Proceedings of the 39th IEEE
Conference on Decision and Control
Sydney, Australia • December, 2000

State-Space Estimation with Uncertain Data: Finite and Infinite-Horizon Results[1]

ALI H. SAYED

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

Abstract

This paper develops a robust estimation procedure for state-space models with parametric uncertainties. Compared with existing robust filters, the proposed filter performs data regularization rather than de-regularization. It is shown that, under certain stabilizability and detectability conditions, the steady-state filter is stable and that, for quadratically-stable models, it guarantees a bounded error variance.

1. INTRODUCTION

The Kalman filter is widely recognized as the optimal linear least-mean-squares state estimator for linear state-space models [1]. When the underlying model is subject to parametric uncertainties, the performance of the filter can deteriorate appreciably. This fact has motivated over the years a variety of ingenious robust estimation techniques such as $\mathcal{H}_\infty$ filters, guaranteed-cost filters, and set-valued estimation methods (see, e.g., [2]-[6] and the references therein).

In recent work [7], we have discussed a robust formulation for state-space estimation. Compared with the standard Kalman filter, which is known to minimize the regularized residual norm at each iteration, the new formulation minimizes the worst-possible regularized residual norm over the class of admissible uncertainties. In addition, compared with other robust formulations, the resulting filter performs data regularization rather than de-regularization; a property that avoids the need for existence conditions. In this paper, we shall first review the filter of [7] and then provide new results concerning its steady-state stability. We shall also show that the filter meets a certain guaranteed-cost property in the case of quadratically-stable models. We start our exposition by reviewing a least-squares problem for uncertain data from [7,8].

2. UNCERTAIN LEAST-SQUARES

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

$$\min_x \left[x^TQx + (Ax-b)^TW(Ax-b)\right] \tag{1}$$

where x^TQx 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

$$\hat{x} = [Q + A^TWA]^{-1}A^TWb. \tag{2}$$

In practice, the nominal data $\{A, b\}$ are often subject to uncertainties. Such errors can degrade the performance of the estimator (2) — see [9]. 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 $J(x,y) = x^TQx + R(x,y)$, where

$$R(x,y) \triangleq \left(Ax - b + Hy\right)^T W \left(Ax - b + Hy\right)$$

Here H is an $N \times m$ known matrix and y is an $m \times 1$ unknown perturbation vector with a bound on its Euclidean norm, say $\|y\| \leq \phi(x)$, for some known nonnegative function $\phi(x)$. Consider then the problem of solving

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

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, the problem (3) trivializes to (1). The statement (3) 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. In the sequel we focus on a special case of (3), namely (see [8]):

$$\min_x \max_{\begin{pmatrix}\delta A \\ \delta b\end{pmatrix}} \left[x^TQx + \left((A+\delta A)x - (b+\delta b)\right)^T W \left(\cdot\right)\right] \tag{4}$$

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

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 $\{\delta A, \delta b\}$ are assumed to satisfy a model of the form

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

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). The following result is proven in [8].

Theorem 1 (Solution) *The problem (4)-(5) has a unique solution $\hat{x}$ that is given by (compare with (2))*

$$\hat{x} = \left[\widehat{Q} + A^T\widehat{W}A\right]^{-1}\left[A^T\widehat{W}b + \hat{\lambda}E_a^T E_b\right] \quad (6)$$

where $\{\widehat{Q}, \widehat{W}\}$ are obtained from $\{Q, W\}$ via

$$\widehat{Q} \triangleq Q + \hat{\lambda}E_a^T E_a \quad (7)$$

$$\widehat{W} \triangleq W + WH(\hat{\lambda}I - H^T WH)^{\dagger}H^T W \quad (8)$$

and the scalar parameter $\hat{\lambda}$ is determined from the optimization

$$\hat{\lambda} = \arg \min_{\lambda \geq \|H^T WH\|} G(\lambda). \quad (9)$$

Here the cost function $G(\lambda)$ is defined as follows:

$$G(\lambda) \triangleq x^T(\lambda)Qx(\lambda) + \lambda\|E_a x(\lambda) - E_b\|^2 + [Ax(\lambda) - b]^T W(\lambda)[Ax(\lambda) - b] \quad (10)$$

where

$$W(\lambda) \triangleq W + WH\big(\lambda I - H^T WH\big)^{\dagger}H^T W \quad (11)$$

$$Q(\lambda) \triangleq Q + \lambda E_a^T E_a \quad (12)$$

and

$$x(\lambda) \triangleq \left[Q(\lambda) + A^T W(\lambda)A\right]^{-1}\left[A^T W(\lambda)b + \lambda E_a^T E_b\right] \quad (13)$$

[The notation $X^{\dagger}$ denotes the pseudo-inverse of X.]

◇

We shall denote the lower bound on λ in (9) by $\lambda_l = \|H^T WH\|$.[1] The function $G(\lambda)$ can be shown to have a unique global minimum (and no local minima) in the interval (λ_l, ∞). Moreover, in the state-space context described below, the matrix W will be positive-definite so that $W(\lambda) > 0$. Therefore, if we restrict the minimization in (9) to the open interval (λ_l, ∞), then the pseudo-inverse operation in (11) can be replaced by the normal matrix inversion, so that it holds that

$$W^{-1}(\lambda) = W^{-1} - \lambda^{-1}HH^T. \quad (14)$$

[1]The notation $\|\cdot\|$ is also used to denote the induced 2-norm of its matrix argument.

3. ROBUST STATE-SPACE ESTIMATION

In the work [7], we described one way to incorporate the uncertain least-squares formulation into a Kalman filtering context. Thus consider a state-space model of the form

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

$$y_i = H_i x_i + v_i \quad (16)$$

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

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

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

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

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

with corresponding error variances P_i and $P_{i|i}$, respectively. The notation l.l.m.s. stands for "linear least-mean-squares". Then $\{\hat{x}_i, \hat{x}_{i|i}\}$ can be constructed recursively via the following time- and measurement-update form of the Kalman filter (see, e.g., [1]):

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

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

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

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

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

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

with initial conditions

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

It can also be verified that these equations are equivalent to the following prediction form of the Kalman filter:

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

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

$$K_i = F_i P_i H_i^T, \quad R_{e,i} = R_i + H_i P_i H_i^T \quad (26)$$

with initial conditions

$$\hat{x}_0 = 0, \quad P_0 = \Pi_0$$

Each step (18)–(23) of the time- and measurement-update form admits a useful deterministic interpretation as the solution to a regularized least-squares problem (e.g., [10]). 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

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

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

$$\begin{aligned} x &\longleftarrow \operatorname{col}\{x_i - \hat{x}_{i|i},\ u_i\}, \quad b \longleftarrow y_{i+1} - H_{i+1}F_i\hat{x}_{i|i} \\ A &\longleftarrow H_{i+1}\begin{bmatrix} F_i & G_i \end{bmatrix}, \quad Q \longleftarrow (P_{i|i}^{-1} \oplus Q_i^{-1}) \\ W &\longleftarrow R_{i+1}^{-1} \end{aligned}$$

The solution of this problem can be shown to lead to (18)–(23).

Consider now an uncertain state-space model of the form

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

$$y_i = H_i x_i + v_i \quad (29)$$

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

for some known matrices $\{M_i, E_{f,i}, E_{g,i}\}$ and for an arbitrary contraction Δ_i, $\|\Delta_i\| \leq 1$. Assume further 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:

$$\begin{aligned} x &\longleftarrow \operatorname{col}\{x_i - \hat{x}_{i|i},\ u_i\}, \quad b \longleftarrow y_{i+1} - H_{i+1}F_i\hat{x}_{i|i} \\ \delta A &\longleftarrow H_{i+1}M_i\Delta_i \begin{bmatrix} E_{f,i} & E_{g,i} \end{bmatrix} \\ \delta b &\longleftarrow -H_{i+1}M_i\Delta_i E_{f,i}\hat{x}_{i|i}, \quad Q \longleftarrow (P_{i|i}^{-1} \oplus Q_i^{-1}) \\ W &\longleftarrow R_{i+1}^{-1}, \ H \longleftarrow H_{i+1}M_i, \ E_a \longleftarrow \begin{bmatrix} E_{f,i} & E_{g,i} \end{bmatrix} \\ E_b &\longleftarrow -E_{f,i}\hat{x}_{i|i}, \ \Delta \longleftarrow \Delta_i, \ A \longleftarrow H_{i+1}\begin{bmatrix} F_i & G_i \end{bmatrix} \end{aligned}$$

This leads, after some algebra, to the equations shown in Table 1 where we defined[2]

$$\lambda_{l,i} \triangleq \|M_i^T H_{i+1}^T R_{i+1}^{-1} H_{i+1} M_i\| \quad (32)$$

The major step in the algorithm of Table 1 is step 3, which consists of recursions that are very similar in nature to the prediction form of the Kalman filter. The main difference is that the new recursions operate on modified parameters rather than on the given nominal values. In addition, the recursion for P_i 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. However, in the special case $E_{g,i} = 0$ (no uncertainties in G_i), it is easy to see that we get

$$\widehat{Q}_i = Q_i, \quad \widehat{G}_i = G_i, \quad \widehat{F}_i = F_i(I - \hat{\lambda}_i\widehat{P}_{i|i}E_{f,i}^T E_{f,i}).$$

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

Likewise, in the case $E_{f,i}^T E_{g,i} = 0$, 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 becomes a standard Riccati recursion. In the work [7], we have further presented alternative equivalent implementations of the robust filter of Table 1 in information form and in time- and measurement-update form.

<u>Assumed uncertain model</u>. Eqs. (28)–(30).

<u>Initial conditions</u>: $\hat{x}_0 = 0$, $P_0 = \Pi_0$, and $\widehat{R}_0 = R_0$.

<u>Step 1a.</u> Using $\{\widehat{R}_i, H_i, P_i\}$ compute $P_{i|i}$:

$$\begin{aligned} P_{i|i} &= (P_i^{-1} + H_i^T \widehat{R}_i^{-1} H_i)^{-1} \\ &= P_i - P_i H_i^T (\widehat{R}_i + H_i P_i H_i^T)^{-1} H_i P_i \end{aligned}$$

<u>Step 1b.</u> 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 $(\lambda_{l,i}, \infty)$.

<u>Step 2.</u> Compute the corrected parameters:

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

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 $\widehat{F}_i = F_i$.

<u>Step 3.</u> Now update $\{\hat{x}_i, P_i\}$ to $\{\hat{x}_{i+1}, P_{i+1}\}$ as follows:

$$\begin{aligned} \hat{x}_{i+1} &= \widehat{F}_i \hat{x}_i + \widehat{F}_i P_i H_i^T R_{e,i}^{-1} e_i \\ e_i &= y_i - H_i \hat{x}_i \\ P_{i+1} &= F_i P_i F_i^T - \overline{K}_i \overline{R}_{e,i}^{-1} \overline{K}_i^T + \widehat{G}_i \widehat{Q}_i \widehat{G}_i^T \\ \overline{K}_i &= F_i P_i \overline{H}_i^T, \quad \overline{R}_{e,i} = I + \overline{H}_i P_i \overline{H}_i^T \end{aligned}$$

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

Observe further that the algorithm of Table 1 requires, 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

$$\hat{\lambda}_i = (1 + \alpha)\lambda_{l,i}. \quad (33)$$

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 (e.g., [7]), we have observed that this approximation leads to good results.

4. STEADY-STATE RESULTS

We now examine the steady-state performance of the filter of Table 1 when the model parameters are constant, say $\{F, G, H, M, E_f, E_g, Q, R\}$. [Only the contraction Δ_i is allowed to vary with time.] In particular we shall establish that, under certain detectability and stabilizability assumptions, the steady-state filter is stable and that, in addition, for quadratically-stable models it guarantees a bounded error variance.

4.1 Stable Performance

We consider first the special case that involves uncertainties in the system dynamics only. That is, we consider an uncertain model of the form

$$\begin{aligned} x_{i+1} &= (F + \delta F_i)x_i + Gu_i, \quad i \geq 0 && (34) \\ y_i &= Hx_i + v_i && (35) \\ \delta F_i &= M\Delta_i E_f && (36) \end{aligned}$$

where only Δ_i (and hence δF_i) is allowed to change with time. This is a model that is often studied in the literature of robust filtering. We further assume that the correction parameter $\hat{\lambda}_i$ is set to a constant value that is equal to a multiple of the admissible lower bound, i.e.,

$$\hat{\lambda}_i = (1+\alpha)\lambda_l = (1+\alpha)\|M^T H^T R^{-1} H M\| \triangleq \hat{\lambda} \quad (37)$$

for some $\alpha > 0$ chosen by the designer, and for all i.

The prediction form of the robust filter in this case becomes (cf. Table 1):[3]

$$\begin{aligned} \hat{x}_{i+1} &= \widehat{F}_i\hat{x}_i + \widehat{F}_i P_i H^T R_{e,i}^{-1}[y_i - H\hat{x}_i] \\ &= \widehat{F}_i[I - P_i H^T R_{e,i}^{-1} H]\hat{x}_i + \widehat{F}_i P_i H^T R_{e,i}^{-1} y_i \end{aligned}$$

where

$$\begin{aligned} \widehat{R} &= R - \hat{\lambda}^{-1} H M M^T H^T \\ R_{e,i} &= \widehat{R} + H P_i H^T \\ P_{i+1} &= F P_i F^T - \overline{K}_i \overline{R}_{e,i}^{-1} \overline{K}_i^T + G Q G^T \\ \overline{K}_i &= F P_i \overline{H}^T \\ \overline{R}_{e,i} &= I + \overline{H} P_i \overline{H}^T \\ \overline{H}^T &= \begin{bmatrix} H^T \widehat{R}^{-T/2} & \sqrt{\hat{\lambda}}\, E_f^T \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} \widehat{F}_i &= F\left[I - \hat{\lambda}\left(P_i^{-1} + H^T\widehat{R}^{-1}H + \hat{\lambda}E_f^T E_f\right)^{-1} E_f^T E_f\right] \\ &= F\left[I - \hat{\lambda}\left(I + P_i\overline{H}^T\overline{H}\right)^{-1} P_i E_f^T E_f\right] \end{aligned}$$

where the second form is independent of P_i^{-1}.

[3] Note that even though the coefficient matrix F is constant, the matrix that appears multiplying $\hat{x}_i$ is time-variant and equal to $\widehat{F}_i$. This is in contrast to a Kalman filtering implementation.

Lemma 1 (Two useful identities) *The following two identities hold:*

$$\begin{aligned} F[I - P_i\overline{H}^T\overline{R}_{e,i}^{-1}\overline{H}] &= \widehat{F}_i[I - P_i H^T R_{e,i}^{-1} H] \\ \widehat{F}_i P_i H^T R_{e,i}^{-1} &= F_{p,i} P_i H^T \widehat{R}^{-1} \end{aligned}$$

Proof: Introduce, for compactness of notation, the two matrices

$$F_{c,i} \triangleq F[I - P_i\overline{H}^T\overline{R}_{e,i}^{-1}\overline{H}], \quad F_{p,i} \triangleq \widehat{F}_i[I - P_i H^T R_{e,i}^{-1} H]$$

We want to show that they coincide. Thus using the identity

$$[I - P_i\overline{H}^T\overline{R}_{e,i}^{-1}\overline{H}]^{-1} = I + P_i\overline{H}^T\overline{H}$$

we get $F_{c,i}[I + P_i\overline{H}^T\overline{H}] = F$. Likewise, a similar argument shows that $F_{p,i}[I + P_i H^T \widehat{R}^{-1} H] = \widehat{F}_i$. It also follows from the expression for $\widehat{F}_i$ that

$$\begin{aligned} \widehat{F}_i &= F - \hat{\lambda}F(I + P_i\overline{H}^T\overline{H})^{-1}P_i E_f^T E_f \\ &= F - \hat{\lambda}F_{c,i}P_i E_f^T E_f \end{aligned}$$

so that

$$\begin{aligned} F_{p,i}(I + P_i H^T R^{-1} H) &= F - \hat{\lambda}F_{c,i}P_i E_f^T E_f \\ &= F_{c,i}(I + P_i H^T R^{-1} H) \end{aligned}$$

Now since the matrix $I + P_i H^T R^{-1} H$ is invertible, we conclude that $F_{p,i} = F_{c,i}$. The second equality of the lemma follows from a straightforward calculation.

◊

Using the second identity in the lemma we can rewrite the recursion for the state estimate as

$$\hat{x}_{i+1} = F_{p,i}\hat{x}_i + F_{p,i}P_i H^T \widehat{R}^{-1} y_i. \quad (38)$$

We are now in a position to establish the main result of this section concerning the convergence of the robust filter to a stable steady-state filter (i.e., we show that $F_{p,i}$ becomes stable in steady-state).

Theorem 2 (Stable steady-state filter) *Consider the uncertain state-space model (34)-(36) with the corresponding robust filter (38). Assume further that $\{F, \overline{H}\}$ is detectable and $\{F, GQ^{1/2}\}$ is stabilizable. Then, for any initial condition $P_0 = \Pi_0 > 0$ and for any $\alpha > 0$ in (37), the Riccati variable P_i tends to the unique stabilizing and positive semi-definite solution P of the DARE*

$$P = FPF^T - FP\overline{H}^T\left(I + \overline{H}P\overline{H}^T\right)^{-1}\overline{H}PF^T + GQG^T.$$

The solution P is stabilizing in the sense that the steady-state closed-loop matrix, $F_p \triangleq \widehat{F}[I - PH^T R_e^{-1} H]$, is stable, where

$$\begin{aligned} \widehat{F} &= F\left[I - \hat{\lambda}(P - P\overline{H}^T\overline{R}_e^{-1}\overline{H}P)E_f^T E_f\right] \\ R_e &= \widehat{R} + HPH^T \\ \overline{R}_e &= I + \overline{H}P\overline{H}^T \end{aligned}$$

Proof: The condition $\alpha > 0$ guarantees a positive-definite matrix $\widehat{R}$ so that its Cholesky factor, and hence $\overline{H}$, are well defined. Now the detectability of $\{F, \overline{H}\}$ and the stabilizability of $\{F, GQ^{1/2}\}$ are known to guarantee the convergence of P_i to the unique positive semi-definite solution P of the DARE that stabilizes the following matrix (see, e.g., [1]):

$$F_c \triangleq F[\, I - P\overline{H}^T(I + \overline{H}P\overline{H}^T)^{-1}\overline{H}\,].$$

But we know from the first identity in the previous lemma that this matrix coincides with F_p, and the result is therefore established.

◇

A similar conclusion can be obtained for the more general uncertain model

$$\begin{aligned} x_{i+1} &= (F + \delta F_i)x_i + (G + \delta G_i)u_i && (39)\\ y_i &= Hx_i + v_i && (40)\\ \begin{bmatrix} \delta F_i & \delta G_i \end{bmatrix} &= M\Delta_i \begin{bmatrix} E_f & E_g \end{bmatrix} && (41)\\ E_f^T E_g &= 0 && (42) \end{aligned}$$

with uncertainties in both F and G that satisfy $E_f^T E_g = 0$. In this case, the same recursions as above will hold with the only exception that the term GQG^T in the recursion for P_{i+1} should be replaced by $G\widehat{Q}G^T$ where $\widehat{Q} = \left(Q^{-1} + \hat{\lambda} E_g E_g^T\right)^{-1}$. The conclusion of Thm. 2 will continue to hold.

4.2 Bounded Steady-State Error Variance

We continue with the model (34)–(36) and further assume now that it is quadratically stable, i.e., that there exists a positive-definite matrix V such that

$$V - [F + M\Delta E_f]^T V [F + M\Delta E_f] > 0$$

for all contractions Δ. By the small gain theorem of [12,13], this condition is equivalent to the combined conditions of a stable F *and* a bounded norm $\|E_f(zI - F)^{-1}M\|_\infty < 1$.[4] For such systems we now argue that the steady-state robust filter of the previous section guarantees a bounded error variance.

Introduce the estimation error $\tilde{x}_i = x_i - \hat{x}_i$. Then subtracting the equations

$$\begin{aligned} x_{i+1} &= (F + M\Delta_i E_f)x_i + Gu_i\\ \hat{x}_{i+1} &= F_p\hat{x}_i + F_pPH^T\widehat{R}^{-1}[Hx_i + v_i] \end{aligned}$$

we arrive at the extended recursion

$$\begin{bmatrix} \tilde{x}_{i+1} \\ \hat{x}_{i+1} \end{bmatrix} = (\mathcal{F} + \delta\mathcal{F}_i)\begin{bmatrix} \tilde{x}_i \\ \hat{x}_i \end{bmatrix} + \mathcal{G}\begin{bmatrix} u_i \\ v_i \end{bmatrix} \qquad (43)$$

where

$$\begin{aligned} \mathcal{F} + \delta\mathcal{F}_i &= \begin{bmatrix} F - F_pPH^T\widehat{R}^{-1}H & F - F_p - F_pPH^T\widehat{R}^{-1}H \\ F_pPH^T\widehat{R}^{-1}H & F_p + F_pPH^T\widehat{R}^{-1}H \end{bmatrix}\\ &\quad + \begin{bmatrix} M\Delta_i E_f & M\Delta_i E_f \\ 0 & 0 \end{bmatrix}\\ \mathcal{G} &= \begin{bmatrix} G & -F_pPH^T\widehat{R}^{-1} \\ 0 & F_pPH^T\widehat{R}^{-1} \end{bmatrix} \end{aligned}$$

[4] Here, $\|\cdot\|_\infty$ denotes the peak singular value of its argument over values of z along the unit circle.

Lemma 2 (Stability of extended system) *The model (43) is quadratically stable.*

Proof: Introduce the similarity transformation

$$\mathcal{T} = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix}, \quad \mathcal{T}^{-1} = \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix}$$

Then the system matrices $\{\mathcal{F} + \delta\mathcal{F}_i, \mathcal{G}\}$ reduce to

$$\begin{aligned} \mathcal{T}(\mathcal{F} + \delta\mathcal{F}_i)\mathcal{T}^{-1} &= \begin{bmatrix} F & 0 \\ F_pPH^T\widehat{R}^{-1}H & F_p \end{bmatrix} + \begin{bmatrix} M\Delta_i E_f & 0 \\ 0 & 0 \end{bmatrix}\\ \mathcal{T}\mathcal{G} &= \begin{bmatrix} G & 0 \\ 0 & F_pPH^T\widehat{R}^{-1} \end{bmatrix} \end{aligned}$$

The stability of F and F_p guarantees that the nominal matrix

$$\begin{bmatrix} F & 0 \\ F_pPH^T\widehat{R}^{-1}H & F_p \end{bmatrix}$$

is stable. Moreover, the equality

$$\begin{bmatrix} E_f & 0 \end{bmatrix}\begin{bmatrix} zI - F & 0 \\ -F_pPH^T\widehat{R}^{-1}H & zI - F_p \end{bmatrix}^{-1}\begin{bmatrix} M \\ 0 \end{bmatrix} = E_f(zI - F)^{-1}M$$

shows that the matrix function on the left-hand side has $\mathcal{H}_\infty$-norm strictly less than one. We thus conclude that the extended system (43) is quadratically stable.

◇

By the result of the above lemma, we also conclude that there exists a positive-definite matrix $\mathcal{V}$ such that

$$\mathcal{V} - (\mathcal{F} + \delta\mathcal{F}_i)\,\mathcal{V}\,(\mathcal{F} + \delta\mathcal{F}_i)^T > 0$$

for any Δ_i. Now let $\mathcal{M}_i$ denote the covariance matrix of the extended vector col$\{\tilde{x}_i, \hat{x}_i\}$:

$$\mathcal{M}_i \triangleq E\begin{bmatrix} \tilde{x}_i \\ \hat{x}_i \end{bmatrix}\begin{bmatrix} \tilde{x}_i \\ \hat{x}_i \end{bmatrix}^T$$

It then follows from (43) that $\mathcal{M}_i$ satisfies the Lyapunov recursion

$$\mathcal{M}_{i+1} = (\mathcal{F} + \delta\mathcal{F}_i)\mathcal{M}_i(\mathcal{F} + \delta\mathcal{F}_i)^T + \mathcal{G}\begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\mathcal{G}^T.$$

Now using arguments that are common in guaranteed-cost designs (e.g., as in [5, 6]), it is immediate to establish the following conclusion.

Theorem 3 (Bounded error-variance) *Under the conditions of Thm. 2, and for a quadratically stable model (34)–(36), the variance of the estimation error of the steady-state robust filter satisfies*

$$\lim_{i\to\infty} E\tilde{x}_i\tilde{x}_i^T \le \mathcal{P}_{11}$$

where $\mathcal{P}_{11}$ is the $(1,1)$ block entry with the smallest trace among all $(1,1)$ block entries of positive-definite matrices $\mathcal{P}$ that satisfy the inequality

$$\mathcal{P} - \left(\mathcal{F} + \begin{bmatrix} M \\ 0 \end{bmatrix} \Delta \begin{bmatrix} E_f & E_f \end{bmatrix}\right) \mathcal{P} \left(\cdot\right)^T - \mathcal{G} \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \mathcal{G}^T \geq 0$$

for all contractive matrices Δ.

Proof: We use an argument similar to the one in [5, pp. 39-40]. The existence of $\mathcal{V} > 0$ also guarantees the existence of a positive scaling parameter μ such that

$$\mu\mathcal{V} - (\mathcal{F} + \delta\mathcal{F}_i)\, \mu\mathcal{V}(\cdot)^T - \mathcal{G} \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \mathcal{G}^T > 0$$

so that a $\mathcal{P}$ exists ($\mathcal{P} = \mu\mathcal{V}$) satisfying

$$\mathcal{P} \geq (\mathcal{F} + \delta\mathcal{F}_i)\, \mathcal{P}\, (\mathcal{F} + \delta\mathcal{F}_i)^T + \mathcal{G} \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \mathcal{G}^T$$

Subtracting this inequality from the recursion for $\mathcal{M}_i$ we get

$$\mathcal{P} - \mathcal{M}_{i+1} \geq (\mathcal{F} + \delta\mathcal{F}_i)\, (\mathcal{P} - \mathcal{M}_i)\, (\mathcal{F} + \delta\mathcal{F}_i)^T$$

or, equivalently,

$$\mathcal{P} - \mathcal{M}_{i+1} = (\mathcal{F} + \delta\mathcal{F}_i)\, (\mathcal{P} - \mathcal{M}_i)\, (\mathcal{F} + \delta\mathcal{F}_i)^T + \mathcal{Q}_i$$

for some $\mathcal{Q}_i \geq 0$. The quadratic stability of $\mathcal{F} + \delta\mathcal{F}_i$ then implies that in the limit, as $i \to \infty$, $\mathcal{P} - \mathcal{M}_{i+1} \geq 0$ or $\mathcal{M}_{i+1} \leq \mathcal{P}$.

$\diamond$

5. Concluding Remarks

There are several issues that deserve investigation. One issue is to pursue more guided selections of the tuning parameter α, e.g., by affecting the value of the bound in Thm. 3. Another issue is to extend the results to other classes of model uncertainties, such as replacing (30) with conditions of the form

$$\|\delta F_i\| \leq \eta_{f,i}, \quad \|\delta G_i\| \leq \eta_{g,i}$$

for some known bounds $\{\eta_{n,i}, \eta_{f,i}\}$. This corresponds to a different choice of the function $\phi(x)$ in (3). A third issue is a closer examination of the stochastic properties of the filter and a more explicit characterization of the error variance. A fourth issue is the development of array variants, in addition to fast algorithms. The former would tend to exhibit better numerical properties while the latter would be more appropriate for large-scale problems.

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. State estimation with uncertain parametric models, *CD-ROM Proceedings of the 8th IEEE Mediterranean Conference on Control & Automation*, Rio, Greece, July 2000.

[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. 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.

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

[11] A. Packard and J. Doyle. Quadratic stability with real and complex perturbations. *IEEE Transactions on Automatic Control*, vol. AC-35, pp. 198–201, 1990.

[12] P. P. Khargonekar, I. R. Petersen, and K. Zhou. Robust stabilization of uncertain systems and $\mathcal{H}^\infty$ optimal control. *IEEE Transactions on Automatic Control*, vol. AC-35, no. 3, pp. 356–361, 1990.