Proceedings of the 38th
Conference on Decision & Control
Phoenix, Arizona USA • December 1999

FINITE HORIZON ROBUST KALMAN FILTER DESIGN

Minyue Fu[1], *Carlos E. de Souza*[2] *and Zhi-Quan Luo*[3]

[1]Dept. Electrical and Computer Eng., Univ. Newcastle, Australia (eemf@ee.newcastle.edu.au)
[2]Dept. Systems and Control, Laboratório Nacional de Computação Científica – LNCC, Brazil
[3]Dept. Electrical and Computer Eng., MacMaster University, Canada

ABSTRACT

In this paper, we study the problem of finite horizon Kalman filtering for systems involving a norm-bounded uncertain block. A new technique is presented for robust Kalman filter design. This technique involves using multiple scaling parameters which can be optimized by solving a semidefinite program. The use of optimized scaling parameters leads to an improved design. Also proposed is a recursive design method which can be applied to real-time applications.

1. INTRODUCTION

Finite horizon Kalman filters, including recursive least-squares filters as a special case, are widely used in signal processing applications. Compared with infinite horizon Kalman filters, the finite horizon ones can offer a better transient performance, which is an important property for applications where signals are non-stationary.

One of the problems with Kalman filters, which has been well recognized now, is that they can be sensitive to system data, or in another word, they may lack robustness. A typical phenomenon is that the performance of the filter, although being optimal for a "nominal" system, may deteriorate very quickly as the system data drift; see, e.g., [4]. This is of course not acceptable for applications where a good system model is hard to obtain or the system drifts. Motivated by this problem, a number of papers have attempted to generalize the classical Kalman filter to systems involving a norm-bounded uncertain block; see [6, 3, 4, 5, 2]. Note that norm-bounded blocks are used to represent inaccuracies in the system model. The resulting filters are often called robust Kalman filters.

The design of robust Kalman filters faces a major obstacle in comparison with the classical Kalman filters. There are two prevailing properties possessed by classical finite horizon Kalman filters. First, an optimal filter at time k leads to an optimal filter at $k+1$. That is, an optimal filter at k produces a minimum state estimation error at k (in the variance sense), which is the best initial condition for the filter design at $k+1$. Secondly, the optimal filter for state estimation is also optimal for estimation of any other signal, provided it is a linear function of the state. Unfortunately, neither of the two properties carries through when the system involves uncertainties. More precisely, a filter which produces a small state estimation error at time k may worsen the state estimation at time $k+1$. Similarly, a filter which minimizes the state estimation error may not be optimal for estimation of the signal of interest, even when it is a linear combination of the state.

A commonly used technique for robust Kalman filter design is to apply the so-called S-Procedure, which replaces the uncertainty block with a scaling parameter. This yields an upper bound for the covariance of the estimation error. Two types of scaling parameters have been used: constant and time-varying. A constant scaling parameter (τ) is used in [3, 6, 4] and is most suitable for infinite horizon or stationary filtering problems. One serious problem with using a constant scaling parameter is that the entailed conservatism can aggregate quickly as time evolves and may lead to a very poor estimator. Time-varying scaling parameters (τ_k) are more flexible, and if they are carefully chosen, the amount of conservatism can be reduced. Two papers have used time-varying scaling parameters. In [5], a simple formula is given but the scaling parameter is not optimized in any way. In [2], the scaling parameter is chosen using a semidefinite program. However, as we shall reveal later, the scaling parameter obtained at time k using [2] may lead to a poor estimation at future times. Also, the semidefinite program to be solved in [2] is quite cumbersome.

In this paper, we intend to carry out some deeper study on finite horizon Kalman filtering for systems involving a norm-bounded uncertain block. Our focus will be on how to choose scaling parameters. A summary of our results is given below.

- We show that optimal scaling parameters for time k may lead to poor estimation at future times. Subsequently, two types of scaling parameters are suggested: one optimal for time k, and one used for the future. In fact, at each time k, all the scaling parameters $\tau_0, \cdots, \tau_k$ need to be re-optimized.
- The design of the estimator has the following separation properties:
  - The covariance of the estimation error at $k+1$ depends only on the scaling parameters $\tau_0, \cdots, \tau_k$ and the system data, not on other parameters in

0-7803-5250-5/99/$10.00

the filter. Thus the scaling parameters can be optimized first. In particular, we note that they depend on the signal to be estimated.

  - Once the scaling parameters are determined, an optimal filter can be generated using an algebraic formula. In particular, we note that the optimal filter does not explicitly depend on the signal to be estimated. Implicit dependence happens only through the scaling parameters.

- We show that optimal scaling parameters can be computed using a semidefinite program. The size of the program is moderate and grows at the rate k. An suboptimal scheme is also given which requires a constant amount of computation at each k.

2. COVARIANCE ANALYSIS

Consider the following uncertain system:

$$\begin{array}{rcl} x_{k+1} & = & (A_k + H_k F_k E_k) x_k + B_k w_k \\ z_k & = & C_k x_k \end{array} \quad (1)$$

where $x_k \in \mathcal{R}^n$ is the state, $z_k \in \mathcal{R}^p$ is a linear combination of x_k, $A_k \in \mathcal{R}^{n\times n}$, $H_k \in \mathcal{R}^{n\times i}$, $E_k \in \mathcal{R}^{j\times n}$, $B_k \in \mathcal{R}^{n\times m}$ and $C_k \in \mathcal{R}^{p\times n}$ are given matrices, $F_k \in \mathcal{R}^{i\times j}$ represents norm-bounded time-varying uncertainty, i.e.,

$$F_k F_k^t \leq I, \ \forall k = 0, 1, \cdots \quad (2)$$

w_k and x_0 are zero-mean, independent and satisfy the following second order statistics:

$$\mathcal{E}(w_k w_l^t) = \begin{cases} I & \text{if } k = l \\ 0 & \text{otherwise} \end{cases}; \quad \mathcal{E}(x_0 x_0^t) = \Sigma_0 > 0 \quad (3)$$

Without loss of generality, $E_k \neq 0$ for all k. To assure that the order of the system is not degenerate, we further assume

$$\text{rank}[A_k \ H_k \ B_k] = n, \ \forall k \quad (4)$$

Denote by Σ_k and $\Sigma_{z,k} = C_k \Sigma_k C_k^t$ the covariance matrices of x_k and z_k, respectively. The *(worst-case) covariance analysis problem* is as follows: Given $T \geq 0$, determine the worst-case $\Sigma_{z,T+1}$, i.e.,

$$L_{T+1} = \max\{L(\Sigma_{z,T+1}) : F_k F_k^t \leq I, 0 \leq k \leq T\} \quad (5)$$

where $L(\Sigma)$ is any given linear function of Σ. In particular, it is common to choose $L(\Sigma) = \text{trace}(\Sigma)$.

We first introduce the so-called S-Procedure (see [1]):

Lemma 2.1 *Given $M, A, \Sigma \in \mathcal{R}^{n\times n}$, $H \in \mathcal{R}^{n\times i}$, and $E \in \mathcal{R}^{j\times n}$ with $\Sigma = \Sigma^t > 0$, the following inequality holds*

$$M - (A + HFE)\Sigma(A + HFE)^t \geq 0, \ \forall F \in \mathcal{R}^{i\times j}, FF^t \leq I \quad (6)$$

if and only if there exists $\tau \geq 0$ such that

$$\begin{bmatrix} M - \tau^{-1} H H^t & A \\ A^t & \Sigma^{-1} - \tau E^t E \end{bmatrix} \geq 0 \quad (7)$$

or equivalently,

$$\begin{bmatrix} M & A & H \\ A^t & \Sigma^{-1} - \tau E^t E & 0 \\ H^t & 0 & \tau I \end{bmatrix} \geq 0 \quad (8)$$

Next, we give a solution to the covariance analysis problem for the case $T = 0$.

Theorem 2.1 *Define*

$$\Sigma_1(\tau_0) = A_0 S_0 A_0^t + B_0 B_0^t + \tau_0^{-1} H_0 H_0^t, \ \tau_0 \in \mathbf{R} \quad (9)$$

where

$$S_0 = \Sigma_0 + \Sigma_0 E_0^t (\tau_0^{-1} I - E_0 \Sigma_0 E_0^t)^{-1} E_0 \Sigma_0 \quad (10)$$

Then,

$$\Sigma_1(\tau_0) \geq \Sigma_1, \ \Sigma_1(\tau_0) > 0, \ \forall\, 0 < \tau_0 < \|E_0 \Sigma_0 E_0^t\|^{-1} \quad (11)$$

Also,

$$L_1 = \inf\{L(C_1 \Sigma_1(\tau_0) C_1^t) : 0 < \tau_0 < \|E_0 \Sigma_0 E_0^t\|^{-1}\} \quad (12)$$

Further, the optimal τ_0 can be found by solving the following semidefinite program:

$$\begin{array}{ll} L_1 & = \min L(C_1 X C_1^t) \\ \text{s.t.} & \begin{bmatrix} X - B_0 B_0^t & A_0 & H_0 \\ A_0^t & \Sigma_0^{-1} - \tau_0 E_0^t E_0 & 0 \\ H_0^t & 0 & \tau_0 I \end{bmatrix} \geq 0 \\ & X = X^t, \tau_0 \geq 0 \end{array} \quad (13)$$

Proof: Follows from Lemma 2.1. Details are omitted. ■

Returning to the problem in (5) for $T > 0$ where more than one F_k terms are involved, it is expected that they will be replaced by additional scaling parameters τ_k to compute L_{T+1}. This can indeed be done except that for $T > 0$ an upper bound $\bar{L}_{T+1}$ for L_{T+1} yields. Nevertheless, this bound can be solved via semidefinite programming. This is detailed as follows:

Theorem 2.2 *Denote $\tau = [\tau_0, \cdots, \tau_T]$ and define*

$$\begin{array}{rcl} \Sigma_0(\tau) & = & \Sigma_0 \\ \Sigma_{k+1}(\tau) & = & B_k B_k^t + \tau_k^{-1} H_k H_k^t + A_k S_k(\tau) A_k^t \end{array} \quad (14)$$

for $k = 0, 1, \cdots, T$, where

$$S_k(\tau) = \Sigma_k(\tau) + \Sigma_k(\tau) E_k^t (\tau_k^{-1} I - E_k \Sigma_k(\tau) E_k^t)^{-1} E_k \Sigma_k(\tau) \quad (15)$$

with

$$S_k^{-1}(\tau) = \Sigma_k^{-1}(\tau) - \tau_k E_k^t E_k \quad (16)$$

Also define

$$\Omega = \{\tau : 0 < \tau_k < \|E_k \Sigma_k(\tau) E_k^t\|^{-1},\ k = 0, \cdots, T\} \tag{17}$$

Then,

$$\Sigma_{k+1}(\tau) \geq \Sigma_{k+1},\ \Sigma_{k+1}(\tau) > 0,\ \forall \tau \in \Omega,\ 0 \leq k \leq T \tag{18}$$

Next, an upper bound for L_{T+1} is given by

$$\bar{L}_{T+1} = \inf_{\tau \in \Omega} L(C_{T+1} \Sigma_{T+1}(\tau) C_{T+1}^t) \tag{19}$$

Further, the optimum above can be found by solving the following semi-definite program:

$$\begin{aligned} \bar{L}_{T+1} = \min \quad & L(C_{T+1} X C_{T+1}^t) \\ \text{s.t.} \quad & \begin{bmatrix} X & U_{T+1} \\ U_{T+1}^t & \Pi_{T+1} \end{bmatrix} \geq 0 \\ & X = X^t,\ \tau \geq 0 \end{aligned} \tag{20}$$

where U_T and Π_T are defined recursively:

$$\begin{cases} U_0 &= I; \\ \Pi_0 &= \Sigma_0^{-1}; \\ \Pi_{k+1} &= \text{diag}\,\{\Pi_k - \tau_k U_k^t E_k^t E_k U_k,\ \tau_k I_i,\ I_m\} \\ U_{k+1} &= [A_k U_k\ H_k\ B_k],\quad k \geq 0 \end{cases} \tag{21}$$

Proof: Omitted due to page limit. ∎

Remark 2.1 *Theorem 2.2 suggests that τ needs to be recomputed as T changes. This is indeed the case. In fact, we will show in Section 6 that an optimal τ at a given time T may not be optimal at a different time. Because of this property, we will denote the optimal τ_k at time T by $\tau_{T,k}$, $k = 0, \cdots, T$ whenever necessary.*

Remark 2.2 *In Theorem 2.2, we have assumed nonsingularity of Σ_0. If Σ_0 is singular, we can always decompose it into $U_0 \Pi_0^{-1} U_0^t$ for some U_0 and Π_0. With these U_0 and Π_0, the recursion in (21) will still be valid.*

3. ROBUST FILTER DESIGN: PROBLEM STATEMENT

We extend the system (1) to the following:

$$\begin{aligned} x_{k+1} &= (A_k + H_{1,k} F_k E_k) x_k + B_k w_k \\ y_k &= (C_{2,k} + H_{2,k} F_k E_k) x_k + v_k \\ z_k &= C_{1,k} x_k \end{aligned} \tag{22}$$

where $y_k \in \mathcal{R}^r$ is a measured output, $C_{2,k} \in \mathbf{R}^{r \times n}$, $H_{2,k} \in \mathbf{R}^{r \times i}$, v_k is a zero-mean measurement noise, which is independent of w_k, and with statistics

$$\mathcal{E}(v_k v_l^t) = \begin{cases} I & \text{if } k = l \\ 0 & \text{otherwise} \end{cases}; \tag{23}$$

The other matrices are defined accordingly. In the design problem, z_k is a linear combination of x_k to be estimated. Similar to (4), it is assumed that

$$\text{rank} \begin{bmatrix} A_k & H_{1,k} & B_k \\ C_{2,k} & H_{2,k} & 0 \end{bmatrix} = n + r,\ \forall\, k \tag{24}$$

The robust linear filter is of the form:

$$\begin{aligned} \hat{x}_{k+1} &= \hat{A}_k \hat{x}_k + \hat{B}_k (y_k - C_{2,k} \hat{x}_k) \\ \hat{x}_0 &= 0 \\ \hat{z}_k &= C_{1,k} \hat{x}_k \end{aligned} \tag{25}$$

Note that the use of the same $C_{1,k}$ and $C_{2,k}$ does not lose any generality.

Given the filter above, the augmented system involving x_k and $\hat{x}_k$ is given by

$$\begin{aligned} \tilde{x}_{k+1} &= (\tilde{A}_k + \tilde{H}_k F_k \tilde{E}_k) \tilde{x}_k + \tilde{B}_k \tilde{w}_k \\ e_k &= \tilde{C}_k \tilde{x}_k \end{aligned} \tag{26}$$

where e_k is the estimation error and

$$\begin{aligned} &\tilde{x}_k = \begin{bmatrix} x_k \\ \hat{x}_k \end{bmatrix},\ \tilde{w}_k = \begin{bmatrix} w_k \\ v_k \end{bmatrix}, \\ &\tilde{A}_k = \begin{bmatrix} A_k & 0 \\ \hat{B}_k C_{2,k} & \hat{A}_k - \hat{B}_k C_{2,k} \end{bmatrix}, \\ &\tilde{B}_k = \begin{bmatrix} B_k & 0 \\ 0 & \hat{B}_k \end{bmatrix},\ \tilde{H}_k = \begin{bmatrix} H_{1,k} \\ \hat{B}_k H_{2,k} \end{bmatrix}, \\ &\tilde{E}_k = [E_k\ \ 0];\ \tilde{C}_k = C_{1,k}[I\ \ -I] \end{aligned}$$

We will denote by $\tilde{\Sigma}_k$, $\Sigma_{x,k}$ and $\Sigma_{e,k}$ the covariance matrices of $\tilde{x}_k$, $x_k - \hat{x}_k$ and e_k.

Similar to the previous section, scaling parameters τ_k will be used to replace the uncertainty block F_k, which yields parameterized covariance matrices $\tilde{\Sigma}_k$, $\Sigma_{x,k}$ and $\Sigma_{e,k}$. With this in mind, a number of technical problems are proposed as follows:

P1 : Given $\tilde{\Sigma}_k$ and τ_k, find the optimal filter at k (i.e., $\hat{A}_k$ and $\hat{B}_k$) such that $L(\Sigma_{e,k+1})$ is minimized.

P2 : Given $\tilde{\Sigma}_k$, find optimal τ_k, $\hat{A}_k$ and $\hat{B}_k$ such that $L(\Sigma_{e,k+1})$ is minimized.

P3 : Given T, Σ_0, find optimal τ_k and the optimal filter at all k, $k = 0, \cdots, T$ such that $L(\Sigma_{e,T+1})$ is minimized.

Obviously, our aim is to solve P3 while P1 and P2 are the immediate steps.

4. ROBUST FILTER DESIGN: SOLUTIONS

Solutions to P1-P3 are given in this section.

Problem P1

Theorem 4.1 *Suppose*

$$\tilde{\Sigma}_k = \begin{bmatrix} \Sigma_{x,k} + \Sigma_{2,k} & \Sigma_{2,k} \\ \Sigma_{2,k} & \Sigma_{2,k} \end{bmatrix}, \ \Sigma_{x,k} > 0, \Sigma_{2,k} \geq 0 \tag{27}$$

(which holds at $k = 0$) and $0\tau_k < \|E_k \Sigma_{1,k} E_k^t\|^{-1}$, where $\Sigma_{1,k}$ denotes $\Sigma_{x,k} + \Sigma_{2,k}$. Then the optimal solution to Problem 1 is given as follows:

$$\hat{A}_k = A_k + \Delta_k \tag{28}$$

$$\Delta_k = (A_k - \hat{B}_k C_{2,k})\Sigma_{x,k} E_k^t V_k E_k \tag{29}$$

$$\hat{B}_k = (\tau_k^{-1} H_{1,k} H_{2,k}^t + A_k S_k C_{2,k}^t) \cdot (I + \tau_k^{-1} H_{2,k} H_{2,k}^t + C_{2,k} S_k C_{2,k}^t)^{-1} \tag{30}$$

where

$$V_k = (\tau_k^{-1} I - E_k \Sigma_{x,k} E_k^t)^{-1} > 0 \tag{31}$$

$$S_k = \Sigma_{x,k} + \Sigma_{x,k} E_k^t V_k E_k \Sigma_{x,k} \tag{32}$$

In particular, the optimal filter is independent of $C_{1,k+1}$ (or the signal to be estimated). Further, the optimal filter given above preserves the structure in (27), i.e.,

$$\tilde{\Sigma}_{k+1}(\tau_k) = \begin{bmatrix} \Sigma_{x,k+1}(\tau_k) + \Sigma_{2,k+1}(\tau_k) & \Sigma_{2,k+1}(\tau_k) \\ \Sigma_{2,k+1}(\tau_k) & \Sigma_{2,k+1}(\tau_k) \end{bmatrix} \tag{33}$$

with

$$\Sigma_{x,k+1}(\tau_k) = B_k B_k^t + \tau_k^{-1} H_{1,k} H_{1,k}^t + A_k S_k A_k^t - Z_k^t \Xi_k^{-1} Z_k \tag{34}$$

and

$$\Sigma_{x,k+1}(\tau_k) + \Sigma_{2,k+1}(\tau_k) = B_k B_k^t + \tau_k^{-1} H_{1,k} H_{1,k}^t + A_k (\Sigma_{1,k}^{-1} - \tau_k E_k^t E_k)^{-1} A_k^t \tag{35}$$

where

$$\Xi_k = I + \tau_k^{-1} H_{2,k} H_{2,k}^t + C_{2,k} S_k C_{2,k}^t \tag{36}$$

$$Z_k = \tau_k^{-1} H_{2,k} H_{1,k}^t + C_{2,k} S_k A_k^t \tag{37}$$

Finally, we have

$$\Sigma_{x,k+1}(\tau_k) > 0, \ \Sigma_{x,k+1}(\tau_k) \geq \Sigma_{x,k+1}, \quad \forall\, 0 < \tau_k < \|E_k \Sigma_{1,k} E_k^t\|^{-1} \tag{38}$$

Proof: Omitted due to page limit. ■

Problem P2

Theorem 4.2 *Under (27), the optimal solution to Problem 2 is given as in Theorem 4.1 with the optimal τ_k solving the following semidefinite program:*

$$\begin{aligned} &\min \ L(C_{1,k+1} X C_{1,k+1}^t) \\ &\text{s.t.} \begin{bmatrix} X - B_k B_k^t & A_k & H_{1,k} \\ A_k^t & S_k^{-1} + C_{2,k}^t C_{2,k} & C_{2,k}^t H_{2,k} \\ H_{1,k}^t & H_{2,k}^t C_{2,k} & \tau_k I + H_{2,k}^t H_{2,k} \end{bmatrix} \geq 0 \\ &X = X^t, \ 0 \leq \tau_k \leq \|E_k^t \Sigma_{1,k} E_k\|^{-1} \end{aligned} \tag{39}$$

where S_k is given in (32) with

$$S_k^{-1} = \Sigma_{x,k}^{-1} - \tau_k E_k^t E_k \tag{40}$$

Proof: Omitted due to page limit. ■

Problem P3

Before we give the solution to P3, several key observations about Theorem 4.2 are needed:

- First, only $\Sigma_{x,k}$, rather the whole $\tilde{\Sigma}_k$, is directly required for the filter design at time k. However, the term $\Sigma_{2,k}$ is used in constraining the range for τ_k.
- The optimal τ_k is solved independently of the optimal $\hat{A}_k$ and $\hat{B}_k$, although the latter depends on τ_k.
- The optimal X in (39) is indeed the optimal $\Sigma_{x,k+1}$. Further, if $\Sigma_{x,k}$ is replaced with any of its upper bound, the resulting optimal $\Sigma_{x,k+1}$ will be worsened.

These observations, together with the results in Section 2, lead us to the main result of the paper:

Theorem 4.3 *Let $\Sigma_0 > 0$ and $T \geq 0$. Denote*

$$\tau = [\tau_0, \cdots, \tau_T]$$

$$L_{e,k} = C_{1,k} \Sigma_{x,k} C_{1,k}^t$$

Define $\Sigma_{x,0}(\tau) = \Sigma_0$. Let $\Sigma_{x,k+1}(\tau)$ and $\Sigma_{2,k+1}(\tau)$ be given as in (31)-(32), (34)-(37) for $k = 0, 1, \cdots$, except that $\Sigma_{x,k}$, $\Sigma_{2,k}$ and $\Sigma_{1,k}$ are replaced by $\Sigma_{x,k}(\tau)$, $\Sigma_{2,k}(\tau)$ and $\Sigma_{1,k}(\tau)$. Then, an upper bound for $L_{e,T+1}$ is given by

$$\bar{L}_{e,T+1} = \inf_{\tau \in \Omega_x} L(C_{1,T+1} \Sigma_{x,T+1}(\tau) C_{1,T+1}^t) \tag{41}$$

where

$$\Omega_x = \{\tau : 0 < \tau_k < \|E_k^t \Sigma_{1,k}(\tau) E_k\|^{-1}, k = 0, \cdots, T\} \tag{42}$$

Also define

$$\begin{cases} U_{x,0} = I; \\ \Pi_{x,0} = \Pi_{1,0} = \Sigma_0^{-1}; \\ W_{x,k} = \Pi_{x,k} - \tau_k U_{x,k}^t E_k^t E_k U_{x,k}; \\ W_{1,k} = \Pi_{1,k} - \tau_k U_{x,k}^t E_k^t E_k U_{x,k}; \\ \Pi_{x,k+1} = \operatorname{diag}\{\begin{bmatrix} W_{x,k} & 0 \\ 0 & \tau_k I \end{bmatrix} + \begin{bmatrix} U_{x,k}^t C_{2,k}^t \\ H_{2,k}^t \end{bmatrix} \cdot [C_{2,k} U_{x,k} \ \ H_{2,k}], \ I_m\} \\ \Pi_{1,k+1} = \operatorname{diag}\{W_{1,k}, \tau_k I, I_m\} \\ U_{x,k+1} = [A_k U_{x,k} \ H_{1,k} \ B_k] \end{cases} \tag{43}$$

Note that $\Pi_{x,k}$ is affine in τ. Then the optimal $\bar{L}_{e,T+1}$ can be found by solving the following semi-definite program:

$$\begin{aligned} \bar{L}_{e,T+1} = \min \ & L(C_{1,T+1} X C_{1,T+1}^t) \\ \text{s.t.} \ & \begin{bmatrix} X & U_{x,T+1} \\ U_{x,T+1}^t & \Pi_{x,T+1} \end{bmatrix} \geq 0 \\ & W_{1,k} \geq 0, \ k = 0, \cdots, T \\ & X = X^t, \ \tau \geq 0 \end{aligned} \tag{44}$$

Once the optimal τ is found, the optimal filter at time T is given as in Theorem 4.1, with $\Sigma_{x,T} = \Sigma_{x,T}(\tau)$.

Proof: Omitted due to page limit. ■

Remark 4.1 *Note that the optimal τ_k for each T may be different. Using a fixed τ_k may lead to conservative designs. But the optimal $\hat{A}_T$ and $\hat{B}_T$ does not explicitly depend on past filters, i.e., they depend on the optimal τ and the system data at T.*

5. RECURSIVE ROBUST FILTER DESIGN

There is one unpleasant feature about the solution in Theorem 4.3. That is, the size of the semidefinite program in (44) grows linearly in k. To avoid this, we propose a suboptimal solution, i.e., a recursive method which optimizes only a fixed number of most recent scaling parameters. The motivation for this approximate solution stems from a simple fact about Kalman filtering that the contribution of the initial covariance Σ_0 to the estimation error at time T decays as time evolves, provided that the augmented system (26) is asymptotically stable. The recursive method involves solving a semidefinite program of a constant size. Therefore, it is suitable for real-time applications where the information of the system dynamics (i.e., A_k, B_k, etc.) may not be available *a priori*.

The recursive algorithm given below is simply modified from Theorem 4.3.

Step 1: Let $N+1$ be the *window size* for recursion, $N \geq 0$. For $0 \leq T \leq N$, apply (43) and (44);

Step 2: For $T > N$, still apply (43) and (44) but replace the constraint $W_{x,k} \geq 0$, $\forall k = 0, \cdots, T$ by $W_{x,k} \geq 0$, $\forall k = T-N, \cdots, T$ and reinitialize $U_{x,T-N} = I$ and $\Pi_{x,T-N} = (\Sigma^*_{T-N})^{-1}$, where Σ^*_{T-N} is the optimal $\Sigma_{T-N}(\tau)$ determined at $T-N$.

6. EXAMPLE

To illustrate the results in this paper, we consider the following example, which has been used as a "benchmark" in [6, 5, 2]:

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 0 & -0.5 \\ 1 & 1+0.3\delta_k \end{bmatrix} x_k + \begin{bmatrix} -6 \\ 1 \end{bmatrix} w_k \\ y_k &= [-100 \ \ 10] x_k + v_k \\ z_k &= [1 \ \ 0] x_k \end{aligned} \tag{45}$$

where $|\delta_k| \leq 1$ is the uncertainty. We assume that the initial state covariance matrix $\Sigma_0 = I$.

To match the system description in (22), the uncertain term is represented by the matrices

$$H_1 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \ H_2 = 0, \ E = [0 \ \ 0.03]. \tag{46}$$

Stationary filters are designed in [6, 5, 2] to compare with the so-called "nominal" Kalman filter where the uncertainty is ignored. An infinite-horizon filter is used in [6] with guaranteed stability, which gives a great improvement over the nominal design. The design in [5] is based on finite-horizon. In our setting, this design is similar to the recursive case with window size equal to one except that the scaling parameter τ is pre-selected. The performance turns out to be superior to [6]. The design in [2] is similar to [5] except that the scaling parameter is optimized at each iteration using a semidefinite programming technique, yielding some small improvement over [5].

For comparison, three new designs are shown below.

Design 1: Recursive Design with $N = 1$

First, we design a filter using the given system data. The resulting filter turns out to be unstable. This demonstrates the inherent instability of finite-horizon designs. The intuitive reason is that the filter only aims at minimizing the cost function at each given time instant without considering its consequence in the future. This problem has been recognized by other researchers. For example, [5] solves this problem by using a fixed (conservative) scaling parameter, while in [2] the augmented covariance matrix $\tilde{\Sigma}_k$ is required to be bounded.

Alternatively, we solve the instability problem by adding an additional term to the performance cost. Indeed, we take

$$C_1 = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix} \tag{47}$$

and the performance cost is trace$\{C_1 \mathcal{E}(e_k e_k^t) C_1^t\}$. It is observed in the simulations that increasing ε can dramatically improve the stability and the steady state performance with a minor tradeoff of the initial performance.

To demonstrate various recursive designs, we select $\varepsilon = 0.2$. The corresponding filter is stable and converges to a stationary one as $k \to \infty$, and it is given by (22) with

$$\hat{A}_k = \begin{bmatrix} 0 & -0.5165 \\ 1 & 1.0362 \end{bmatrix}, \ \hat{B}_k = \begin{bmatrix} -0.003044 \\ -0.003304 \end{bmatrix} \tag{48}$$

Design 2: Recursive Design with $N = 2$

Recall that with $N = 2$, two scaling parameters, $\tau_{T,T-1}$ and $\tau_{T,T}$, are involved at each T. The first one is for estimating $\Sigma_{x,T}$ from $\Sigma_{x,T-1}$ and the second one is used to estimate $\Sigma_{x,T+1}$ and to design the filter.

For the same fix of C_1 in (47), the steady state filter has

$$\hat{A}_k = \begin{bmatrix} 0 & -0.5175 \\ 1 & 1.0386 \end{bmatrix}, \ \hat{B}_k = \begin{bmatrix} -0.002569 \\ -0.004348 \end{bmatrix} \tag{49}$$

The steady state scaling parameters are $\tau_{T,T-1} = 4.3966$, $\tau_{T,T} = 1.5101$.

Design 3: Recursive Design with $N = 3$

The filter for $N = 3$ and the same fix of C_1 is also stable and has steady state matrices

$$\hat{A}_k = \begin{bmatrix} 0 & -0.5563 \\ 1 & 1.1238 \end{bmatrix}, \ \hat{B}_k = \begin{bmatrix} -0.002613 \\ -0.004252 \end{bmatrix} \tag{50}$$

Now, the scaling parameters converge to $\tau_{T,T-2} = 4.7172$, $\tau_{T,T-1} = 4.8874$, $\tau_{T,T} = 1.6699$.

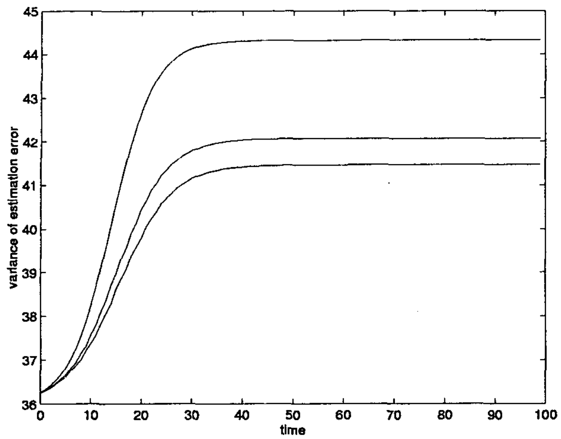

Figure 1: Performances for Recursive Robust Filter Designs. Design 1: The top curve; Design 2: The middle curve; Design 3: The bottom curve

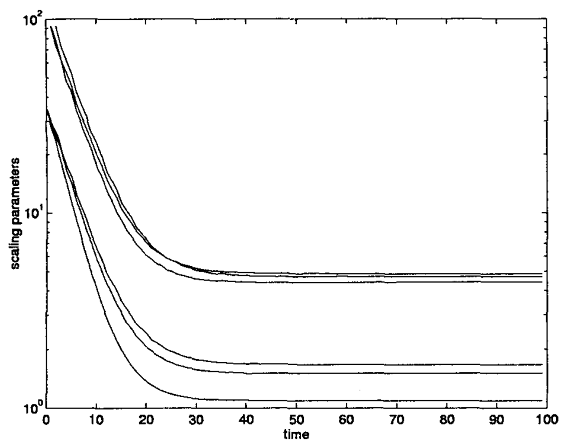

Figure 2: Scaling Parameters for Recursive Robust Filter Designs. Design 1: $\tau_{T,T}$–the bottom bottom; Design 2: $\tau_{T,T-1}$–the third curve from the top; $\tau_{T,T}$–the second curve from the bottom; Design 3: $\tau_{T,T-2}, \tau_{T,T-1}$–the two top curves; $\tau_{T,T}$–the third curve from the bottom

7. CONCLUDING REMARKS

In this paper, we have proposed a new design technique for finite horizon robust Kalman filters. This technique allows us to effectively treat systems with norm-bounded uncertainty blocks. The uncertainties are dealt with using the so-called S-Procedure, which yields a set of scaling parameters to optimize. The corresponding optimization problem is convex and can be solved either directly or via semidefinite program. Also presented is a recursive design method which is mostly suitable to applications with non-stationary processes or signals. The proposed technique gives less conservative designs in comparison with existing techniques for robust Kalman filtering. This property has been demonstrated using an example.

Filter	$\delta = -1$	$\delta = 0$	$\delta = 1$	Bound
Nominal Filter	551.2	36.0	8352.8	-
Robust Filter of [6]	64.0	61.4	64.4	98.7
Robust Filter of [5]	46.6	45.2	54.1	54.3
Robust Filter of [2]	50.8	49.4	53.5	N/A
Proposed Design 1	39.22	39.68	40.29	44.33
Proposed Design 2	38.13	38.68	39.42	42.07
Proposed Design 3	37.75	38.19	38.82	41.47

Table 1: Steady State Performance Comparison

Acknowledgement. The authors wish to thank Dr. Lihua Xie for his help in indentifying a technical error in an earlier version of the paper.

8. REFERENCES

[1] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, Philadelphia: SIAM, 1994.

[2] L. Li, Z.-Q. Luo, K. M. Wong and E. Bossé, "Semidefinite programming solutions to the robust state estimation problem with applications to multi-target tracking," preprint.

[3] I. Petersen and D. C. McFarlane, "Optimal guaranteed cost filtering for uncertain discrete-time linear systems," *Int. J. Robust and Nonlinear Control*, vol. 6, pp. 267-280, 1996.

[4] U. Shaked and C. E. de Souza, "Robust Minimum Variance Filtering," *IEEE Trans. Signal Processing*, vol. 43, no. 11, pp. 2474-2483, 1995.

[5] Y. Theodor and U. Shaked, "Robust Discrete-Time Minimum-Variance Filtering," *IEEE Trans. Signal Processing*, vol. 44, no. 2, pp. 181-189, 1996.

[6] L. Xie, Y. C. Soh and C. E. de Souza, "Robust Kalman filtering for uncertain discrete-time systems," *IEEE Trans. Auto. Contr.*, vol. 39, no. 6, pp. 1310-1314, 1994.