size

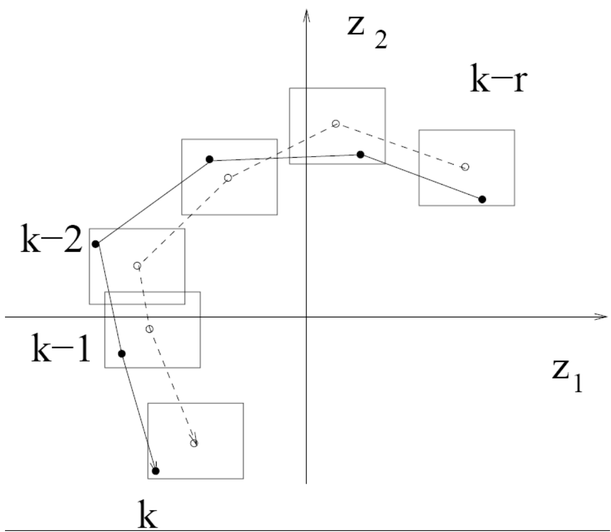

Fig. 1. The equalized filtering idea: full dots are the true trajectory the empty dots are the estimated trajectory

μ, which will be explicitly computed later in the paper, centered around $\hat{z}_{k-j}$ (see Figure 1). As we show in the sequel, this reduced complexity scenario leads to computationally tractable, bounded complexity filters. Moreover, exploiting the concepts from Information Based Complexity briefly described in section II-B allows for establishing that these filters are globally (worst–case) optimal. The first step towards establishing this result is to compute the radius of information for filters operating in the simplified information scenario considered here. Formally, this scenario amounts to replacing, at time t, the set of all past measurement older than $t-r$ by the set $\mathcal{Z}_{t-1}^r \doteq \{\hat{z}_{t-1}, \ldots, \hat{z}_{t-r}\}$ of "noisy" measurements of z given by

$$\hat{z}_k = z_k + e_k,\ k = t-1, \ldots, t-r,\ |e_k| \leq \mu \tag{7}$$

where $\hat{z}_k$ denote the past estimates of z. Thus, in this context, the filtering problem under consideration can be simply restated as:

Problem 3.1: Given the $r+1$ measurement $\mathcal{Y}_t^r$, past r filter estimates $\hat{z}_k$ and a bound μ on the past estimation errors, find $\hat{z}_t(\mathcal{Y}_t^r, \mathcal{Z}_{t-1}^r, \Sigma_t^r)$ to minimize $\|z_t - \hat{z}_t\|_\infty$.

Note that the problem above minimizes the estimation error only at time t. Clearly, applying this algorithm in a receding horizon fashion leads to an ℓ^∞ optimal filter.

Lemma 3.1: For the system (4)–(6) with measurements $(\mathcal{Y}_t^r,\ \mathcal{Z}_{t-1}^r)$, a given mode variable trajectory Σ_t^r, and a given bound μ on the past estimation errors, the radius of information $r(I, \Sigma_t^r, \mu)$ is given by the solution to the following optimization problem:

$$r(I, \Sigma_t^r, \mu) = \max_{x,v,w} |z_t| \tag{8}$$

$$\begin{aligned} &\text{subject to (4) (5) and} \\ &|z_k| \leq \mu,\ k = t-1, \ldots, t-r, \\ &\|C(\sigma_k)x_k\|_\infty \leq \|D(\sigma_k)\|_1 \eta_w,\ k = t, \ldots, t-r \end{aligned}$$

Proof: Begin by noting that the consistency set $\mathcal{T}(0)$ for the system defined by (4)–(6), and (7) is given by:

$$\begin{aligned} \mathcal{T}(0) \quad = \quad & \{z_t \colon \hat{z}_k = 0,\ k = t-1, ., t-r \text{ and } y_k = 0, \\ & k = t, .., t-r \text{ for some sequences } \{x_k\}, \{v_k\}, \\ & \{e_k\}, \{w_k\} \text{subject to } (4)-(6) \text{ and } (7), \text{ with} \\ & \|v_k\|_\infty \leq \eta_v,\ \|e_k\|_\infty \leq \mu,\ \|w_k\|_\infty \leq \eta_w\} \end{aligned}$$

Since the sets $\|v\|_\infty \leq \eta_v$, $\|e\|_\infty \leq \mu$ and $\|w\|_\infty \leq \eta_w$ are convex, balanced, direct application of Theorem 2.1 yields:

$$r(I, \Sigma_t^r, \mu) \quad = \quad \sup_{z_t \in \mathcal{T}(0)} |z_t| \tag{9}$$

which is precisely (8) where the last inequality follows from the fact that $\{w_k\}$ can be arbitrarily chosen. ■

Lemma 3.2: If the inequality

$$r(I, \Sigma_t^r, \mu_o) \leq \mu_o \tag{10}$$

holds for some μ_o, then it holds for all $\mu > \mu_o$.

Proof: Given $\mu_1 > \mu_o$, let $(x^o, z^o, y^o, v^o, w^o)$ and $(x^1, z^1, y^1, v^1, w^1)$ denote the optimizing sequences associated with $r(I, \Sigma_t^r, \mu_o)$ and $r(I, \Sigma_t^r, \mu_1)$, respectively. Define $\rho = \frac{\mu_1}{\mu_o}$. Since by assumption $\rho > 1$, the sequences $\frac{x^1}{\rho}$, $\frac{z^1}{\rho}$, $\frac{y^1}{\rho}$ $\frac{v^1}{\rho}$ and $\frac{w^1}{\rho}$ are a feasible solution for (8) with $\mu = \mu_o$. Thus, it follows that $\mu_1 = \rho\mu_o \geq \rho r(I, \Sigma_t^r, \mu_o) \geq r(I, \Sigma_t^r, \mu_1)$ ■

The next result shows that if the system is observable, then, for each switching trajectory Σ^r, there exist $\mu_r^* < \infty$ such that (10) holds.

Lemma 3.3: Assume that the system (4)-(6) is observable in the sense that, for each switching trajectory Σ^r with $r \geq n$, the following holds:

$$\text{rank}(\Gamma^r) = n,\ \Gamma \doteq \begin{bmatrix} C(\sigma_{r-n+1}) \\ C(\sigma_{r-n+2})\Phi(r-n+2, r-n+1) \\ \vdots \\ C(\sigma_r)\Phi(r, r-n+1) \end{bmatrix}$$

where $\Phi(.,.)$ denotes the transition matrix of (4)-(6) corresponding to Σ^r. Then, there exists a finite μ_r^* such that (10) holds.

Proof: Let

$$\begin{aligned} \mu^* \quad &\doteq \max_{x,v,w} z_t \text{ subject to } (4), (5), (6) \text{ and} \\ y_k \quad &= 0,\ k = t-1, \ldots, t-r \end{aligned} \tag{11}$$

Since Γ has full column rank, and v and w are bounded, then the set of all initial conditions compatible with $y_k = 0$ is a compact set. Hence μ^* is finite. Next, note that the optimization problems (11) and (8) have the same objective function, but the constraints in the former are a subset of the constraints in the latter. Thus $\mu^* \geq r(I, \Sigma^r, \mu^*)$. ■

The results above justify defining, for each possible switching trajectory of length r, Σ^r, a local worst-case performance index as the solution to the implicit equation

$$\mu(\Sigma^r) = r(I, \Sigma^r, \mu) \tag{12}$$

Note that, for a *fixed* Σ^r, computing μ entails a combination of the convex optimization (8) (in fact a linear programming problem) and a scalar line search. Further, from Lemmas 3.2 and 3.3 it follows that if the system is observable, then $\mu(\Sigma^r)$ is finite. Finally, define

$$\mu_{opt} = \max_{\Sigma^r} \mu(\Sigma^r) \tag{13}$$

Lemma 3.4: μ_{opt} is a lower bound on the worst case estimation error achievable by any filter acting on the restricted measurements set $(\mathcal{Y}_t^r, \mathcal{Z}_{t-1}^r)$.

Proof: From Theorem 2.1 and Lemma 3.1 it follows that for any fixed switching trajectory the worst case estimation error achieved by any filter satisfies

$$\sup_{\mathcal{Y}_t^r, \mathcal{Z}_{t-1}^r} |z_t - \hat{z}_t| \geq r(I, \Sigma^r, \mu)$$

The proof follows by simply maximizing both sides of this equation over all possible r-length switching trajectories. ■

B. Bounded complexity filters

Next, we present the proposed filter, based upon the online solution of two optimization problems. Given Σ_t^r, $\mathcal{Y}_t^r$, $\mathcal{Z}_{t-1}^r$, the past r values of the mode variable σ, measurements, and filter estimates, respectively, define z_t^+ and z_t^- as the solutions to the following optimization problems:

$$z_t^+ = max_{z,v,w,x,e} z_t \tag{14}$$

subject to

$$\begin{array}{l} x_{k+1} = A(\sigma_k)x_k + B(\sigma_k)v_k \\ e_k = \hat{z}_k - H(\sigma_k)x_k, \; k = t-1,..,t-r \\ y_k = C(\sigma_k)x_k + D(\sigma_k)w_k, \; k = t,..,t-r \\ \|v_k\|_\infty \leq \eta_v, \; \|w\|_\infty \leq \eta_w, \|e_k\|_\infty \leq \mu_{opt} \end{array} \tag{15}$$

and

$$z_t^- = min_{z,v,w,x_{t-r}} z \text{ subject to (15)} \tag{16}$$

where μ_{opt}, defined in (13), is found *off-line* prior to running the filter.

Theorem 3.1: Assume that $|z_k - \hat{z}_k| \leq \mu_{opt}, \; k = t-1, \ldots, t-r$. Then the *central* estimator defined by

$$\hat{z}_t \doteq \frac{z_t^+ + z_t^-}{2} \tag{17}$$

is a globally optimal estimator of z_t.

Proof: Denote by $\{x_k^+, z_k^+, v_k^+, w_t^+\}$ and $\{x_k^-, z_k^-, v_k^-, w_t^-\}$ the optimizing sequences in (14) and (16), respectively. Define $\delta x_k \doteq \frac{x_k^+ - x_k^-}{2}$, $\delta v_k \doteq \frac{v_k^+ - v_k^-}{2}$, $\delta e_k \doteq \frac{e_k^+ - e_k^-}{2}$ and $\delta z_k \doteq \frac{z_k^+ - z_k^-}{2}$. It can be easily shown that these quantities satisfy equations of the form:

$$\begin{array}{rcl} \delta x_{k+1} & = & A(\sigma_k)\delta_k + B(\sigma_k)\delta v_k; \|\delta v_k\|_\infty \leq \eta_v \\ \delta z_k & = & H(\sigma_k)\delta x_k = \delta x_k; \; \|\delta e_k\|_\infty \leq \mu_{opt} \\ 0 & = & C(\sigma_k)\delta x_k + D(\sigma_k)\delta w_k, \; \|\delta w_t\|_\infty \leq \eta_w \end{array} \tag{18}$$

Hence, from (9) it follows that

$$\max \frac{z_t^+ - z_t^-}{2} = r(I, \Sigma_t^r, \mu_{opt}) = \mu_{opt} \tag{19}$$

Combining this inequality with (17) yields:

$$\begin{array}{rcl} z_t - \hat{z}_t & \leq & z_t^+ - \hat{z}_t = \frac{z_t^+ - z_t^-}{2} \leq \mu_{opt} \\ z_t - \hat{z}_t & \geq & z_t^- - \hat{z}_t = \frac{z_t^- - z_t^+}{2} \geq -\mu_{opt} \\ & \Longleftrightarrow & |z_t - \hat{z}_t| \leq \mu_{opt} \end{array} \tag{20}$$

The proof follows now from Lemma 3.4 establishing that μ_{opt} is a lower bound on the worst case estimation error. ■

Corollary 3.1: If $|z_k - \hat{z}_k| \leq \mu_{opt}$ for $k = t, \ldots, t-r+1$, then $|z_k - \hat{z}_k| \leq \mu_{opt}$ for all $k \geq t$

Proof: Follows directly from (20) by an induction argument ■

The result above shows that once filter is operating in *steady state*, in the sense that the estimation error has been confined to a μ_{opt} sized hyperrectangle for the past r time instants, this situation will persist into the future. Thus, the questions arise of how to initialize the filter and what happens when the past estimation error exceeds μ_{opt}. These issues are addressed next.

C. Filter Initialization

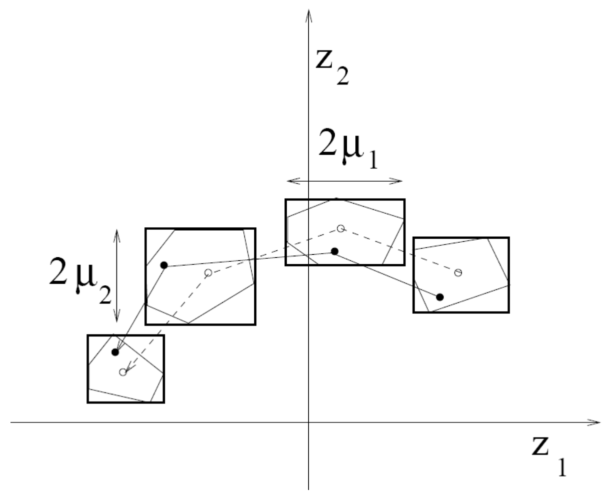

Fig. 2. The filter initialization

In this section we consider the problem of filter initialization. The main result shows that, given an initial set of r measurements, $y \doteq [y_o, y_1, \ldots, y_{r-1}]$ and switching trajectory Σ^r, there exists a finite μ and a set of r filter estimates $\hat{z}_k$, $t = 0, \ldots, r-1$ such that the estimation error satisfies $e_k \in \mathcal{B}\ell^\infty(\mu)$ for all k. The main idea is to compute the central estimator for the first r time steps, using all available measurements[2] That is,

1) at time $0 \leq t \leq r-1$, compute:

$$\begin{array}{rl} z_t^+ & \doteq \max\limits_{x,v,w} z_t \text{ subject to (4) – (6)}, \\ & j = 0, 1, \ldots, r-1 \\ z_t^- & \doteq \min\limits_{x,v,w} z_t \text{ subject to (4) – (6)}, \\ & j = 0, 1, \ldots, r-1 \end{array} \tag{21}$$

[2]The estimate z_c can be thought off as a smoothing problem equivalent of the central estimator introduced in [12].

2) Define:

$$\begin{array}{rcl}\hat{z}_k^{init} & \doteq & \frac{z_k^+ + z_k^-}{2}, \\ \mu_k & \doteq & \frac{1}{2}|z_k^+ - z_k^-|\end{array} \quad (22)$$

3) Let $\mu_{init} = \max_{0 \le t \le r-1} \{\mu_t\}$ and use the estimates $\hat{z}_t,\ t = r = 1, \ldots, 0$ as center of the hyperrectangles (see Fig. 2)

Lemma 3.5: The filter (17), using as initial estimates $\hat{z}_k^{init},\ k = 0, \ldots, r-1$ and error bound μ_{init} has the following properties:

1) If $\mu_{init} \le \mu_{opt}$, then the estimation error satisfies $|e_k| \le \mu_{opt}$ for all k.
2) If $\mu_{init} > \mu_{opt}$, then the estimation error satisfies $|e_k| \le \mu_{init}$ for all k and, given any $\epsilon > 0$, there exist a finite $T(\epsilon)$ such that $|e_k| \le \mu_{opt} + \epsilon$ for all $k \ge T$.

Proof: The proof of Property 1 follows directly from Corollary 3.1. To prove the first part of property 2, consider a fixed switching sequence Σ^r and its associated worst case error $\mu(\Sigma^r)$ defined in (12). Denote by (x^+, z^+, v^+, w^+) and (x^-, z^-, v^-, w^-) the optimizing sequences in (21) and define $\delta x^* = \frac{1}{2}(x^+ - x^-)$, $\delta z^* = \frac{1}{2}(z^+ - z^-)$, $\delta v^* = \frac{1}{2}(v^+ - v^-)$, and $\delta w^* = \frac{1}{2}(w^+ - w^-)$. Let $\rho \doteq \frac{\mu_{init}}{\mu(\Sigma^r)}$. Note that since $\mu(\Sigma^r) \le \mu_{opt} < \mu_{init}$, then $\rho > 1$. Hence the sequences $x^\dagger \doteq \frac{\delta x^*}{\rho}$, $z^\dagger \doteq \frac{\delta z^*}{\rho}$, $v^\dagger \doteq \frac{\delta v^*}{\rho}$ and $w^\dagger \doteq \frac{\delta w^*}{\rho}$ are a feasible solution for (8). Thus, it follows that $\mu(\Sigma^r) \ge |z_r^\dagger| = |\frac{\delta z_r}{\rho}|$, or, equivalently,

$$\mu_{init} = \rho\mu(\Sigma^r) \ge |\delta z_r| \quad (23)$$

By induction, this last inequality implies that $|z_t - \hat{z}_t| \le |z_t^+ - z_t^-| \le \mu_{init}$ for all $t \ge r$. To prove that strict inequality holds in (23), begin by writing $z_r^\dagger$ in terms of $x_o^\dagger$ and $v_k^\dagger$:

$$z_r^\dagger = T_{zx}(r, 0)x_o^\dagger + \sum_j T_{z,v}(r, j)v_j^\dagger \quad (24)$$

where $T_{z,x}(.,.)$ and $T_{z,v}(.,.)$ are the (time varying) operators mapping the initial condition x_o and disturbance sequence $\{v_k\}$ to z. If the system is controllable, then there exists at least one j such that $T_{z,v}(r, j) \ne 0$. Since $\rho > 1$, $\|v^\dagger\|_\infty < \eta_v$, $\|z^\dagger\|_{\ell^\infty} < \mu(\Sigma^r)$ and $\|w^\dagger\|_\infty < \eta_w$. Hence, there exists some $0 < \epsilon$ such that the sequence

$$\tilde{v}_k \doteq \begin{cases} v_k^\dagger & k \ne j \\ v_k^\dagger + \epsilon * \text{sign}(T_{z,v}(r, j)) & k = j \end{cases}$$

is an admissible sequence. Since the cost associated with this sequence is $\tilde{\mu} = |z_r^\dagger| + \epsilon|T_{z,v}(r, j)| \le \mu(\Sigma^r)$, it follows that

$$|z_r - \hat{z}_r| \le \rho|z_r^\dagger| \le \rho\mu(\Sigma^r) - \epsilon_1 = \mu^{init} - \epsilon_1 \quad (25)$$

where $\epsilon_1 \doteq \rho\epsilon|T_{z,v}(r, j)|$. An induction argument shows that as long as $|z_t - \hat{z}_t| > \mu_{opt} \ge \mu(\Sigma^r)$, then the sequence $|z_t - \hat{z}_t|$ is strictly decreasing. Since this sequence is bounded below, it has a limit $\tilde{\mu}$. Assume that $\tilde{\mu} > \mu_{opt}$. Then, proceeding as above, it can be shown that the subsequence $|z_{t+nr} - \hat{z}_{t+nr}| \le \mu^{init} - n\epsilon_1(\tilde{\mu})$, for some $\epsilon_1(\tilde{\mu}) > 0$. However, this contradicts the assumption that $|z_t - \hat{z}_t| \ge \tilde{\mu}$ for all t ■

D. The multi–output case

In the previous sections we considered the case where z, the quantity to be estimated, is a scalar. However, the results of the previous sections directly apply to the case where $z \in R^{n_z}$ by simply adding more constraints to the optimization problems (8), (14) and (16).

IV. ILLUSTRATIVE EXAMPLES

In this section we illustrate our results with some simple examples

Example 1. In this example we consider a system that switches between two unstable plants with state space realizations:

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1.1 & 0.1 \end{pmatrix}, \ B_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \ C_1^T = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ H_1^T = \begin{pmatrix} 1 \\ 1.5 \end{pmatrix}$$

and

$$A_2 = \begin{pmatrix} 0 & 1 \\ 0.2 & 1.1 \end{pmatrix}, \ B_2 = \begin{pmatrix} 0.5 \\ -1 \end{pmatrix}, \ C_2^T = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ H_2^T = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

with $\|v\|_{\ell^\infty} \le 1$ and $\|w\|_{\ell^\infty} \le 1$

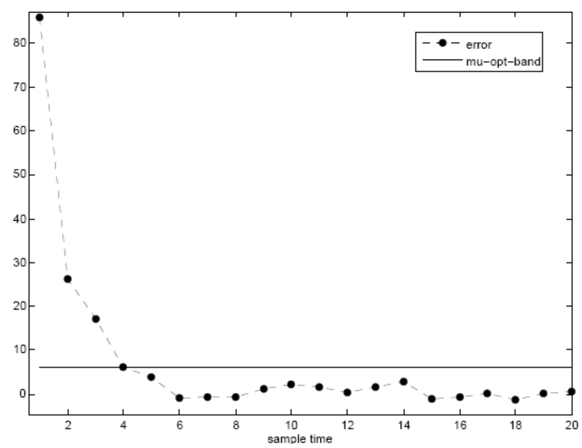

Fig. 3. Estimation Error for Example 1 and initial error $\mu_{init} > \mu_{opt}$

Computing the value of $\mu(\Sigma^r)$ for all possible switching trajectories of length $r = 2$ yields $\mu_{opt} = \max\{\mu(\Sigma^r)\} = 6.15$. Figure 3 shows a sample trajectory for the estimation error of an equalized filter with memory $r = 2$. For this particular trajectory, the initialization procedure outlined in section III-C yields $\mu_{init} = 112.25 > \mu_{opt}$. As shown in the figure, the estimation error quickly converges to within a band of size μ_{opt} and remains there.

Example 2. This example considers the case of a system that switches between two stable plants such that the overall system is not switched stable. The state space realizations of the plants are given by:

$$A_1 = \begin{pmatrix} 0.5 & 0 \\ 2 & 0.5 \end{pmatrix} \text{ and } A_2 = \begin{pmatrix} 0.5 & 2 \\ 0 & 0.5 \end{pmatrix}$$

with $B_1 = B_2 = \begin{pmatrix} 1 & 1 \end{pmatrix}^T$, $C_1 = C_2 = \begin{pmatrix} 1 & 1 \end{pmatrix}$,$\|v\|_{\ell^\infty} \le 1$ and $\|w\|_{\ell^\infty} \le 1$. Using the filter proposed in this paper leads to a value of $\mu_{opt} = 3.56$. A sample trajectory for the case where the initial error falls below μ_{opt} is shown in Figure 4. As expected, the error remains below μ_{opt} for all times.

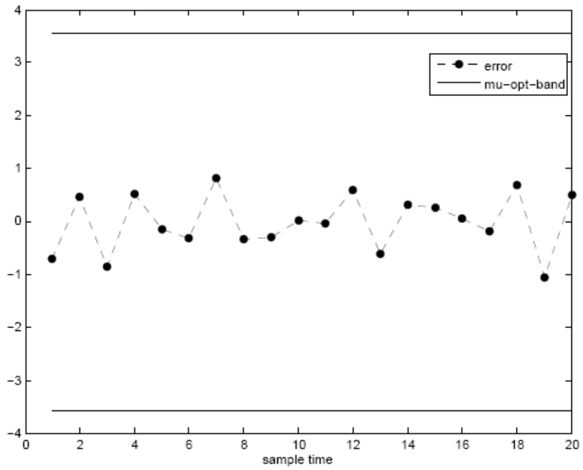

Fig. 4. Estimation Error for Example 2, with $\mu_{init} < \mu_{opt}$

V. CONCLUSION AND DISCUSSION

Filtering in the presence of unknown but bounded noise aims at confining the estimation error within a bounded set. Previous work dealing with the problem, 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), led to filters whose complexity can be arbitrarily large, and potentially grows online. Overbounding these sets (using for instance ellipsoids or the approach in [11], [15]), leads to conservative filters with hard to ascertain optimality properties. The receding horizon approach to filtering (see for instance [1]) requires the solution of non–trivial optimization problems online.

To avoid these difficulties, in this paper we propose a different approach, based on a generalization to switched systems of the idea of *equalized* filtering presented in [4]. 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 by solving on–line a linear programming problem with $\mathcal{O}(r)$ number of constraints, where r is the memory of the filter.

These results were illustrated with two examples involving systems that are not switched–stable. In both cases the estimation error quickly converges to the region $\|e\|_\infty \leq \mu_{opt}$.

REFERENCES

[1] A. Alessandri, M. Baglietto, and G. Battistelli. "Receding-horizon estimation for switching discrete-time linear system", IEEE Trans. Automat Control, Vol. 50, No 11, pp. 1736–1748, 2005.

[2] A. Alessandri and P. Coletta, "Design of Luenberger observer for a class of hybrid linear systems," Hybrid systems: computation and control. Lecture notes in computer science, vol 2034, pp.7–18. Springer-verlag, 2001.

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

[4] F. Blanchini and M. Sznaier, "A Convex Optimization Approach to Synthesizing Bounded Complexity ℓ^∞ Filters," Proc. 48th IEEE Conf. Dec. Control, Shanghai, China, Dec. 16-18, 2009, pp. 217-222.

[5] A. Juloski, W.P.M.H. Heemels and S. Weiland, "Observer design for a class of piecewise linear systems," Int. J. of Robust and Nonlinear Control, 17(15), pp. 1387-1404, 2007.

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

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

[8] 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. Automat Control, Vol. 27, no. 2, pp. 408–414, Apr. 1982.

[9] C. A. Micchelli and T. J. Rivlin, "A Survey of Optimal Recovey," in Optimal Recovery in Approximation Theory, C. A. Micchelli and T. J. Rivlin Eds., Plenum, N.Y., 1977, pp. 1–54.

[10] S. Pettersson, "Observer design for switched systems using multiple quadratic Lyapunov functions," Proc. 13th Mediterranean conference on control and automation. Limassol, Cyprus. June 2005.

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

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

[13] 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.

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

[15] A. A. Stoorvogel, "ℓ^1 state estimation for linear systems using non-linear observers,", Proc. 35$^{\rm th}$ IEEE CDC, Kobe, Japan, pp. 2407–2411, Dec. 1996.

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

[17] J. B. Traub, H. Wozniakowski and G. W. Wasilkowski, Information Based Complexity, Academic Press, 1988.