Proceedings of the 32nd Conference
on Decision and Control
San Antonio, Texas • December 1993

RECURSIVE LINEAR ESTIMATION IN KREIN SPACES - PART II: APPLICATIONS *

BABAK HASSIBI, ALI H. SAYED† AND THOMAS KAILATH

Information Systems Laboratory, Stanford University, CA 94305, USA.

ABSTRACT

We show that several applications recently considered in the context of H^∞-filtering and game theory, risk sensitive control and estimation, follow as special cases of the Krein space Kalman filter developed in the companion paper [1]. We show that these problems can be cast into the problem of calculating the stationary points of certain second order forms, and that by considering appropriate state space models and error Gramians, we can use the Krein space Kalman filter to recursively compute these stationary points and to study their properties.

I. INTRODUCTION

Classical results in linear least-squares estimation and Kalman filtering are based on an L^2-criterion and require apriori knowledge of the statistical properties of the noise signals. In some applications however, one is often faced with model uncertainties and lack of statistical information on the exogenous signals, which has led to an increasing interest in minimax estimation (see, *e.g.*, [2]–[8] and the references therein), with the belief that the resulting so-called H^∞ algorithms will be more robust and less sensitive to parameter variations.

Although the general consensus is that the filters obtained for H^∞ estimation are totally different from the conventional Kalman filter, we shall presently show that they are nothing more than certain Krein space Kalman filters. In other words, the H^∞ filters can be viewed as recursively performing a (Gram-Schmidt) orthogonalization (or projection) procedure on a convenient set of observation data that obey a state-space model with entries in an indefinite metric space. This is of significance since it yields a derivation of the H^∞ filters that is, in several respects, simpler than those given in the literature, and because it unifies H_2 and H^∞ estimation in a simple framework. Moreover, once this connection has been made explicit, many known alternative and more efficient algorithms, such as square-root arrays and Chandrasekhar equations, can be used in the H^∞-setting as well. Finally, we should note that our results deal directly with the time-variant scenario, and that we restrict ourselves here, for brevity, to the discrete-time case. However, one should be able to reproduce the continuous time analogs following the same principles. Many of the results discussed here were obtained earlier by several other authors, and using different methods and arguments. Our approach, we believe, provides a powerful unification, with immediate insights to various extensions.

As was done in the companion paper [1], we shall use bold letters to denote elements in a Krein space. We shall also use $\check{z}$ to denote the estimate of z based on some observations (according to a certain error criterion), and $\hat{z}$ to denote the estimate of z provided by the Krein space Kalman filter (*cf.* Theorem 2 in [1]), thereby stressing the fact that they need not coincide. This distinction will become clear later.

II. H_∞ ESTIMATION

We begin with the definition of the H^∞-norm of a transfer operator. Let h_2 denote the vector space of square summable causal sequences $\{f_k,\ 0 \le k < \infty\}$, with inner product $< \{f_k\}, \{g_k\} > = \sum_{k=0}^{\infty} f_k^* g_k$, where $*$ denotes complex conjugation. Let T be a transfer operator that maps a causal input sequence $\{u_k\}$ to a causal output sequence $\{y_k\}$. The H^∞ norm of T is equal to

$$\|T\|_\infty = \sup_{u \in h_2, u \neq 0} \frac{\|y\|_2}{\|u\|_2} ,$$

where the notation $\|u\|_2$ denotes the h_2-norm of

---

†Dr. Sayed is currently with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106, USA. This work was performed while he was a Research Associate in the Information Systems Laboratory, Stanford University, and on Leave from Escola Politécnica da Universidade de São Paulo, Brazil.

* This work was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command under Contract AFOSR91-0060 and by the Army Research Office under contract DAAL03-89-K-0109.

0191-2216/93/$3.00 © 1993 IEEE

the causal sequence $\{u_k\}$, viz., $\|u\|_2^2 = \sum_{k=0}^{\infty} u_k^* u_k$. The H^∞ norm can thus be regarded as the maximum *energy gain* from the input u to the output y.

II.1 Formulation of the $\mathbf{H^\infty}$ Problem

We consider a state-space model of the form

$$\begin{aligned} x_{i+1} &= F_i x_i + G_i u_i \quad , \quad x_0 \\ y_i &= H_i x_i + v_i \ , \end{aligned} \tag{1}$$

where x_0, $\{u_i\}$, and $\{v_i\}$ are *unknown* quantities and y_i is the measured output. Let z_i be linearly related to the state x_i via a given matrix L_i, viz., $z_i = L_i x_i$.

We shall be interested in the following two problems. Let $\check{z}_{i|i} = \mathcal{F}_f(y_0, y_1, \ldots, y_i)$ denote the estimate of z_i given observations $\{y_j\}$ from time 0 up to and including time i, according to a certain error criterion to be made precise ahead, and let $\check{z}_i = \mathcal{F}_p(y_0, y_1, \ldots, y_{i-1})$ denote the estimate of z_i given observations $\{y_j\}$ from time 0 to time $i-1$. This defines two estimation errors: the *filtered* error $e_{f,i} = \check{z}_{i|i} - L_i x_i$, and the *predicted* error $e_{p,i} = \check{z}_i - L_i x_i$.

Let T_f (T_p) denote the transfer operator that maps the unknown disturbances $\{\Pi_0^{-1/2}(x_0 - \hat{x}_0), u_i, v_i\}$ to the filtered (predicted) error $e_{f,i}$ ($e_{p,i}$), where $\hat{x}_0$ denotes an initial guess of x_0 and Π_0 denotes a positive definite matrix that reflects apriori knowledge as to how close x_0 is to the initial guess $\hat{x}_0$. The H^∞ estimation problem(s) can now be stated as follows.

Optimal H^∞ Problem. *Find H^∞ optimal estimation strategies $\check{z}_{i|i} = \mathcal{F}_f(y_0, y_1, \ldots, y_i)$ and $\check{z}_i = \mathcal{F}_p(y_0, y_1, \ldots, y_{i-1})$ that respectively minimize $\| T_f \|_\infty$ and $\| T_p \|_\infty$, and obtain the resulting $\gamma_{f,o}^2 = \inf_{\mathcal{F}_f} \| T_f \|_\infty^2$, and $\gamma_{p,o}^2 = \inf_{\mathcal{F}_p} \| T_p \|_\infty^2$, where $\gamma_{f,o}^2 =$*

$$\inf_{\mathcal{F}_f} \sup_{\substack{x_0 \\ u \in h_2 \\ v \in h_2}} \frac{\| e_f \|_2^2}{(x_0 - \hat{x}_0)^* \Pi_0^{-1} (x_0 - \hat{x}_0) + \| u \|_2^2 + \| v \|_2^2}$$

and $\gamma_{p,o}^2 =$

$$\inf_{\mathcal{F}_p} \sup_{\substack{x_0 \\ u \in h_2 \\ v \in h_2}} \frac{\| e_p \|_2^2}{(x_0 - \hat{x}_0)^* \Pi_0^{-1} (x_0 - \hat{x}_0) + \| u \|_2^2 + \| v \|_2^2}$$

The distinction between the *strictly* causal $\mathcal{F}_p$ and the causal $\mathcal{F}_f$ is significant since the solution to the H^∞ problem, as we shall see, depends on the structure of the information available to the estimator. We can also infer from the above problem that the robust behaviour of H^∞ optimal estimators is because they guarantee the smallest estimation error energy over all possible disturbances of fixed energy.

A closed form solution of the optimal H^∞ problem is available only for some special cases (one of which is the adaptive filtering problem studied in [9]), and a simpler problem results if one relaxes the minimization condition and settles for a suboptimal solution.

Sub-optimal H^∞ Problem. *Given scalars $\gamma_f > 0$ and $\gamma_p > 0$, find estimation strategies $\check{z}_{i|i} = \mathcal{F}_f(y_0, y_1, \ldots, y_i)$ and $\check{z}_i = \mathcal{F}_p(y_0, y_1, \ldots, y_{i-1})$ that respectively achieve $\| T_f \|_\infty \leq \gamma_f$ and $\| T_p \|_\infty \leq \gamma_p$. This clearly requires checking whether $\gamma_f \geq \gamma_{f,o}$ and $\gamma_p \geq \gamma_{p,o}$.*

We shall replace the condition $\| T_f \|_\infty \leq \gamma_f$ by the following procedure: let $T_{f,i}$ be the transfer operator that maps the disturbances $\left\{\Pi_0^{-1/2}(x_0 - \hat{x}_0), \{u_j\}_{j=0}^i, \{v_j\}_{j=0}^i\right\}$ to the filtered errors $\{e_{f,j}\}_{j=0}^i$. We shall find a γ_f that ensures $\| T_{f,i} \|_\infty < \gamma_f$ for all i. Likewise we shall find a γ_p that ensures $\| T_{p,i} \|_\infty < \gamma_p$ for all i.

III. RELATION TO QUADRATIC FORMS

We first show how to reduce the suboptimal H^∞ problem(s) to that of determining a stationary point of a second-order scalar form. For this purpose, we first note that $\|T_{f,i}\|_\infty < \gamma_f$ iff for all nonzero complex vectors x_0 and for all nonzero causal sequences $\{u_k, v_k\} \in h_2$, we have

$$\frac{\sum_{k=0}^i (\check{z}_{k|k} - L_k x_k)^* (\check{z}_{k|k} - L_k x_k)}{(x_0 - \hat{x}_0)^* \Pi_0^{-1} (x_0 - \hat{x}_0) + \sum_{k=0}^i u_k^* u_k + \sum_{k=0}^i v_k^* v_k} < \gamma_f^2.$$

Different choices of $\hat{x}_0$ will only change the initial condition of the estimators and therefore, without loss of generality, we shall assume that $\hat{x}_0 = 0$. We thus have the following.

Lemma 1 *Given a scalar $\gamma_f > 0$, then $\|T_{f,i}\|_\infty < \gamma_f$ iff there exists $\check{z}_{k|k}$ (for all $k \leq i$) such that, for all nonzero complex vectors x_0 and for all nonzero causal sequences $\{u_j, v_j\}_{j=0}^i$, the scalar second order form* $J_{f,i} = x_0^* \Pi_0^{-1} x_0 + \sum_{k=0}^i u_k^* u_k +$

$$+ \sum_{k=0}^{i} (y_k - H_k x_k)^* (y_k - H_k x_k) -$$

$$\gamma_f^{-2} \sum_{k=0}^{i} (\check{z}_{k|k} - L_k x_k)^* (\check{z}_{k|k} - L_k x_k)$$

satisfies $J_{f,i} > 0$.

Now observe that we can rewrite J_i as follows: $J_{f,i} = x_0^* \Pi_0^{-1} x_0 + \sum_{k=0}^{i} u_k^* u_k +$

$$\sum_{k=0}^{i} \left(\begin{bmatrix} y_k \\ \check{z}_{k|k} \end{bmatrix} - \begin{bmatrix} H_k \\ L_k \end{bmatrix} x_k \right)^* R_j^{-1} \left(\begin{bmatrix} y_k \\ \check{z}_{k|k} \end{bmatrix} - \begin{bmatrix} H_k \\ L_k \end{bmatrix} x_k \right)$$

where $R_j = (I \oplus -\gamma_f^2 I)$. This is a special case of the quadratic expression of Lemma 2 in the companion paper [1], and it suggests that we introduce the following *auxiliary* state-space model:

$$\begin{aligned} \mathbf{x}_{i+1} &= F_i \mathbf{x}_i + G_i \mathbf{u}_i \\ \begin{bmatrix} \mathbf{y}_i \\ \check{\mathbf{z}}_{i|i} \end{bmatrix} &= \begin{bmatrix} H_i \\ L_i \end{bmatrix} \mathbf{x}_i + \mathbf{w}_i , \end{aligned} \quad (2)$$

where the disturbances $\{\mathbf{x}_0, \mathbf{u}_i, \mathbf{w}_i\}$ are assumed to be elements in a Krein space $\mathcal{K}$ with $< \mathbf{u}_i, \mathbf{u}_j >_{\mathcal{K}} = I\delta_{ij}$, $< \mathbf{x}_0, \mathbf{x}_0 >_{\mathcal{K}} = \Pi_0$, and $< \mathbf{w}_i, \mathbf{w}_j >_{\mathcal{K}} = R_j \delta_{ij}$. Observe that we have to consider elements in a Krein space since the Gramian matrix R_j is indefinite. It is thus clear that the corresponding recursive estimation algorithm of Theorem 2 in the companion paper [1] computes a stationary point of the above $J_{f,i}$. Comparing the above state-space model with that in Theorem 1 in [1] we see that we can identify the quantities $(\mathbf{y}_i, \mathbf{v}_i, H_i, Q_i, R_i)$ in Theorem 1 with the quantities

$$\left(\begin{bmatrix} \mathbf{y}_i \\ \check{\mathbf{z}}_{i|i} \end{bmatrix}, \mathbf{w}_i, \begin{bmatrix} H_i \\ L_i \end{bmatrix}, I, \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -\gamma_f^2 I \end{bmatrix} \right),$$

respectively. The following result is then expected, and shows why finding a stationary (in fact, a minimum) point for $J_{f,i}$ is necessary to guarantee $J_{f,i} > 0$.

Lemma 2 *The scalar second-order form $J_{f,i}$ satisfies $J_{f,i} > 0$ iff $J_{f,i}$ has a minimum with respect to $\{x_0, u_0, \ldots, u_i\}$, and the value at the minimum is positive, viz.,*

$$\sum_{k=0}^{i} \begin{bmatrix} e_{y,k}^* & e_{z,k}^* \end{bmatrix} R_{e,k}^{-1} \begin{bmatrix} e_{y,k} \\ e_{z,k} \end{bmatrix} > 0 ,$$

where $\begin{bmatrix} e_{y,k} \\ e_{z,k} \end{bmatrix} = \begin{bmatrix} y_k \\ \check{z}_{k|k} \end{bmatrix} - \begin{bmatrix} \hat{y}_{k|k-1} \\ \hat{z}_{k|k-1} \end{bmatrix}$ *is the innovations obtained by writing down the Krein space Kalman filter of Theorem 2 in [1] that corresponds to the state-space model (2), and* $\begin{bmatrix} \hat{y}_{k|k-1} \\ \hat{z}_{k|k-1} \end{bmatrix}$ *is the estimate of* $\begin{bmatrix} y_k \\ \check{z}_{k|k} \end{bmatrix}$ *that is obtained via the Krein space Kalman filter.*

A similar argument applies to $\|T_{p,i}\|_\infty < \gamma_p$. In this case we consider the quadratic form: $J_{p,i} = x_0 \Pi_0^{-1} x_0^* + \sum_{k=0}^{i} u_k^* u_k +$

$$\sum_{k=0}^{i} (y_k - H_k x_k)^* (y_k - H_k x_k) -$$
$$-\gamma_p^{-2} \sum_{k=0}^{i+1} (\check{z}_k - L_k x_k)^* (\check{z}_k - L_k x_k) ,$$

and construct the auxiliary Krein state-space model

$$\begin{aligned} \mathbf{x}_{i+1} &= F_i \mathbf{x}_i + G_i \mathbf{u}_i \\ \begin{bmatrix} \check{\mathbf{z}}_i \\ \mathbf{y}_i \end{bmatrix} &= \begin{bmatrix} L_i \\ H_i \end{bmatrix} \mathbf{x}_i + \mathbf{w}_i \end{aligned} \quad (3)$$

with $R_i = (-\gamma_p^2 I \oplus I)$, Q_i, and Π_0. Following the same reasoning as before, an H^∞ apriori estimator of level γ_p will exist iff $\check{z}_k$ can be chosen so as to guarantee $J_{p,i} > 0$ for all possible disturbances.

IV. THE H^∞–FILTERS

We are now in a position to write down the filtered (aposteriori) and predicted (apriori) H^∞– filters.

Aposteriori Filter. *For a given $\gamma_f > 0$, if the $\{F_j\}_{j=0}^{i}$ are nonsingular, then the second-order form satisfies $J_{f,i} > 0$ iff for all $j = 0, \ldots, i$, we have $P_j^{-1} + H_j^* H_j - \gamma^{-2} L_j^* L_j > 0$, where $P_0 = \Pi_0$ and P_j satisfies the Riccati recursion:* $P_{j+1} = F_j P_j F_j^* + G_j G_j^* -$

$$F_j P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} \left\{ R_j + \begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j \begin{bmatrix} H_j^* & L_j^* \end{bmatrix} \right\}^{-1}$$
$$\begin{bmatrix} H_j \\ L_j \end{bmatrix} P_j F_j^* , \quad R_j = (I \oplus -\gamma_f^2 I).$$

If this is the case, then one possible filtered H^∞ filter with level γ_f is given by $\check{z}_{j|j} = L_j \hat{x}_{j|j}$, where $\hat{x}_{-1|-1} = 0$,

$$\hat{x}_{j+1|j+1} = F_j \hat{x}_{j|j} + K_{1,j}(y_{j+1} - H_{j+1} F_j \hat{x}_{j|j}), \quad (4)$$

and $K_{1,j} = P_{j+1} H_{j+1}^* (I + H_{j+1} P_{j+1} H_{j+1}^*)^{-1}$.

We should remark that the above filter is one among *many possible* filters with level γ_f. All filters that guarantee $J_{f,i} > 0$ can be parametrized as follows.

Theorem 1 *All aposteriori H^∞ estimators that achieve a level γ_f (assuming they exist) are given by* $\check{z}_{i|i} = L_i \hat{x}_{i|i} + [I - L_i (P_i^{-1} + H_i^* H_i)^{-1} L_i^*]^{\frac{1}{2}}$.

$$\mathcal{S}_i \left((I + H_i P_i H_i^*)^{\frac{1}{2}} (y_i - H_i \hat{x}_{i|i}), \ldots, \right.$$

$$,\ldots,(I+H_0P_0H_0^*)^{\frac{1}{2}}(y_0-H_0\hat{x}_{0|0})\Big)\ ,$$

where $\mathcal{S}$ is any (possibly nonlinear) contractive causal mapping of the form

$$\mathcal{S}(a_i,\ldots,a_0)=\begin{bmatrix}\mathcal{S}_0(a_0)\\ \mathcal{S}_1(a_1,a_0)\\ \vdots\\ \mathcal{S}_i(a_i,\ldots,a_0)\end{bmatrix}$$

and satisfies $\sum_{j=0}^{i}|\mathcal{S}_j(a_j,\ldots,a_0)|^2<\sum_{j=0}^{i}|a_j|^2$.

Note that although the aposteriori filter given in the beginning of this Section is linear, the full parametrization of all H^∞ aposteriori filters with level γ_f is given by a *nonlinear* causal contractive mapping $\mathcal{S}$. The filter (4) is known as the *central* filter, and corresponds to $\mathcal{S}=0$. It has a number of other interesting properties: it corresponds to the risk-sensitive optimal filter (see, *e.g.*, [10]), and can be also shown to be the *maximum entropy* filter [11]. Moreover, in the game theoretic formulation of the H^∞ problem, the central filter corresponds to the solution of the game.

Apriori Filter. *For a given $\gamma_p>0$, if the $\{F_j\}$ are nonsingular, then the second-order form satisfies $J_{p,i}>0$ iff for all $j=0,1,\ldots i+1$, we have $\tilde{P}_j^{-1}=P_j^{-1}-\gamma_p{}^{-2}L_j^*L_j>0$, where P_j is the same as in the aposteriori filter. If this is the case, then one possible H^∞ apriori filter with level γ_p is given by $\hat{x}_0=0$, $\check{z}_i=L_i\hat{x}_i$,*

$$\hat{x}_{j+1}=F_j\hat{x}_j+K_{2,j}(y_j-H_j\hat{x}_j)\ , \qquad (5)$$

where $K_{2,j}=F_j\tilde{P}_jH_j^*(I+H_j\tilde{P}_jH_j^*)^{-1}$.

We again have a full parametrization of all H^∞ apriori estimators.

Theorem 2 *All H^∞ apriori estimators that achieve a level γ_p (assuming they exist) are given by* $\check{z}_i=L_i\hat{x}_i+[I-L_iP_iL_i^*]^{\frac{1}{2}}$.

$$\mathcal{S}_i\Big((I+H_{i-1}\tilde{P}_{i-1}H_{i-1}^*)^{-\frac{1}{2}}(y_{i-1}-H_{i-1}\hat{x}_{i-1}),\ldots,$$

$$,\ldots,(I+H_0\tilde{P}_0H_0^*)^{-\frac{1}{2}}(y_0-H_0\hat{x}_0)\Big)$$

where $\mathcal{S}$ is as before.

IV.1 The H$^\infty$−Smoother

If instead of $J_{f,i}$ and $J_{p,i}$, which correspond to the aposteriori and apriori filters, respectively, we consider the quadratic form $J_{s,i}=x_0\Pi_0^{-1}x_0^*+\sum_{k=0}^{i}u_k^*u_k+$

$$\sum_{k=0}^{i}(y_k-H_kx_k)^*(y_k-H_kx_k)-$$

$$-\gamma_s{}^{-2}\sum_{k=0}^{i}(\check{z}_{k|i}-L_kx_k)^*(\check{z}_{k|i}-L_kx_k),$$

then we are led to H^∞ smoothers. The main difference between $J_{s,i}$ and the previous second-order forms is that $\check{z}_{k|k}$ and $\check{z}_{k|k-1}$ have been replaced by $\check{z}_{k|i}$, *i.e.*, filtered and predicted estimates have been replaced by smoothed estimates. Once more, an H^∞ smoother of level γ_s is said to exist iff there exists some $\check{z}_{k|i}$ such that $J_{s,i}>0$. The rather interesting result stated below, and which has already been pointed out in the literature (see *e.g.*, [4, 7, 12]), is that the H^∞ smoother is identical to the conventional H^2 smoother.

Theorem 3 *For a given $\gamma_s>0$, if the $\{F_j\}_{j=0}^{i}$ are nonsingular, then the second-order form satisfies $J_{s,i}>0$ iff for all $j=0,1,\ldots i$, we have $P_j^{-1}+H_jH_j^*-\gamma_s{}^{-2}L_j^*L_j>0$, where P_j is the same as in the aposteriori case. If this is the case, then one possible H_∞ smoother is the H^2 smoother.*

V. RISK SENSITIVE FILTERS

The classical Kalman filtering algorithm can be viewed as a recursive procedure that minimizes a convenient quadratic form. There has also been increasing interest in an alternative so-called *exponential cost function* [10, 13, 14], which is risk sensitive, in the sense that it depends on a real parameter that determines whether more or less weight should be given to higher or smaller errors. The corresponding filters have been termed risk-sensitive and include the Kalman filter as a special case. In what follows, we shall show that the risk-sensitive filters are also special cases of the Krein space Kalman filter derived in [1].

V.1 The Exponential Cost Function

We start with a state-space model of the form

$$\begin{aligned}x_{i+1}&=F_ix_i+G_iu_i\ ,\\ y_i&=H_ix_i+v_i\ ,\end{aligned}$$

where x_0, $\{u_i\}$, and $\{v_i\}$ are now zero mean independent Gaussian random variables with covariances Π_0, Q_i, and R_i, respectively. We further assume that the $\{u_i\}$ and $\{v_i\}$ are white-noise processes. The conventional Kalman filter that estimates the quantity $z_i=L_ix_i$ from $\{y_0,y_1,\ldots,y_i\}$ is a linear estimator that performs the following minimization

$$\min_{\check{z}_j}E\left[\sum_{j=0}^{i}(\check{z}_j-L_jx_j)^*(\check{z}_j-L_jx_j)\right]\ ,$$

where $\check{z}_j$ denotes the estimate of z_j given the observations up to and including time $j-1$. Moreover, the expectation is taken over the Gaussian random variables x_0 and $\{u_i\}$ whose joint conditional distribution is given by $p(x_0, U_i|Y_i) \propto exp(-\frac{1}{2}J_i)$, where the symbol $\propto$ stands for 'proportional to' and J_i is equal to (using the fact that x_0, $\{u_i\}$, and $\{v_i\}$ are independent, and that $v_j = y_j - H_j x_j$):

$$J_i = x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{i} u_j^* Q_j^{-1} u_j +$$

$$\sum_{j=0}^{i} (y_j - H_j x_j)^* R_j^{-1} (y_j - H_j x_j).$$

The conventional Kalman filter computes the (one step-ahead estimates), $\hat{x}_{i+1}$, as well as the innovations signals $e_i = y_i - H_i \hat{x}_i$, and the minimizing solution is given by $\check{z}_i = L_i \hat{x}_i$. In the terminology of [10], this filter is known as a *risk-neutral* filter.

An alternative criterion that is risk-sensitive has been extensively studied in [10]-[14] and corresponds to the following minimization problem

$$\min_{\check{z}_j} \mu_i(\theta) = \min_{\check{z}_j} \left(-\frac{2}{\theta} log \left[E exp(-\frac{\theta}{2} C_i) \right] \right) ,$$

where $C_i = \sum_{j=0}^{i} (\check{z}_j - L_j x_j)^* (\check{z}_j - L_j x_j)$. This criterion is known as an *exponential cost function*, and any filter that minimizes $\mu_i(\theta)$ is referred to as a *risk-sensitive* filter. The scalar parameter θ is correspondingly called the *risk-sensitivity* parameter. Some intuition concerning the nature of this modified criterion is obtained by expanding $\mu_i(\theta)$ in terms of θ and writing,

$$\mu_i(\theta) = E(C_i) - \frac{\theta}{4} Var(C_i) + O(\theta^2).$$

The above equation shows that for $\theta = 0$, we have the risk-neutral case (*i.e.*, the conventional Kalman filter). When $\theta > 0$, we seek to maximize $E exp(-\frac{\theta}{2} C_i)$, which is convex and decreasing in C_i. Such a criterion is termed *risk-seeking* (or optimistic). When $\theta < 0$, we seek to minimize $E exp(-\frac{\theta}{2} C_i)$, which is convex and increasing in C_i. Such a criterion is termed *risk-averse* (or pessimistic). In what follows, we shall see that in the risk-averse case $\theta < 0$, the limit at which the minimization makes sense is the H^∞ criterion.

V.2 Minimizing the Risk-Sensitive Criterion

Using the conditional distribution density function we can easily verify that

$$E(exp(-\frac{\theta}{2} C_i)) \propto \int exp(-\frac{\theta}{2} C_i) exp(-\frac{J_i}{2}) \, dx_0 \, dU_i ,$$

which shows that the risk-sensitive criterion can be alternatively written as follows: if $\theta > 0$ then it is equivalent to

$$\max_{\check{z}_j} \int exp(-\frac{\theta}{2} C_i - \frac{1}{2} J_i) \, dx_0 \, dU_i.$$

If $\theta < 0$ then it is equivalent to

$$\min_{\check{z}_j} \int exp(-\frac{\theta}{2} C_i - \frac{1}{2} J_i) \, dx_0 \, dU_i.$$

This suggests that we define the second-order scalar form $\bar{J}_i = J_i + \theta C_i =$

$$x_0^* \Pi_0^{-1} x_0 + \sum_{j=0}^{i} u_j^* Q_j^{-1} u_j +$$

$$\sum_{j=0}^{i} (y_j - H_j x_j)^* R_j^{-1} (y_j - H_j x_j) +$$

$$\theta \sum_{j=0}^{i} (\check{z}_j - L_j x_j)^* (\check{z}_j - L_j x_j)$$

If we now introduce the *auxiliary* state-space model

$$\begin{aligned} \mathbf{x}_{i+1} &= F_i \mathbf{x}_i + G_i \mathbf{u}_i \\ \begin{bmatrix} \check{\mathbf{z}}_i \\ \mathbf{y}_i \end{bmatrix} &= \begin{bmatrix} L_i \\ H_i \end{bmatrix} \mathbf{x}_i + \mathbf{w}_i \end{aligned} \qquad (6)$$

with $< \mathbf{u}_i, \mathbf{u}_j >_{\mathcal{K}} = Q_i \delta_{ij}$, $< \mathbf{x}_0, \mathbf{x}_0 >_{\mathcal{K}} = \Pi_0$, and $< \mathbf{w}_i, \mathbf{w}_j >_{\mathcal{K}} = (\theta^{-1} I \oplus R_i) \delta_{ij}$, Then the recursive estimation algorithm of Theorem 2 in [1] that corresponds to (6) will compute a stationary point of $\bar{J}_i$. We are thus led to the following result.

Theorem 4 *For a given $\theta > 0$, the risk-sensitive aposteriori estimation problem always has a solution. For a given $\theta < 0$, a solution exists iff $P_i^{-1} + H_i^* H_i + \theta L_i^* L_i > 0$, where $P_0 = \Pi_0$ and* $P_{i+1} = F_i P_i F_i^* -$

$$F_i P_i \begin{bmatrix} L_i^* & H_i^* \end{bmatrix} \left\{ \begin{bmatrix} \theta^{-1} I & 0 \\ 0 & R_i \end{bmatrix} + \right.$$

$$\left. \begin{bmatrix} L_i \\ H_i \end{bmatrix} P_i \begin{bmatrix} L_i^* & H_i^* \end{bmatrix} \right\}^{-1} \begin{bmatrix} L_i \\ H_i \end{bmatrix} P_i F_i^*$$

In both cases the optimal risk-sensitive filter with parameter θ is given by $\check{z}_{i|i} = L_i \hat{x}_{i|i}$, $\hat{x}_{-1|-1} = 0$,

$$\hat{x}_{i+1|i+1} = F_i \hat{x}_{i|i} + K_{1,i}(y_{i+1} - H_{i+1} F_i \hat{x}_{i|i}) \quad ,$$

and $K_{1,i} = P_{i+1} H_{i+1}^* (I + H_{i+1} P_{i+1} H_{i+1}^*)^{-1}$.

We can also construct an *apriori* filter similar to what was done in the H^∞ case.

Theorem 5 *For a given $\theta > 0$, the apriori risk-sensitive estimation problem always has a solution. For a given $\theta < 0$, a solution exists iff $\tilde{P}_i^{-1} = P_i^{-1} + \theta L_i^* L_i > 0$, where P_i is the same as in the aposteriori case. In both cases the apriori risk-sensitive filter with parameter θ is given by $\check{z}_i = L_i \hat{x}_i$,*

$$\hat{x}_{i+1} = F_i \hat{x}_i + K_{2,i}(y_i - H_i \hat{x}_i) \ , \ \hat{x}_0 = 0,$$

where $K_{2,i} = F_i \tilde{P}_i H_i^* (I + H_i \tilde{P}_i H_i^*)^{-1}$.

We can now state the striking resemblances between the H^∞ and the risk-sensitive filters. The H^∞ filters obtained earlier are essentially risk-sensitive filters with parameter $\theta^{-1} = -\gamma_f{}^2$ $(-\gamma_p^2)$. Note, however, that at each level $\gamma_f (\gamma_p)$, the H^∞ filters are not unique, whereas for each θ, the risk-sensitive filters are unique. Also, the risk-sensitive filters generalize to the $\theta > 0$ case. It is also noteworthy that the optimal H^∞ filter corresponds to the risk-sensitive filter with $\bar{\theta}^{-1} = -\gamma_{f,o}^2$ $(-\gamma_{p,o}^2)$, and that $\bar{\theta}$ is that value for which the minimizing property of $\bar{J}_i$ breaks down and $\mu_i(\theta)$ becomes infinite. This relationship between the optimal H^∞ filter and the corresponding risk-sensitive filter was first noted in [15].

VI. CONCLUDING REMARKS

We have discussed problems in $H^\infty-$ and risk-sensitive estimation within the framework of the Krein space Kalman filter theory developed in the companion paper [1]. Several other applications fit into the same framework such as finite memory adaptive filtering, $H^\infty-$control, and will be discussed elsewhere.

REFERENCES

[1] B. Hassibi, A. H. Sayed, and T. Kailath, "Recursive linear estimation in Krein spaces - Part I: Theory," *in Proc. of this Conference.*

[2] H. Kwakernaak, "A polynomial approach to mimimax frequency domain optimization of multivariable feedback systems," *Int. J. of Control*, vol. 44, pp. 117–156, 1986.

[3] J. C. Doyle, K. Glover, P. Khargonekar, and B. Francis, "State-space solutions to standard H_2 and H_∞ control problems," *IEEE Transactions on Automatic Control*, vol. AC-34, no. 8, pp. 831–847, August 1989.

[4] P.P. Khargonekar and K. M. Nagpal, "Filtering and smoothing in an $H^\infty-$ setting," *IEEE Trans. on Automatic Control*, AC-36, pp. 151–166, 1991.

[5] T. Basar, "Optimum performance levels for minimax filters, predictors and smoothers," *Systems and Control Letters*, vol. 16, pp. 309–317, 1991.

[6] D. J. Limebeer and U. Shaked, "New results in h^∞-filtering," In *Proc. Int. Symp. on MTNS*, pp. 317–322, June 1991.

[7] U. Shaked and Y. Theodor, "$H_\infty-$optimal estimation: A tutorial," In *Proc. IEEE Conference on Decision and Control*, pp. 2278–2286, Tucson, AZ, Dec. 1992.

[8] M. J. Grimble, "Polynomial matrix solution of the H^∞ filtering problem and the relationship to Riccati equation state-space results," *IEEE Trans. on Signal Processing*, vol. 41, no. 1, pp. 67–81, January 1993.

[9] B. Hassibi, A. H. Sayed, and T. Kailath, "LMS is $H^\infty-$optimal," *in Proc. of this Conference.*

[10] P. Whittle, *Risk Sensitive Optimal Control*, John Wiley and Sons, New York, 1990.

[11] K. Glover and D. Mustafa, "Derivation of the maximum entropy H^∞ controller and a state space formula for its entropy," *Int. J. Control*, vol. 50, pp. 899-916, 1989.

[12] J.L. Speyer, C. Fan, and R. N. Banavar, "Optimal stochastic estimation with exponential cost criteria," In *Proc. IEEE Conference on Decision and Control*, pp. 2293–2298, Tucson, Arizona, December 1992.

[13] D. H. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to determenistic games," *IEEE Trans. Automatic Control*, vol. AC-18, no. 2, April 1973.

[14] J. Speyer, J. Deyst, and D. H. Jacobson, "Optimization of stochastic linear systems with additive measurement and process noise using exponential performance criteria," *IEEE Trans. Automatic Control*, vol. AC-19, pp. 358–366, August 1974.

[15] K. Glover and J. C. Doyle, "State-space formulae for all stabilizing controllers that satisfy an H^∞-norm bound and relations to risk sensitivity," *System and Control Letters*, vol. 11, no. 167–172, 1988.