Proceedings of the 8th IEEE Mediterranean Conference
on Control and Automation (MED 2000)
Rio, Patras, GREECE 17-19 July, 2000

STATE ESTIMATION WITH UNCERTAIN PARAMETRIC MODELS

ALI H. SAYED *

* Department of Electrical Engineering
University of California
Los Angeles, CA 90095
sayed@ee.ucla.edu
www.ee.ucla.edu/asl

Abstract: This paper develops robust estimation algorithms for state-space models that are subject to bounded parametric uncertainties. Compared with existing robust filters, the new filters perform data regularization rather than de-regularization and they do not require existence conditions. The resulting filter structures also turn out to be similar to various (time- and measurement-update, prediction, and information) forms of the Kalman filter, albeit ones that operate on corrected parameters rather than on the given nominal parameters.

Keywords: Estimation, parametric uncertainty, set-valued estimation, Kalman filtering, $\mathcal{H}_\infty$ filtering, guaranteed-cost design, steady-state filter, regularized least-squares.

1. INTRODUCTION

A central premise in the Kalman filter theory is that the underlying state-space model is accurate (see, e.g., [1]). When this assumption is violated, the performance of the filter can deteriorate significantly. This filter sensitivity to modeling errors has led over the years to the development of several robust filters; robust in the sense that they attempt to limit, in certain ways, the effect of model uncertainties on the overall filter performance. Three of the most distinguished approaches to state-space estimation in this regard are $\mathcal{H}_\infty$ filtering, set-valued estimation, and guaranteed-cost designs.

One limitation of $\mathcal{H}_\infty$ designs for on-line (i.e., recursive) filter operation is that they require continuous testing of certain existence conditions (see, e.g., [2]). When a condition fails at any particular iteration, the desired $\mathcal{H}_\infty$ performance is lost and the filter can diverge. In addition, the design of $\mathcal{H}_\infty$ filters requires accurate state-space models.

In robust set-valued estimation, one attempts to construct ellipsoids around state estimates that are consistent with the observations (see, e.g., [3,4,5]). Here again one is faced with existence conditions that can at times be violated more often than in a typical $\mathcal{H}_\infty$ implementation.

In guaranteed-cost designs, one attempts to construct state space estimators that lead to bounded error variances (see, e.g., [5,6]). The results that are available in this framework turn out to involve an observer structure that is similar to Kalman filtering. The arguments and the derivation, however, are limited to time-invariant quadratically-stable nominal models and hold for steady-state operation (i.e., only over infinite-time horizons).

In this paper we introduce a procedure for robust state-space estimation in the presence of modeling uncertainties. Compared with the standard Kalman filter, which is known to minimize the regularized residual norm at each iteration, the new filters minimize the worst-possible regularized residual norm over the class of admissible uncertainties. In addition, the framework proposed herein distinguishes itself from earlier robust designs in the following ways: a) it performs data regularization rather than de-regularization. In this way, no existence conditions are required — they are automatically satisfied just like standard Kalman filtering, b) it applies to finite-horizon problems and to time-variant state-space models, c) it provides some geometric insights into the nature of the solution, and d) it exhibits promising performance when compared with $\mathcal{H}_\infty$, set-valued state estimation, and guaranteed-cost filters.

We start our exposition by formulating a least-squares problem for uncertain data. Once this is done, we shall then focus on the state-space estimation problem.

[1] This work was partially supported by NSF ECS-9820765.

2. UNCERTAIN LEAST-SQUARES

As is well-known, many estimation techniques rely on solving regularized least-squares problems of the form

$$\min_x \left[x^T Qx + (Ax-b)^T W(Ax-b)\right] \qquad (1)$$

where $x^T Qx$ is a regularization term with $Q = Q^T > 0$, and $W = W^T \geq 0$ is a weighting matrix. The unknown vector x is $n-$dimensional, while A is $N \times n$ and b is $N \times 1$. Both A and b are assumed known with A called the data matrix and b the measurement vector. The solution of (1) is

$$\hat{x} = [Q + A^T WA]^{-1} A^T Wb. \qquad (2)$$

In practice, the nominal data $\{A, b\}$ are often subject to uncertainties. Such errors can degrade the performance of the estimator (2) — see [7]. This motivated us to introduce in [8] a generalization of (1) that can account for uncertainties in $\{A, b\}$. Thus let $J(x, y)$ denote a cost function of the form $J(x, y) = x^T Qx + R(x, y)$, where

$$R(x,y) \stackrel{\Delta}{=} \left(Ax - b + Hy\right)^T W \left(Ax - b + Hy\right)$$

Here H is an $N \times m$ known matrix and y is an $m \times 1$ unknown perturbation vector. Comparing the expression for $R(x, y)$ with the term $(Ax - b)^T W(Ax - b)$ that appears in (1), we see that we are representing possible sources of uncertainties in A and b by the additional term Hy. The incorporation of the matrix H provides the designer with the freedom to restrict the uncertainty y to certain range spaces.

Now, while y itself is not known, we shall assume that we know a bound on its Euclidean norm, say $||y|| \leq \phi(x)$, for some known nonnegative function $\phi(x)$. Observe that the bound on y is allowed to depend on x. Consider then the problem of solving

$$\hat{x} = \arg\min_x \max_{||y|| \leq \phi(x)} J(x,y) \qquad (3)$$

This statement can be interpreted as a constrained two-player game problem, with the designer trying to pick an estimate $\hat{x}$ that minimizes the cost while the opponent $\{y\}$ tries to maximize the cost. The game problem is constrained since it imposes a limit (through $\phi(x)$) on how large (or how damaging) the opponent can be. Observe further that the strength of the opponent can vary with the choice of x. We shall assume that H and $\phi(x)$ are not identically zero, i.e., $H \neq 0$ and $\phi(\cdot) \neq 0$, since if either is zero, then the game problem (3) trivializes to the standard regularized least-squares problem (1).

In the sequel we shall focus on the following specialization of (3),

$$\min_x \max_{\substack{\delta A \\ \delta b}} \left[x^T Qx + \Big((A + \delta A)x - (b + \delta b)\Big)^T W \Big(\cdot\Big)\right] \qquad (4)$$

where the compact notation $(\cdot)$ refers to the term $(A + \delta A)x - (b + \delta b)$. Here $\{\delta A\}$ denotes an $N \times n$ perturbation to A, δb denotes an $N \times 1$ perturbation to b, and $\{\delta A, \delta b\}$ satisfy a model of the form

$$\begin{bmatrix} \delta A & \delta b \end{bmatrix} = H\Delta \begin{bmatrix} E_a & E_b \end{bmatrix} \qquad (5)$$

where Δ is an arbitrary contraction, $||\Delta|| \leq 1$, and $\{H, E_a, E_b\}$ are known quantities of appropriate dimensions (e.g., E_b is a column vector). Perturbation models of this form are common in robust filtering and control and can arise from tolerance specifications on physical parameters. In order to verify that (4) is a special case of (3), simply let $y = \Delta(E_a x - E_b)$ and $\phi(x) = ||E_a x - E_b||$. One can also handle the case in which the uncertainties $\{\delta A, \delta b\}$ in (4) are bounded, say $||\delta A|| \leq \eta$ and $||\delta b|| \leq \eta_b$ for some nonnegative scalars $\{\eta, \eta_b\}$, instead of (5). In this case, problem (4) would be a special case of (3) with the choices $H = I$ and $\phi(x) = \eta||x|| + \eta_b$.

2.1 Solution

Let λ be a nonnegative scalar parameter and define the following functions of λ:

$$W(\lambda) \stackrel{\Delta}{=} W + WH\left(\lambda I - H^T WH\right)^{\dagger} H^T W \qquad (6)$$

$$Q(\lambda) \stackrel{\Delta}{=} Q + \lambda E_a^T E_a \qquad (7)$$

$$x(\lambda) \stackrel{\Delta}{=} \left[Q(\lambda) + A^T W(\lambda) A\right]^{-1} \left[A^T W(\lambda) b + \lambda E_a^T E_b\right] \qquad (8)$$

$$\begin{aligned} G(\lambda) \stackrel{\Delta}{=} {} & x^T(\lambda) Q x(\lambda) \; + \; \lambda ||E_a x(\lambda) - E_b||^2 \; + \\ & + [Ax(\lambda) - b]^T W(\lambda)[Ax(\lambda) - b] \end{aligned} \qquad (9)$$

where the notation $X^{\dagger}$ denotes the pseudo-inverse of X. The following result is from [8].

<u>THEOREM</u> *Let*

$$\hat{\lambda} = \arg \min_{\lambda \geq ||H^T WH||} G(\lambda). \qquad (10)$$

Then problem (4)–(5) has a unique solution $\hat{x}$ that is given by

$$\hat{x} = \left[\widehat{Q} + A^T \widehat{W} A\right]^{-1} \left[A^T \widehat{W} b + \hat{\lambda} E_a^T E_b\right] = x(\hat{\lambda}) \qquad (11)$$

where

$$\widehat{Q} \stackrel{\Delta}{=} Q + \hat{\lambda} E_a^T E_a \equiv Q(\hat{\lambda}), \qquad (12)$$

$$\widehat{W} \stackrel{\Delta}{=} W + WH(\hat{\lambda} I - H^T WH)^{\dagger} H^T W \equiv W(\hat{\lambda}). \qquad (13)$$

$\diamond$

We shall denote the lower bound on λ in (10) by $\lambda_l = ||H^T WH||$. Then note that $W(\lambda) \geq 0$ for any $\lambda \in [\lambda_l, \infty)$, so that $G(\lambda)$ is nonnegative for all such λ. The function $G(\lambda)$ can also be shown to have a unique global minimum in the interval (λ_l, ∞).

Compared with the solution (2) of the standard regularized least-squares problem (1), we see that the expression for $\hat{x}$

in (11) is distinct in two important ways: a) the weighting matrices $\{Q, W\}$ need to be replaced by corrected versions $\{\widehat{Q}, \widehat{W}\}$ and b) the right-hand side of (11) contains an additional term that is equal to $\hat{\lambda} E_a^T E_b$. This means that, with $\hat{\lambda}$ given, the $\hat{x}$ in (11) can be interpreted as the solution to a regularized least-squares problem with *cross-coupling* between x and unity.

Finally, we should mention that in the state-space context further ahead, the matrix W will be positive-definite so that $W(\lambda)$ itself will always be positive-definite. Therefore, if we restrict the minimization in (10) to the open interval (λ_l, ∞), then the pseudo-inverse operation in (6) can be replaced by the normal matrix inversion, so that

$$W^{-1}(\lambda) = W^{-1} - \lambda^{-1} H H^T. \tag{14}$$

3. THE KALMAN FILTER

Our objective now is to describe one way to incorporate the uncertain least-squares formulation into a Kalman filtering context. We start by reviewing the standard Kalman filter. Thus consider a state-space description of the form

$$x_{i+1} = F_i x_i + G_i u_i, \quad i \geq 0, \tag{15}$$

$$y_i = H_i x_i + v_i, \tag{16}$$

where $\{x_0, u_i, v_i\}$ are uncorrelated zero-mean random variables with variances

$$E\left(\begin{bmatrix} x_0 \\ u_i \\ v_i \end{bmatrix} \begin{bmatrix} x_0 \\ u_j \\ v_j \end{bmatrix}^T \right) = \begin{bmatrix} \Pi_0 & 0 & 0 \\ 0 & Q_i \delta_{ij} & 0 \\ 0 & 0 & R_i \delta_{ij} \end{bmatrix} \tag{17}$$

that satisfy $\Pi_0 > 0$, $R_i > 0$, and $Q_i > 0$.

The Kalman filter provides the optimal linear least-mean-squares estimate of the state variable given prior observations. More specifically, introduce the following predicted and filtered estimates:

$$\hat{x}_i \stackrel{\Delta}{=} \text{l.l.m.s. estimate of } x_i \text{ given } \{y_0, \ldots, y_{i-1}\}$$

$$\hat{x}_{i|i} \stackrel{\Delta}{=} \text{l.l.m.s. estimate of } x_i \text{ given } \{y_0, \ldots, y_i\}$$

and the corresponding error variances P_i and $P_{i|i}$, respectively. Then $\{\hat{x}_i, \hat{x}_{i|i}\}$ can be constructed recursively as follows (see, e.g., [1]):

$$\hat{x}_{i+1} = F_i \hat{x}_{i|i}, \quad i \geq 0 \tag{18}$$

$$\hat{x}_{i+1|i+1} = \hat{x}_{i+1} + P_{i+1|i+1} H_{i+1}^T R_{i+1}^{-1} e_{i+1} \tag{19}$$

$$e_{i+1} = y_{i+1} - H_{i+1} \hat{x}_{i+1} \tag{20}$$

$$P_{i+1} = F_i P_{i|i} F_i^T + G_i Q_i G_i^T \tag{21}$$

$$P_{i+1|i+1} = P_{i+1} - P_{i+1} H_{i+1}^T R_{e,i+1}^{-1} H_{i+1} P_{i+1} \tag{22}$$

$$R_{e,i+1} = R_{i+1} + H_{i+1} P_{i+1} H_{i+1}^T \tag{23}$$

with initial conditions

$$\hat{x}_{0|0} = P_{0|0}^{-1} H_0^T R_0^{-1} y_0, \ P_{0|0} = (\Pi_0^{-1} + H_0^T R_0^{-1} H_0)^{-1}$$

Equations (18)–(23) are known collectively as the time- and measurement-update form of the Kalman filter. It can be further seen from these equations that the following prediction form of the Kalman filter also holds:

$$\hat{x}_{i+1} = F_i \hat{x}_i + K_i R_{e,i}^{-1} e_i, \quad \hat{x}_0 = 0 \tag{24}$$

$$P_{i+1} = F_i P_i F_i^T + G_i Q_i G_i^T - K_i R_{e,i}^{-1} K_i^T \tag{25}$$

$$K_i = F_i P_i H_i^T, \ \ P_0 = \Pi_0. \tag{26}$$

3.1 A Deterministic Interpretation

Each step (18)–(23) of the time- and measurement-update form of the Kalman filter admits a useful deterministic interpretation as the solution to a regularized least-squares problem, as we shall now explain (a related simplified discussion can be found in [9]). This interpretation will suggest a robust extension that will enable us to employ the result of the earlier Theorem.

Given $\{\hat{x}_{i|i}, P_{i|i} > 0, y_{i+1}\}$, consider the problem of estimating x_i again, along with u_i, by solving

$$\min_{\{x_i, u_i\}} \begin{pmatrix} (x_i - \hat{x}_{i|i})^T P_{i|i}^{-1}(\cdot) + u_i^T Q_i^{-1} u_i + \\ (y_{i+1} - H_{i+1} x_{i+1})^T R_{i+1}^{-1}(\cdot) \end{pmatrix} \tag{27}$$

If we make the substitution $x_{i+1} = F_i x_i + G_i u_i$, then the cost in (27) reduces to a regularized least-squares cost of the form (1) with the identifications

$$x \longleftarrow \text{col}\{x_i - \hat{x}_{i|i}, \ u_i\}, \quad b \longleftarrow y_{i+1} - H_{i+1} F_i \hat{x}_{i|i}$$

$$A \longleftarrow H_{i+1} \begin{bmatrix} F_i & G_i \end{bmatrix}, \ Q \longleftarrow (P_{i|i}^{-1} \oplus Q_i^{-1}), \ W \longleftarrow R_{i+1}^{-1}$$

We shall denote the minimizing arguments of (27) by $\hat{x}_{i|i+1}$ and $\hat{u}_{i|i+1}$. From the solution (2) of any such regularized least-squares problem, we obtain that $\hat{x}_{i|i+1}$ and $\hat{u}_{i|i+1}$ can be determined by solving

$$\left(\begin{bmatrix} P_{i|i}^{-1} & \\ & Q_i^{-1} \end{bmatrix} + \begin{bmatrix} F_i^T \\ G_i^T \end{bmatrix} H_{i+1}^T R_{i+1}^{-1} H_{i+1} \left[\ \cdot \ \right] \right) \begin{bmatrix} \hat{x}_{i|i+1} - \hat{x}_{i|i} \\ \hat{u}_{i|i+1} \end{bmatrix}$$

$$= \begin{bmatrix} F_i^T \\ G_i^T \end{bmatrix} H_{i+1}^T R_{i+1}^{-1} \left(y_{i+1} - H_{i+1} F_i \hat{x}_{i|i} \right)$$

If we now define the quantities (in agreement with the state-space constraint (15)):

$$\hat{x}_{i+1|i+1} \stackrel{\Delta}{=} F_i \hat{x}_{i|i+1} \ + \ G_i \hat{u}_{i|i+1}, \quad \hat{x}_{i+1} \stackrel{\Delta}{=} F_i \hat{x}_{i|i},$$

as well as $P_{i+1} = F_i P_{i|i} F_i^T + G_i Q_i G_i^T$, then the above equations collapse to the time- and measurement-update form (18)–(23) of the Kalman filter.

4. ROBUST STATE-SPACE ESTIMATION

Consider now an uncertain state-space model of the form

$$x_{i+1} = (F_i + \delta F_i) x_i + (G_i + \delta G_i) u_i, \quad i \geq 0, \tag{28}$$

$$y_i = H_i x_i + v_i, \tag{29}$$

$$\begin{bmatrix} \delta F_i & \delta G_i \end{bmatrix} = M_i \Delta_i \begin{bmatrix} E_{f,i} & E_{g,i} \end{bmatrix} \tag{30}$$

for some known matrices $\{M_i, E_{f,i}, E_{g,i}\}$ and for an arbitrary contraction Δ_i, $\|\Delta_i\| \leq 1$. Observe that for generality we are allowing the quantities $\{M_i, E_{f,i}, E_{g,i}\}$ to vary with time. Also, due to space limitations, we shall explain how to handle uncertainties in H_i elsewhere.

Now assume that at step i we are given an a priori estimate for x_i, say $\hat{x}_{i|i}$, and a positive-definite weighting matrix $P_{i|i}$. Using y_{i+1}, we propose to update the estimate of x_i from $\hat{x}_{i|i}$ to $\hat{x}_{i|i+1}$ by solving

$$\min_{\{x_i,u_i\}} \max_{\begin{pmatrix}\delta F_i \\ \delta G_i\end{pmatrix}} \begin{pmatrix} (x_i - \hat{x}_{i|i})^T P_{i|i}^{-1}(\cdot) + u_i^T Q_i^{-1} u_i + \\ (y_{i+1} - H_{i+1}x_{i+1})^T R_{i+1}^{-1}(\cdot) \end{pmatrix} \quad (31)$$

subject to (28)–(30). This problem can be seen to be the robust version of (27) in the same way that (4)–(5) is the robust version of (1). Now (31) can be written more compactly in the form (4)–(5) with the identifications:

$$x \longleftarrow \text{col}\{x_i - \hat{x}_{i|i},\ u_i\}, \quad b \longleftarrow y_{i+1} - H_{i+1}F_i\hat{x}_{i|i} \quad (32)$$

$$\delta A \longleftarrow H_{i+1}M_i\Delta_i \begin{bmatrix} E_{f,i} & E_{g,i} \end{bmatrix} \quad (33)$$

$$\delta b \longleftarrow -H_{i+1}M_i\Delta_i E_{f,i}\hat{x}_{i|i}, \quad Q \longleftarrow (P_{i|i}^{-1} \oplus Q_i^{-1}) \quad (34)$$

$$W \longleftarrow R_{i+1}^{-1},\ H \longleftarrow H_{i+1}M_i,\ E_a \longleftarrow \begin{bmatrix} E_{f,i} & E_{g,i} \end{bmatrix} \quad (35)$$

$$E_b \longleftarrow -E_{f,i}\hat{x}_{i|i},\ \Delta \longleftarrow \Delta_i,\ A \longleftarrow H_{i+1}\begin{bmatrix} F_i & G_i \end{bmatrix} \quad (36)$$

According to our earlier Theorem, the solution $\{\hat{x}_{i|i+1}, \hat{u}_{i|i+1}\}$ is then found by solving the system of equations:

$$\left(\widehat{Q} + A^T\widehat{W}A\right)\hat{x} = \left(A^T\widehat{W}b + \hat{\lambda}_i E_a^T E_b\right) \quad (37)$$

where [2]

$$\widehat{Q} = \begin{bmatrix} P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^T E_{f,i} & \hat{\lambda}_i E_{f,i}^T E_{g,i} \\ \hat{\lambda}_i E_{g,i}^T E_{f,i} & Q_i^{-1} + \hat{\lambda} E_{g,i}^T E_{g,i} \end{bmatrix} \quad (38)$$

$$\widehat{W} \triangleq \widehat{R}_{i+1}^{-1} = \left(R_{i+1} - \hat{\lambda}_i^{-1} H_{i+1}M_iM_i^TH_{i+1}^T\right)^{-1} \quad (39)$$

Moreover, $\hat{\lambda}_i$ is the minimizing argument in the interval

$$\hat{\lambda}_i > \|M_i^T H_{i+1}^T R_{i+1}^{-1} H_{i+1} M_i\| \triangleq \lambda_{l,i} \quad (40)$$

of the corresponding scalar-valued function $G(\lambda)$ in (9) constructed with the identifications (32)–(36). [The expression for $G(\lambda)$ is of course time-dependent. However, for simplicity of notation, we have not indicated this time-dependence explicitly.]

Now substituting (38)–(39) into (37), we can solve for $\{\hat{x}_{i|i+1}, \hat{u}_{i|i+1}\}$ and obtain, after some algebra, the time- and measurement-update robust algorithm listed in Table 1. The major step in the algorithm is step 3, which consists of recursions that are very similar in nature to the time- and measurement-update form of the Kalman filter (cf. Equations (18)–(23)). The main difference is that the new recursions operate on modified parameters rather than on the given nominal values. Note further from the listing in Table 1 that $\widehat{Q}_i^{-1} \geq Q_i^{-1}$ and $\widehat{R}_{i+1} \leq R_{i+1}$.

[2] Without much loss in generality, we are considering here the scenario described at the end of Sec. 2.1, viz., that the minimization of $G(\lambda)$ is performed over the open interval $(\lambda_{l,i}, \infty)$ defined in (40).

Assumed uncertain model: Eqs. (28)–(30).

Initial conditions

$$\hat{x}_{0|0} = P_{0|0}H_0^TR_0^{-1}y_0,\ P_{0|0} = \left(\Pi_0^{-1} + H_0^TR_0^{-1}H_0\right)^{-1}.$$

Step 1. If $H_{i+1}M_i = 0$, then set $\hat{\lambda}_i = 0$. Otherwise, determine $\hat{\lambda}_i$ by minimizing $G(\lambda)$ over the interval $(\lambda_{l,i}, \infty)$.

Step 2. Compute the corrected parameters:

$$\begin{aligned}
\widehat{Q}_i^{-1} &= Q_i^{-1} + \hat{\lambda}_i E_{g,i}^T\left[I + \hat{\lambda}_i E_{f,i}P_{i|i}E_{f,i}^T\right]^{-1}E_{g,i} \\
\widehat{R}_{i+1} &= R_{i+1} - \hat{\lambda}_i^{-1}H_{i+1}M_iM_i^TH_{i+1}^T \\
\widehat{P}_{i|i} &= \left(P_{i|i}^{-1} + \hat{\lambda}_i E_{f,i}^TE_{f,i}\right)^{-1} \\
&= P_{i|i} - P_{i|i}E_{f,i}^T(\hat{\lambda}_i^{-1}I + E_{f,i}P_{i|i}E_{f,i}^T)^{-1}E_{f,i}P_{i|i} \\
\widehat{G}_i &= G_i - \hat{\lambda}_iF_i\widehat{P}_{i|i}E_{f,i}^TE_{g,i} \\
\widehat{F}_i &= (F_i - \hat{\lambda}_i\widehat{G}_i\widehat{Q}_iE_{g,i}^TE_{f,i})(I - \hat{\lambda}_i\widehat{P}_{i|i}E_{f,i}^TE_{f,i})
\end{aligned}$$

If $\hat{\lambda}_i = 0$, then simply set $\widehat{Q}_i = Q_i$, $\widehat{R}_{i+1} = R_{i+1}$, $\widehat{P}_{i|i} = P_{i|i}$, $\widehat{G}_i = G_i$, and $\widehat{F}_i = F_i$.

Step 3. Now compute:

$$\begin{aligned}
\hat{x}_{i+1} &= \widehat{F}_i\hat{x}_{i|i} \\
\hat{x}_{i+1|i+1} &= \hat{x}_{i+1} + P_{i+1|i+1}H_{i+1}^T\widehat{R}_{i+1}^{-1}e_{i+1} \\
e_{i+1} &= y_{i+1} - H_{i+1}\hat{x}_{i+1} \\
P_{i+1} &= F_i\widehat{P}_{i|i}F_i^T + \widehat{G}_i\widehat{Q}_i\widehat{G}_i^T \\
P_{i+1|i+1} &= P_{i+1} - P_{i+1}H_{i+1}^TR_{e,i+1}^{-1}H_{i+1}P_{i+1} \\
R_{e,i+1} &= \widehat{R}_{i+1} + H_{i+1}P_{i+1}H_{i+1}^T
\end{aligned}$$

Table 1: Listing of the proposed robust filtering algorithm in time- and measurement-update form.

Some simple algebra will show that the recursions of Table 1 can be manipulated into an alternative so-called prediction form, which propagates the quantities $\{\hat{x}_i, P_i\}$ directly, as shown in Table 2. We should remark that the recursion for P_i in the table is not a standard Riccati recursion since the product $\widehat{G}_i\widehat{Q}_i\widehat{G}_i^T$ is also dependent on P_i.

When the F_i are invertible, the robust algorithm can also be rewritten in an alternative so-called information form that propagates the inverses of the matrices $P_{i|i}$ rather than the matrices themselves. The recursions are shown in Table 3.

4.1 Two Special Cases

The recursions in all forms can be further simplified in two special cases: $E_{g,i} = 0$ (i.e., no uncertainty in G_i) and $E_{f,i}^TE_{g,i} = 0$ (i.e., the uncertainty in G_i is orthogonal to that in F_i). In the first case ($E_{g,i} = 0$), it is easy to see that we get $\widehat{Q}_i = Q_i$, $\widehat{G}_i = G_i$, and $\widehat{F}_i = F_i(I - \hat{\lambda}_i\widehat{P}_{i|i}E_{f,i}^TE_{f,i})$. In the second case ($E_{f,i}^TE_{g,i} = 0$), we obtain the same simplifications for $\{\widehat{G}_i, \widehat{F}_i\}$ while $\widehat{Q}_i$ becomes

$$\widehat{Q}_i = \left(Q_i^{-1} + \hat{\lambda}_i E_{g,i}^T E_{g,i}\right)^{-1}.$$

In both cases, the recursion for P_i in Table 2 now becomes a standard Riccati recursion.

Assumed uncertain model. Same as in Table 1.
Initial conditions: $\hat{x}_0 = 0$, $P_0 = \Pi_0$, and $\widehat{R}_0 = R_0$.
Step 1a. Using $\{\widehat{R}_i, H_i, P_i\}$ compute $P_{i\|i}$:
$P_{i\|i} = (P_i^{-1} + H_i^T \widehat{R}_i^{-1} H_i)^{-1} = P_i - P_i H_i^T (\widehat{R}_i + H_i P_i H_i^T)^{-1} H_i P_i$
Step 1b. Determine the optimal scalar parameter $\hat{\lambda}_i$ as in step 1 of Table 1.
Step 2. Same as in Table 1.
Step 3. Now update $\{\hat{x}_i, P_i\}$ to $\{\hat{x}_{i+1}, P_{i+1}\}$ as follows:
$\hat{x}_{i+1} = \widehat{F}_i \hat{x}_i + \widehat{F}_i P_i H_i^T R_{e,i}^{-1} e_i$
$e_i = y_i - H_i \hat{x}_i$
$P_{i+1} = F_i P_i F_i^T - \overline{K}_i \overline{R}_{e,i}^{-1} \overline{K}_i^T + \widehat{G}_i \widehat{Q}_i \widehat{G}_i^T$
$\overline{K}_i = F_i P_i \overline{H}_i^T, \quad \overline{R}_{e,i} = I + \overline{H}_i P_i \overline{H}_i^T$
where $\overline{H}_i^T = \begin{bmatrix} H_i^T \widehat{R}_i^{-T/2} & \sqrt{\hat{\lambda}_i} E_{f,i}^T \end{bmatrix}$.

Table 2: Listing of the proposed robust filtering algorithm in prediction form.

Assumed uncertain model. Same as in Table 1 with the additional assumption that F_i is invertible.
Initial conditions:
$P_{0\|0}^{-1} \hat{x}_{0\|0} = H_0^T R_0^{-1} y_0, \; P_{0\|0}^{-1} = \Pi_0^{-1} + H_0^T R_0^{-1} H_0.$
Steps 1 and 2. Same as in Table 1.
Step 3. Compute:
$P_{i+1\|i+1}^{-1} \hat{x}_{i+1\|i+1} = H_{i+1}^T \widehat{R}_{i+1}^{-1} y_{i+1} + \left[\left(P_{i+1\|i+1}^{-1} - H_{i+1}^T \widehat{R}_{i+1}^{-1} H_{i+1} \right) \widehat{F}_i P_{i\|i} \right] P_{i\|i}^{-1} \hat{x}_{i\|i}$
$P_{i+1\|i+1}^{-1} = F_i^{-T} \widehat{P}_{i\|i}^{-1} F_i^{-1} - K_{\nu,i} R_{\nu,i}^{-1} K_{\nu,i}^T + H_{i+1}^T \widehat{R}_{i+1}^{-1} H_{i+1}$
$K_{\nu,i} = F_i^{-T} \widehat{P}_{i\|i}^{-1} F_i^{-1} \widehat{G}_i$
$R_{\nu,i} = \widehat{Q}_i^{-1} + \widehat{G}_i^T F_i^{-T} \widehat{P}_{i\|i}^{-1} F_i^{-1} \widehat{G}_i$

Table 3: Listing of the proposed robust filtering algorithm in information form.

4.2 Suboptimal Implementations

The algorithms of Tables 1–3 require, at each iteration i, the minimization of $G(\lambda)$ over $(\lambda_{l,i}, \infty)$. It turns out that a reasonable approximation that avoids these repeated minimizations is to choose

$$\hat{\lambda}_i \;=\; (1+\alpha)\lambda_{l,i}. \qquad (41)$$

That is, we set $\hat{\lambda}_i$ at a multiple of the lower bound — if the lower bound is zero, we set $\hat{\lambda}_i$ to zero and replace $\hat{\lambda}_i^{-1}$ by $\hat{\lambda}_i^{\dagger}$ (which is also zero). The parameter α could be made time-variant; it serves as a "tuning" parameter that can be adjusted by the designer. In our simulations, we have observed that this approximation leads to good results.

4.3 Steady-State Results

It can be verified that in the time-invariant case, with $E_f^T E_g = 0$, and under a detectability assumption on $\{F, \overline{H}\}$ and a stabilizability assumption on $\{F, G\widehat{Q}^{1/2}\}$, the above suboptimal filters provide stable steady-state performance for any $\Pi_0 > 0$ and $\alpha > 0$. In addition, with uncertainties only in F, if the nominal model is quadratically stable, then the extended estimator-error system can be shown to also be quadratically stable, and the variance of the estimation error can be shown to be bounded. Details will be provided elsewhere.

5. SIMULATION RESULTS

We consider the following example from [5]: $R = 1$ and

$$F = \begin{bmatrix} 0.9802 & 0.0196 \\ 0 & 0.9802 \end{bmatrix}, \; G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; H = \begin{bmatrix} 1 & -1 \end{bmatrix},$$

$$M = \begin{bmatrix} 0.0198 \\ 0 \end{bmatrix}, \; E_f = \begin{bmatrix} 0 & 5 \end{bmatrix}, \; Q = \begin{bmatrix} 1.9608 & 0.0195 \\ 0.0195 & 1.9605 \end{bmatrix}.$$

We also use $\Pi_0 = I$ and $\hat{x}_0 = 0$. These values describe a quadratically-stable time-invariant model of the form

$$x_{i+1} = (F + M\Delta_i E_f)x_i + u_i, \;\; y_i = Hx_i + v_i.$$

[We are simulating a time-invariant nominal model in order to be able to compare with guaranteed-cost filters [5,6], which are only available for steady-state operation.]

In the results shown here, each point in each curve is the average over 500 experiments. Each experiment j fixes Δ at a random value between -1 and 1 and generates 1000-long random measurements $\{y_i\}$. The data is then filtered by a particular algorithm leading to an estimated trajectory $\{\hat{x}_i^{(j)}\}$ for the experiment j. At the end of the 500 experiments, we have 500 such trajectories (of length 1000 points each) for each algorithm and we can use them to approximate the actual error variance curve by computing the ensemble-average. Figure 1 shows the resulting curves. [We don't show a curve for the robust set-valued estimation algorithm since it exhibited poor performance and suffered from divergence problems.] In our simulations, we used the approximation of Sec. 4.2 with $\alpha = 0.5$.

The top part of the figure highlights the degradation in performance by the Kalman filter due to modeling errors (approx. 4dB relative to an optimal implementation that uses the actual model). It is of course not hard to find other examples where the performance of the Kalman filter is significantly worse. [Try using $M = \text{col}\{0.1980, 0\}$.] The plots in the middle row compare the performance of the proposed filter to that of optimized guaranteed-cost designs from [5,6]. In this example, the new filter shows better transient performance.

The plots in the last row of Fig. 1 compare the performance of the new robust filter with that of an $\mathcal{H}_\infty$ filter. While the leftmost plot suggests a good performance by the $\mathcal{H}_\infty$ filter, this result is actually a bit deceiving; if the $\mathcal{H}_\infty$ is allowed to run for a longer period of time, the existence conditions will be violated (starting at iteration 1541) and divergence occurs (as shown in the bottom rightmost plot of Fig. 1).

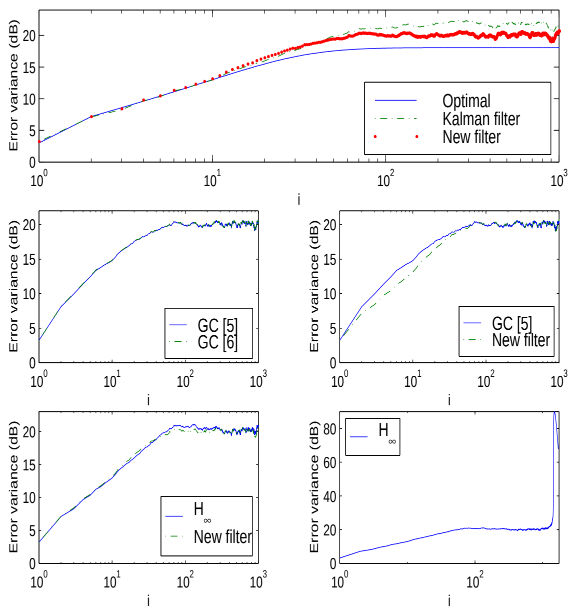

Fig. 1. Error variance curves for different filters with Δ selected uniformly within $[-1, 1]$.

Figure 2 shows the resulting variance curves when Δ is chosen from within the interval $[-1, 0]$. We see that for this case the new filter exhibits the best performance among the robust filters. We also see that the performance of the Kalman filter is comparable to, or even better than, the other filters. We should remember that robust filters are by design well-suited for worst-case scenarios. Hence, their performance will not be superior to non-robust (e.g., Kalman) filters for all possible uncertainties. There will be situations in which non-robust filters will perform better. The purpose of a robust filter should be twofold. On one hand, it avoids degradation in performance under worst-case conditions and, on the other hand, it should exhibit a performance that is comparable to that of non-robust filters under "normal" operating conditions.

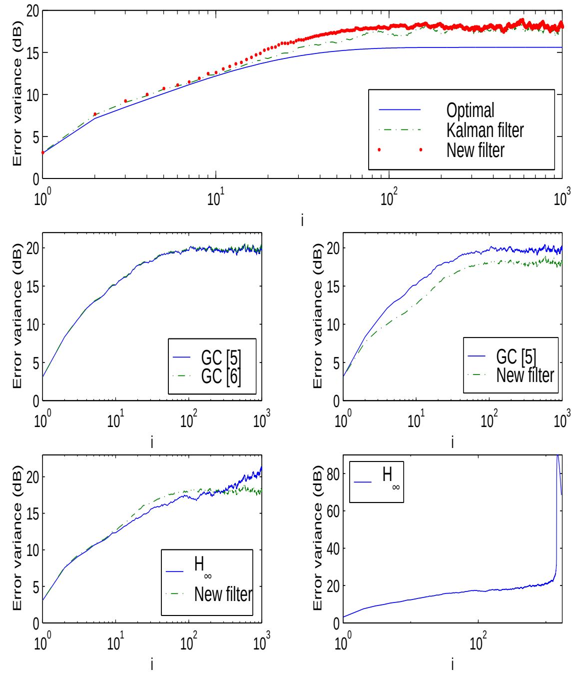

Fig. 2. Error variance curves for different filters with Δ selected uniformly within $[-1, 0]$.

References

[1] T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation.* Prentice-Hall, NJ, 2000.

[2] B. Hassibi, A. H. Sayed, and T. Kailath. *Indefinite Quadratic Estimation and Control: A Unified Approach to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Theories.* SIAM, PA, 1999.

[3] M. Milanese et al., editors. *Bounding Approaches to System Identification*, Plenum, NY, 1996.

[4] D. P. Bertsekas and I. B. Rhodes. Recursive state estimation for a set-membership description of uncertainty. *IEEE Trans. Automat. Contr.*, vol. 16, no. 2, pp. 117–128, 1971.

[5] I. R. Petersen and A. V. Savkin. *Robust Kalman Filtering for Signals and Systems with Large Uncertainties.* Birkhauser, Boston, 1999.

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

[7] A. H. Sayed, V. H. Nascimento, and S. Chandrasekaran. Estimation and control with bounded data uncertainties. *Linear Algebra and Its Applications*, vol. 284, pp. 259–306, Nov. 1998.

[8] A. H. Sayed and V. H. Nascimento. Design criteria for uncertain models with structured and unstructured uncertainties, in *Robustness in Identification and Control*, A. Garulli, A. Tesi, and A. Vicino, editors, vol. 245, pp. 159–173, Springer Verlag, London, 1999.

[9] A. E. Bryson and Y.-C. Ho. *Applied Optimal Control: Optimization, Estimation, and Control.* Taylor & Francis, revised printing, 1975.