International Journal of Innovative
Computing, Information and Control
Volume **8**, Number **1(A)**, January **2012**
 ISSN 1349-4198
pp. 1–12

RISK SENSITIVE FIR FILTERS FOR STOCHASTIC DISCRETE-TIME STATE SPACE MODELS

Soohee Han
Department of Electrical Engineering
Konkuk University
Hwayang-dong, Gwangjin-gu, Seoul 143-701, Korea
shhan@konkuk.ac.kr

Received June 2010; revised November 2010

Abstract. *In this paper, the finite impulse response (FIR) filter based on an exponential quadratic cost function is proposed for a stochastic discrete-time state space model. The joint probability density function of the current state and the external noises on the recent finite horizon is introduced and the corresponding expected value of the exponential quadratic cost function is minimized with respect to the current state. According to the sign of the scalar real parameter in the cost function, we have a risk averse or seeking criterion, from which the optimal FIR filter, called a risk sensitive FIR filter (RSFF), is derived. Being risk averse means that large weights are put on large estimation errors which are suppressed as much as possible. Being risk seeking means that large weights are put on moderate estimation errors. It is also shown via simulation that the proposed FIR filter has better performance than the conventional infinite impulse response (IIR) robust Kalman filter.*
Keywords: Risk sensitive, Risk averse, Risk seeking, FIR structure, State estimation

1. Introduction.

The estimation of the unknown values from given measurements arises in many fields such as control, signal processing and communications. Specially, how to estimate a state from measurements on a state space model has been extensively exploited since most dynamic systems can be easily described over state spaces.

For a long time, the optimal estimators or filters for state estimation have been developed on the basis of the Luenberger-type filters such as the Kalman filter [1] and the H_∞ filter [2, 3, 4, 5, 6]. The duration of impulse response of the conventional Kalman and H_∞ filters is infinite, which means that these filters belong to infinite impulse response (IIR) filters in a signal processing area. Actually, these days, these IIR filters give way to finite impulse response (FIR) filters in the signal processing area. It is generally known that the FIR filters are robust against temporary modelling uncertainties or round-off errors. Furthermore, FIR filters can resolve the divergence and the slow convergence known as demerits of IIR filters. As the main disadvantage of FIR filters, they have the more computation load than the conventional IIR filters. However, this computation burden can be alleviated by the recent fast computer technology.

While IIR filters for state estimation have been widely used for a long time, FIR filters for that purpose have not received much attention and have not been researched much. As in the signal processing area, undesirable effects of the IIR filters for state estimation may be alleviated by using the FIR structure. In this paper, we consider an FIR filter given by

$$\hat{x}_k = \sum_{i=k-N}^{k-1} H_{k-i} y_i + \sum_{i=k-N}^{k-1} L_{k-i} u_i, \tag{1}$$

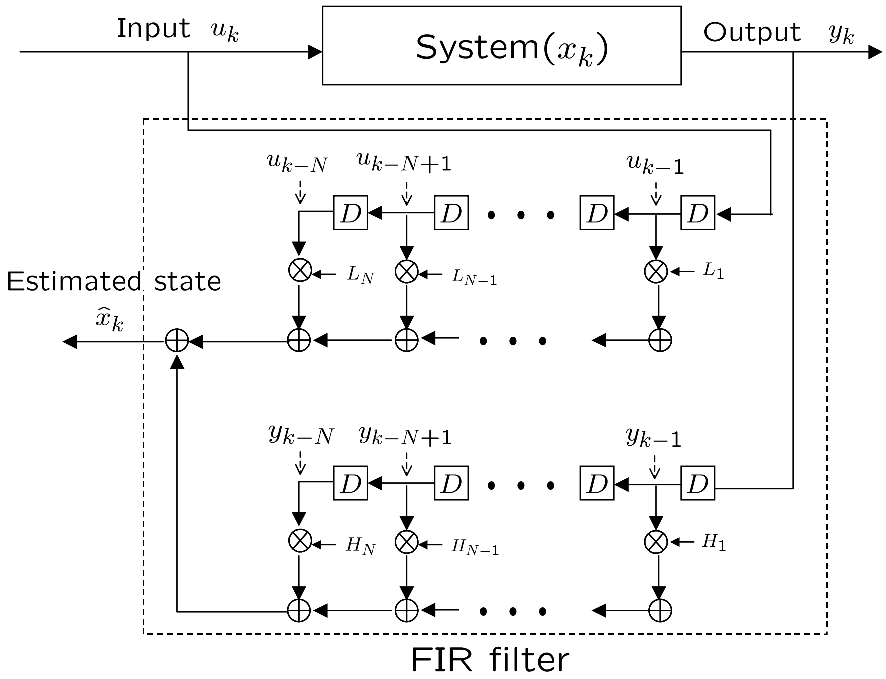

Figure 1. Block diagram of FIR filter: D is a unit delay component

for some gains $H_\cdot$ and $L_\cdot$. The basic block diagram of the FIR filter (1) is depicted in Figure 1. If a forgetting factor is employed, conventional IIR filters such as the Kalman and H_∞ filters can be approximated to the form (1) [7], which may be called a soft FIR filter if the FIR filter (1) is called a hard FIR filter. As anther method, the Kalman filter is forced to put more weights on the recent data like the FIR filter (1), if necessary, by increasing a system noise covariance [8]. However, these methods are very heuristic and how much the optimality is spoiled for the given performance criteria is not clear. In this paper, filter coefficients $H_\cdot$ and $L_\cdot$ in (1) will be computed to optimize the given performance criterion. Among linear FIR filters of the form (1), we will obtain the filter for the following performance criterion:

$$\min_{\hat{x}_k} -\frac{2}{\alpha}\log\left[\mathbf{E}e^{-\frac{\alpha}{2}e_k^Te_k}\right], \tag{2}$$

where α is a constant, $\mathbf{E}(\cdot)$ denotes the expectation, and $e_k \triangleq \hat{x}_k - x_k$ is the estimation error at time k. $-\frac{2}{\alpha}\log$ in (2) is just a scaling factor. The criterion (2) is equivalent to minimizing $\mathbf{E}\left[-\frac{2}{\alpha}\left(e^{-\frac{\alpha}{2}e_k^Te_k}-1\right)\right]$ since a logarithmic function is monotonic increasing and a constant term is not involved with the operation of the expectation. How $-\frac{2}{\alpha}\left(e^{-\frac{\alpha}{2}e_k^Te_k}-1\right)$ varies with $e_k^Te_k$ for different values of α is shown in Figure 2. Sharpness and dullness of the graph can be varied with the value of α. As α goes to zero, $-\frac{2}{\alpha}\left(e^{-\frac{\alpha}{2}e_k^Te_k}-1\right)$ reduces to $e_k^Te_k$ so that the criterion (2) is equivalent to the minimum variance one. It can be said that the criterion (2) is a general version of the minimum variance one. For $\alpha < 0$, the cost function (2) is called a risk averse criterion since large weights are put on large estimation errors and thus the large or risky estimation errors would be suppressed as much as possible. This also means that the designer is pessimistic about the estimation errors so that the filter based on this criterion will work well when large estimation errors often happen. For $\alpha > 0$, the cost function (2) is called a risk seeking criterion since large weights are put on moderate estimation errors and large estimation errors are less weighted compared with the risk averse criterion for $\alpha < 0$. It is useful when the occasional occurrence of a large estimation error is tolerable. This also means that the designer is optimistic about the estimation errors so that the filter based on this criterion will work well when estimation errors are mostly moderate. The FIR filter based on the risk averse or seeking criterion is called a risk sensitive FIR filter (RSFF).

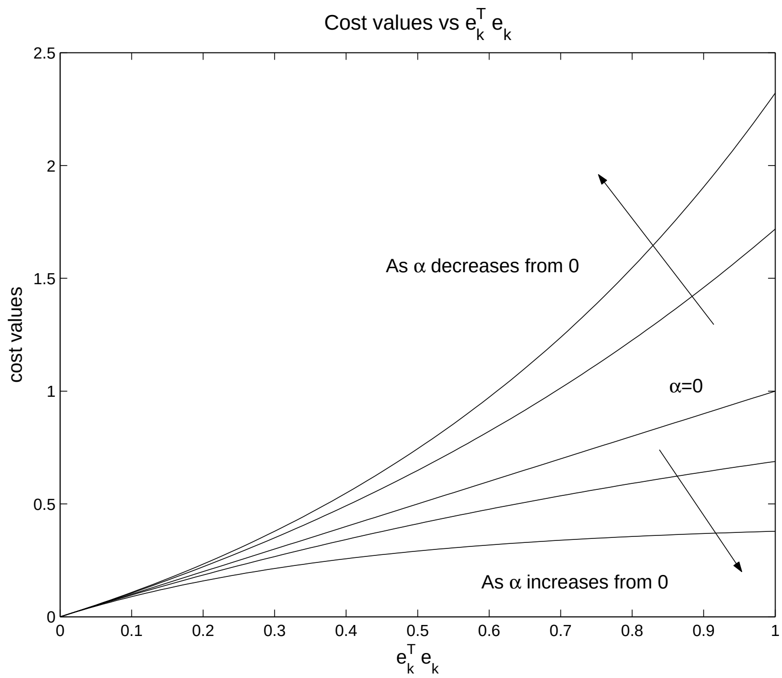

FIGURE 2. Cost functions vs $e_k^T e_k$

There have been a few results on FIR filtering for limited models or heuristic approaches. For deterministic discrete-time systems without noises, a moving horizon least-square filter of the form (1) was given in [9]. For special discrete stochastic systems without system noises, a linear FIR filter was introduced from a maximum likelihood criterion [10]. Since the system noise is not considered, the FIR filter is of a simple form and easy to derive. For general discrete-time stochastic systems, FIR filters were introduced by a modification from the Kalman filter [11] where the infinite covariance of the initial state information is difficult to handle and the efficiency of the filters is not clear. Besides, this work brings out a limitation that the system matrix is required to be nonsingular. In [12], the optimal FIR filter with the unbiased condition was given under an assumption that the system matrix is nonsingular. Even though the unbiased condition leads to an easy derivation, it may go against the optimality of the performance criterion. In [9], the FIR filter was derived without this assumption. Instead, the system noise was assumed not to exist.

To the authors' knowledge, there is no result about FIR filters for general state space models without any artificial restrictions or conditions which may prevent FIR filters from applying to real applications. In this paper, we derive FIR filters without these constraints. General systems with the system and measurement noises will be considered and the inverse of the system matrix is not required, i.e., $H.$ and $L.$ of (1) will be represented without using the inverse of the system matrix. The unbiased condition for easy derivation will not be employed during the design so that the optimality is not affected. While the existing results were based on the minimum variance or least square criteria, this paper deals with the more general performance criterion (2) which includes a minimum variance criterion with $\alpha = 0$ as mentioned before.

For a long time, robustness has been addressed for the analysis and the design of the IIR filters for state estimation. It was shown in [13, 14, 15, 16, 17] that the conventional IIR filter, i.e., the Kalman filter can diverge and have the poor performance due to model uncertainties. In order to build up robustness, robust Kalman and H_∞ filters were proposed in [18, 19]. In several works [20, 21, 22, 25, 26], it was shown through simulation and a quantitative analysis that the FIR filtering for state estimation could also be a good substitute to achieve a high degree of robustness as in the signal processing area. Through simulation, we will show that the proposed RSFF has the robustness to model uncertainties.

In Section 2, the RSFF is derived for a risk averse or seeking criterion. In Section 3, it is shown via simulation that the performance of the proposed RSFF is compared with that of the robust Kalman filter. Finally, conclusions are presented in Section 4.

2. **Risk Sensitive FIR Filters.** Consider a linear discrete-time state space model with control input:

$$x_{i+1} = Ax_i + Bu_i + Gw_i, \tag{3}$$
$$y_i = Cx_i + v_i, \tag{4}$$

where $x_i \in \Re^n$, $u_i \in \Re^l$ and $y_i \in \Re^q$ are the state, the input and the measurement, respectively. At the initial time i_0 of the system, the state x_{i_0} is a random variable with a mean $\bar{x}_{i_0}$ and a covariance P_{i_0}. The system noise $w_i \in \Re^p$ and the measurement noise $v_i \in \Re^q$ are zero-mean white Gaussian and mutually uncorrelated. These noises are uncorrelated with the initial state x_{i_0}. The covariances of w_i and v_i are denoted by Q and R, respectively. Through this paper, k denotes the current time.

The systems (3) and (4) will be represented in a batch form on the most recent time interval $[k-N, k]$, called the horizon. On the horizon $[k-N, k]$, the finite number of measurements is expressed in terms of the state x_{k-N}, the input, and the noise on the horizon as follows:

$$Y_{k-1} = \tilde{C}_N x_{k-N} + \tilde{B}_N U_{k-1} + \tilde{G}_N W_{k-1} + V_{k-1}, \tag{5}$$

where Y_{k-1}, U_{k-1}, W_{k-1} and V_{k-1} are defined as:

$$Y_{k-1} \triangleq \left[y_{k-N}^T \; y_{k-N+1}^T \; \cdots \; y_{k-1}^T\right]^T, \tag{6}$$
$$U_{k-1} \triangleq \left[u_{k-N}^T \; u_{k-N+1}^T \; \cdots \; u_{k-1}^T\right]^T, \tag{7}$$
$$W_{k-1} \triangleq \left[w_{k-N}^T \; w_{k-N+1}^T \; \cdots \; w_{k-1}^T\right]^T,$$
$$V_{k-1} \triangleq \left[v_{k-N}^T \; v_{k-N+1}^T \; \cdots \; v_{k-1}^T\right]^T,$$

and $\tilde{C}_N$, $\tilde{B}_N$ and $\tilde{G}_N$ are given by

$$\tilde{C}_N \triangleq \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N-1} \end{bmatrix}, \tag{8}$$

$$\tilde{B}_N \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ CB & 0 & \cdots & 0 & 0 \\ CAB & CB & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}B & CA^{N-3}B & \cdots & CB & 0 \end{bmatrix}, \tag{9}$$

$$\tilde{G}_N \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ CG & 0 & \cdots & 0 & 0 \\ CAG & CG & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{N-2}G & CA^{N-3}G & \cdots & CG & 0 \end{bmatrix}. \tag{10}$$

The noise term $\tilde{G}_N W_{k-1} + V_{k-1}$ in (5) can be shown to be zero-mean with the covariance Π_N given by

$$\Pi_N = \tilde{G}_N Q_N \tilde{G}_N^T + R_N, \tag{11}$$

where Q_N and R_N are defined as:

$$Q_N \triangleq \left[\mathrm{diag}(\overbrace{Q\ Q\ \cdots\ Q}^{N})\right], \tag{12}$$

$$R_N \triangleq \left[\mathrm{diag}(\overbrace{R\ R\ \cdots\ R}^{N})\right]. \tag{13}$$

The current state x_k can be represented in terms of the state x_{k-N}, the input, and the noise on the horizon as:

$$\begin{aligned} x_k &= A^N x_{k-N} + \begin{bmatrix} A^{N-1}G & A^{N-2}G & \cdots & G \end{bmatrix} W_{k-1} \\ &\quad + \begin{bmatrix} A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix} U_{k-1}, \\ &= A^N x_{k-N} + M_B U_{k-1} + M_G W_{k-1}, \end{aligned} \tag{14}$$

where M_B and M_G are given by

$$\begin{aligned} M_B &\triangleq \begin{bmatrix} A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix}, \\ M_G &\triangleq \begin{bmatrix} A^{N-1}G & A^{N-2}G & \cdots & G \end{bmatrix}. \end{aligned}$$

Now, we compute the cost function (2). Note that U_{k-1} and Y_{k-1} are known variables and x_{k-N}, W_{k-1}, and V_{k-1} are random variables. The expectation of the exponential quadratic cost function (2) will be taken over the jointly Gaussian random variables $\{x_{k-N}, W_{k-1}, V_{k-1}\}$. Since x_{k-N}, W_{k-1} and V_{k-1} are independent Gaussian random variables, their joint probability density function (pdf) $p(x_{k-N}, W_{k-1}, V_{k-1})$ can be written as:

$$p(x_{k-N}, W_{k-1}, V_{k-1}) = \frac{1}{\sqrt{(2\pi)^{n+pN+qN} D}} e^{-\frac{1}{2}J_k}, \tag{15}$$

where $D \triangleq \det P \det Q_N \det R_N$ and J_k is given by

$$J_k \triangleq (x_{k-N} - \bar{m})^T \bar{P}^{-1} (x_{k-N} - \bar{m}) + W_{k-1} Q_N^{-1} W_{k-1}^T \\ + V_{k-1} R_N^{-1} V_{k-1}^T, \tag{16}$$

with the mean $\bar{m}$ and the variance $\bar{P}$ of the random variable x_{k-N}. $\bar{m}$ and $\bar{P}$ can be computed from measured inputs and outputs on the recent horizon according to the least mean square criterion. More details can be seen in [23] where $\bar{m}$ and $\bar{P}$ are written as:

$$\bar{m} = \left(\tilde{C}_N^T \Pi_N^{-1} \tilde{C}_N\right)^{-1} \tilde{C}_N^T \Pi_N^{-1} \left(Y_{k-1} - \tilde{B}_N U_{k-1}\right),$$
$$\bar{P} = \left(\tilde{C}_N^T \Pi_N^{-1} \tilde{C}_N\right)^{-1},$$

where $\bar{P}$ always exists because of the observability condition. If V_k in (16) is replaced with $Y_{k-1} - \tilde{C}_N x_{k-N} - \tilde{B}_N U_{k-1} - \tilde{G}_N W_{k-1}$, the joint pdf of the $\{x_{k-N}, W_{k-1}, Y_{k-1}\}$ is obtained and J_k in (16) can be written as:

$$J_k = (x_{k-N} - \bar{m})^T \bar{P}^{-1} (x_{k-N} - \bar{m}) + W_{k-1}^T Q_N^{-1} W_{k-1} \\ + (\bar{Y}_{k-1} - \tilde{C}_N x_{k-N} - \tilde{G}_N W_{k-1})^T R_N^{-1} \\ \times (\bar{Y}_{k-1} - \tilde{C}_N x_{k-N} - \tilde{G}_N W_{k-1}),$$

where $\bar{Y}_{k-1} = Y_{k-1} - \tilde{B} U_{k-1}$. By using the joint pdf (15) of $\{x_{k-N}, W_{k-1}, Y_{k-1}\}$, the exponential quadratic cost functions (2) can be computed as:

$$\mathbf{E}\left[e^{-\frac{\alpha}{2}(\hat{x}_k - x_k)^T(\hat{x}_k - x_k)}\right] \\ = K_1 \int \exp\left[-\frac{1}{2}\bar{J}_k\right] dx_{k-N} dW_{k-1} \\ = K_2 \exp\left[-\frac{1}{2} \min_{x_{k-N}, W_{k-1}} \bar{J}_k\right], \tag{17}$$

for some constants K_1 and K_2, where $\bar{J}_k \triangleq J_k + \alpha(\hat{x}_k - x_k)^T(\hat{x}_k - x_k)$ and the second equality comes from the fact that J_k is quadratic with respect to all integration variables and the integral of an exponential quadratic function from negative infinity to positive infinity is easily computed using the formula:

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^T \Sigma^{-1} x} dx = \sqrt{(2\pi)^N \det(\Sigma)}, \quad \Sigma \in \Re^{N \times N}. \tag{18}$$

By using (5) and (14), $\bar{J}_k$ in (17) can be written as:

$$\bar{J}_k = \left(x_{k-N} - T\tilde{Y}_{k-1}\right)^T \bar{P}^{-1} \left(x_{k-N} - T\tilde{Y}_{k-1}\right) \\ + W_{k-1}^T Q_N^{-1} W_{k-1} \\ + \left(\tilde{Y}_{k-1} - \tilde{C}_N x_{k-N} - \tilde{G}_N W_{k-1}\right)^T R_N^{-1} \\ \times \left(\tilde{Y}_{k-1} - \tilde{C}_N x_{k-N} - \tilde{G}_N W_{k-1}\right) \\ + \alpha\left(\hat{x}_k - A^N x_{k-N} - M_B U_{k-1} - M_G W_{k-1}\right)^T \\ \times \left(\hat{x}_k - A^N x_{k-N} - M_B U_{k-1} - M_G W_{k-1}\right), \tag{19}$$

where $T = \left(\tilde{C}_N^T \Pi_N^{-1} \tilde{C}_N\right)^{-1} \tilde{C}_N^T \Pi_N^{-1}$ and $\tilde{Y}_{k-1} = Y_{k-1} - \tilde{B} U_{k-1}$. Note that $\bar{J}_k$ in (19) is quadratic with respect to variables x_{k-N}, W_{k-1}, $\hat{x}_k$, $\tilde{Y}_{k-1}$ and U_{k-1}. $\bar{J}_k$ in (19) can be written in a compact form as:

$$\bar{J}_k = \Lambda_k^T \Xi \Lambda_k, \tag{20}$$

where Λ_k and Ξ are given by

$$\Xi = \begin{bmatrix} (1,1) & (1,2) & \alpha A^{NT} & (1,4) & -\alpha A^{NT} M_B \\ * & (2,2) & -\alpha M^T & G^T & \alpha M_G^T M_B \\ * & * & \alpha I & 0 & -\alpha M_B \\ * & * & * & (4,4) & 0 \\ * & * & * & * & \alpha M_B^T M_B \end{bmatrix},$$

$$\begin{aligned}
(1,1) &= \bar{P}^{-1} + \tilde{C}_N^T R_N^{-1} \tilde{C}_N + \alpha A^{NT} A^N, \\
(1,2) &= \tilde{C}_N^T \tilde{G}_N - \alpha A^{NT} M_G, \\
(1,4) &= -\tilde{C}_N^T - \bar{P}^{-1} T, \\
(2,2) &= Q_N^{-1} + \tilde{G}_N^T \tilde{G}_N + \alpha M_G^T M_G, \\
(4,4) &= R_N^{-1} + T^T \bar{P}^{-1} T,
\end{aligned}$$

$$\Lambda_k = \begin{bmatrix} x_{k-N} \\ W_{k-1} \\ \hat{x}_k \\ \tilde{Y}_{k-1} \\ U_{k-1} \end{bmatrix}.$$

Now, we are in a position to find out $\hat{x}_k$ to optimize $\bar{J}_k$ in (20) according to the criterion (2). First, we consider the case of $\alpha < 0$, which is related to the risk averse criterion.

2.1. **Risk averse criterion.** For the case of $\alpha < 0$, the optimization problem (2) reduces to the following one:

$$\min_{\hat{x}_k} \mathbf{E}\left[e^{-\frac{\alpha}{2}(\hat{x}_k - x_k)^T(\hat{x}_k - x_k)}\right]. \tag{21}$$

According to the relation (17), we can change the problem (21) to one of finding the optimal values optimizing a quadratic cost function. The final problem to solve can be thus formulated as follows:

$$\max_{\hat{x}_k} \min_{x_{k-N}, W_{k-1}} \bar{J}_k, \tag{22}$$

where $\bar{J}_k$ is given by (20). Note that minimization problems are changed to maximized ones if the sign in front of a cost function is switched. In order to obtain the solution to minimize $\bar{J}_k$ in (22) with respect to x_{k-N} and W_{k-1}, and maximize it with respect to $\hat{x}_k$, we introduce a useful result.

Lemma 2.1. [24] *Consider a cost function $J(a,b,y)$ given by*

$$J(a,b,y) = \begin{bmatrix} a \\ b \\ y \end{bmatrix}^T \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{12}^T & M_{22} & M_{23} \\ M_{13}^T & M_{23}^T & M_{33} \end{bmatrix} \begin{bmatrix} a \\ b \\ y \end{bmatrix}, \tag{23}$$

where a and b are vector variables and y is a given vector constant. When the following conditions are satisfied:

$$M_{11} > 0, \quad M_{22} - M_{12}^T M_{11}^{-1} M_{12} < 0, \tag{24}$$

the optimal values a and b minimizing $J(a,b,y)$ with respect to a and maximizing $J(a,b,y)$ with respect to b exist and are given by

$$\begin{bmatrix} a^* \\ b^* \end{bmatrix} = - \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix}^{-1} \begin{bmatrix} M_{13} \\ M_{23} \end{bmatrix} y. \tag{25}$$

Besides, a^* and b^* have the property that

$$J(a, b^*, y) \geq J(a^*, b^*, y) \geq J(a^*, b, y), \tag{26}$$

for any a and b.

If a, b, y, M_{11}, M_{12}, M_{13}, M_{22}, M_{23} and M_{33} in Lemma 2.1 are given by the following matrices and vectors:

$$M_{11} = \begin{bmatrix} (1,1) & (1,2) \\ (1,2)^T & (2,2) \end{bmatrix}, \quad M_{12} = \begin{bmatrix} \alpha A^{NT} \\ -\alpha M_G^T \end{bmatrix}, \tag{27}$$

$$M_{13} = \begin{bmatrix} (1,4) & -\alpha A^{NT} M_B \\ G^T & \alpha M_G^T M_B \end{bmatrix}, \quad M_{22} = \alpha I, \tag{28}$$

$$M_{23} = \begin{bmatrix} 0 & \alpha M_B \end{bmatrix}, \quad M_{33} = \begin{bmatrix} (4,4) & 0 \\ 0 & \alpha M_B^T M_B \end{bmatrix}, \tag{29}$$

$$a = \begin{bmatrix} x_{k-N} \\ W_{k-1} \end{bmatrix}, \quad b = \hat{x}_k, \quad y = \begin{bmatrix} \bar{Y}_{k-1} \\ U_{k-1} \end{bmatrix}, \tag{30}$$

the solution (25) gives us the optimal one with respect to the cost function (22), which minimize $\bar{J}_k$ in (22) with respect to x_{k-N}, and W_{k-1}, and maximize it with respect to $\hat{x}_k$.

Now, we check the existence of the solution according to the condition (24).

$$\begin{bmatrix} \bar{P}^{-1} + \tilde{C}_N^T \tilde{C}_N & \tilde{C}_N^T \tilde{G}_N \\ \tilde{C}_N^T \tilde{G}_N & Q_N^{-1} + \tilde{G}_N^T \tilde{G}_N \end{bmatrix} + \alpha \begin{bmatrix} A^{NT} \\ M_G^T \end{bmatrix} \begin{bmatrix} A^N & M_G \end{bmatrix} > 0. \tag{31}$$

For a given value α, it is easy to check whether the condition (31) is met. We have only to compute the eigenvalues of the left side of the inequality (31). If all eigenvalues are positive, the inequality (31) is guaranteed to be satisfied. If $\alpha = 0$, the inequality (31) always holds.

What we have done until now is summarized in the following theorem.

Theorem 2.1. *Suppose that α satisfies the inequality (31). For the risk averse criterion (2) in case of $\alpha < 0$, the risk sensitive FIR filter of the form (1) is given by (25), where M_{11}, M_{12}, M_{22}, M_{13} and M_{23} are defined in (27)-(30).*

Next, we consider the case of $\alpha > 0$, which is related to the risk seeking criterion.

2.2. **Risk seeking criterion.** For the case of $\alpha > 0$, we shall solve the following optimization problem:

$$\max_{\hat{x}_k} \mathbf{E}\left[e^{-\frac{\alpha}{2}(\hat{x}_k - x_k)^T(\hat{x}_k - x_k)} \right]. \tag{32}$$

According to the relation (17), we can change the problem (32) to one of finding the minimum value of a quadratic cost function. The problem to solve can be thus formulated as follows:

$$\min_{\hat{x}_k, x_{k-N}, W_{k-1}} \bar{J}_k \tag{33}$$

where $\bar{J}_k$ is given by (20). In case of $\alpha > 0$, the problem is much easier since only minimization is required, not mixing with maximization as in the risk averse criterion.

In order to obtain the solution to minimize $\bar{J}_k$ with respect to $\hat{x}_k$, x_{k-N} and W_{k-1}, we introduce a useful result.

Lemma 2.2. *If the cost function $J(a, y)$ is given by*

$$J(a,y) = \begin{bmatrix} a \\ y \end{bmatrix}^T \begin{bmatrix} N_{11} & N_{12} \\ N_{12}^T & N_{22} \end{bmatrix} \begin{bmatrix} a \\ y \end{bmatrix}, \tag{34}$$

where $N_{11} > 0$, a is a vector variable, and y is a given vector constant, then the optimal value minimizing $J(a, y)$ is given by

$$a_{opt} = -N_{11}^{-1} N_{12} y. \tag{35}$$

If a, y, N_{11} and N_{12} are given by the following matrices or vectors:

$$a = \begin{bmatrix} x_{k-N} \\ W_{k-1} \\ \hat{x}_k \end{bmatrix}, \quad y = \begin{bmatrix} \bar{Y}_{k-1} \\ U_{k-1} \end{bmatrix}, \tag{36}$$

$$N_{11} = \begin{bmatrix} (1,1) & (1,2) & \alpha A^{NT} \\ * & (2,2) & -\alpha M_G^T \\ * & * & \alpha I \end{bmatrix}, \tag{37}$$

$$N_{12} = \begin{bmatrix} (1,4) & -\alpha A^{NT} M_B \\ G^T & \alpha M_G^T M_B \\ 0 & 0 \end{bmatrix}, \tag{38}$$

$$N_{22} = \begin{bmatrix} (4,4) & 0 \\ 0 & \alpha M_B^T M_B \end{bmatrix}, \tag{39}$$

then the solution (35) gives us the optimal one with respect to the cost function (33), which minimizes $\bar{J}_k$ in (33) with respect to $\hat{x}_k$, x_{k-N} and W_{k-1}. It is noted that N_{11} in (37) is nonsingular since it is positive definite.

What we have done in this section can be summarized in the following theorem.

Theorem 2.2. *For the risk seeking criterion (2) in case of $\alpha > 0$, the RSFF of the form (1) is given by (35), where a, y, N_{11}, N_{12} and N_{22} are defined in (36)-(39).*

3. **Simulation Results.** To demonstrate the validity of the proposed RSFF, the numerical example on the model of an F-404 engine is presented via simulation studies. This model is a discrete-time version sampled by 0.05 sec from a continuous one.

As mentioned in Introduction, IIR filters can have drawbacks such as a slow convergence and divergence. In this section, it is shown via simulation that the RSFF can overcome these problems due to FIR structure. The uncertain model is represented as:

$$x_{i+1} = \begin{bmatrix} 0.931 + \delta_k & 0 & 0.111 \\ 0.008 + 0.05\delta_k & 0.98 + 1.11\delta_k & -0.017 \\ 0.014 & 0 & 0.895 + \delta_k \end{bmatrix} x_i + \begin{bmatrix} 0.051 \\ 0.049 \\ 0.048 \end{bmatrix} w_i,$$

$$y_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x_i + v_i,$$

where $\mathbf{E}[w_i^2] = 0.002$, $\mathbf{E}[v_i v_i^T] = 0.002 I_2$, and the parameter δ_k is given by

$$\delta_k = \begin{cases} 1, & 50 \le k \le 100, \\ 0, & \text{otherwise}. \end{cases} \tag{40}$$

α is taken as -1. To begin with, we check the impulse responses of the RSFF and the robust Kalman IIR filter. Figure 3 shows that the proposed RSFF has the finite duration of impulse responses while the robust Kalman filter has the infinite duration. This implies that the RSFF guarantees a fast convergence to a normal state within a finite time when temporary uncertainties happen.

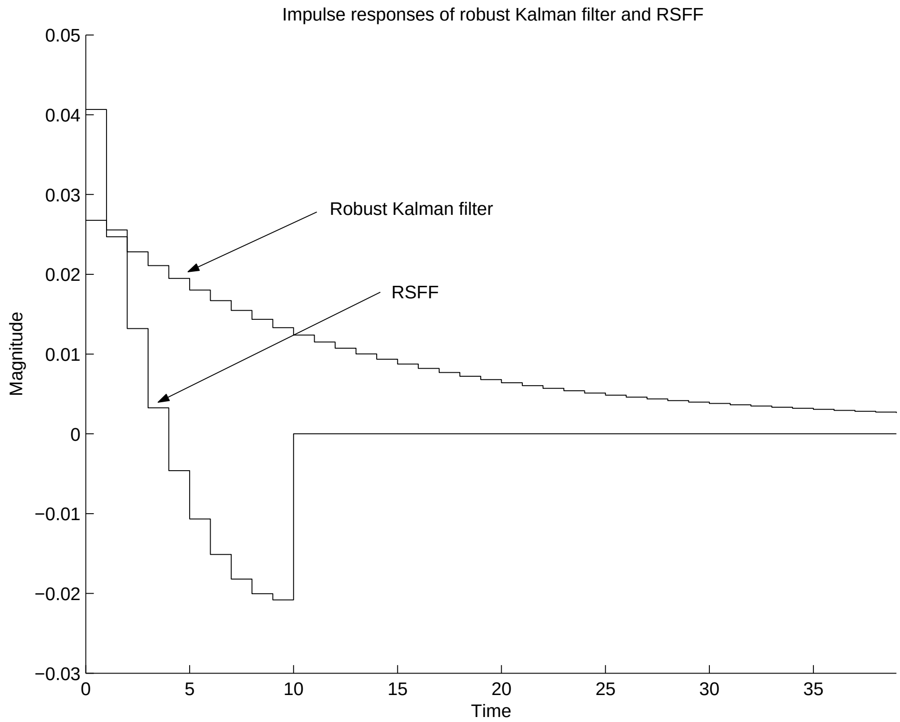

FIGURE 3. Impulse responses of robust Kalman filter and RSFF

Figure 4 compares how the Kalman filter and RSFF respond to temporarily modeling uncertainties. The horizon size N of the RSFF is set to 10. The figure shows that the estimation errors of the RSFF are remarkably smaller than that of the robust Kalman filter on the interval where modeling uncertainties exist. Actually, poles of the robust Kalman filter is close to a unit circle. $0.8893 \pm 0.0225i$ and 0.9712. Due to these poles and uncertainties, the estimation error blows up between 50 and 100 while only a little deviation is shown in the RSFF. In addition, it is shown that the convergence of estimation errors of the RSFF is much faster than that of the Kalman filter after temporary modeling uncertainties disappear. Therefore, it can be seen that the suggested RSFF are more robust than the robust Kalman filter when applied to systems with model parameter uncertainties. Actually, the good performance of the proposed RSFF is significant when the optimal IIR filter is slow.

4. **Conclusions.** In this paper, we introduced a risk averse or seeking performance criterion for state estimation, which is represented as the expectation of an exponential quadratic cost function. Based on this performance criterion, a risk sensitive FIR filter (RSFF) was proposed for a general stochastic discrete-time state space model. The proposed RSFF is linear with the most recent finite measurements and inputs. The RSFF was obtained to optimize the risk averse or seeking performance criterion, together with prior constraints such as linearity and FIR structure. Nonsingularity of the system matrix and the unbiased condition are not required. System and measurement noises are considered simultaneously. It is shown via simulation that, due to FIR structure, the RSFF has a better estimation ability for temporary modelling uncertainties compared with a conventional robust IIR filter, i.e., robust Kalman filter.

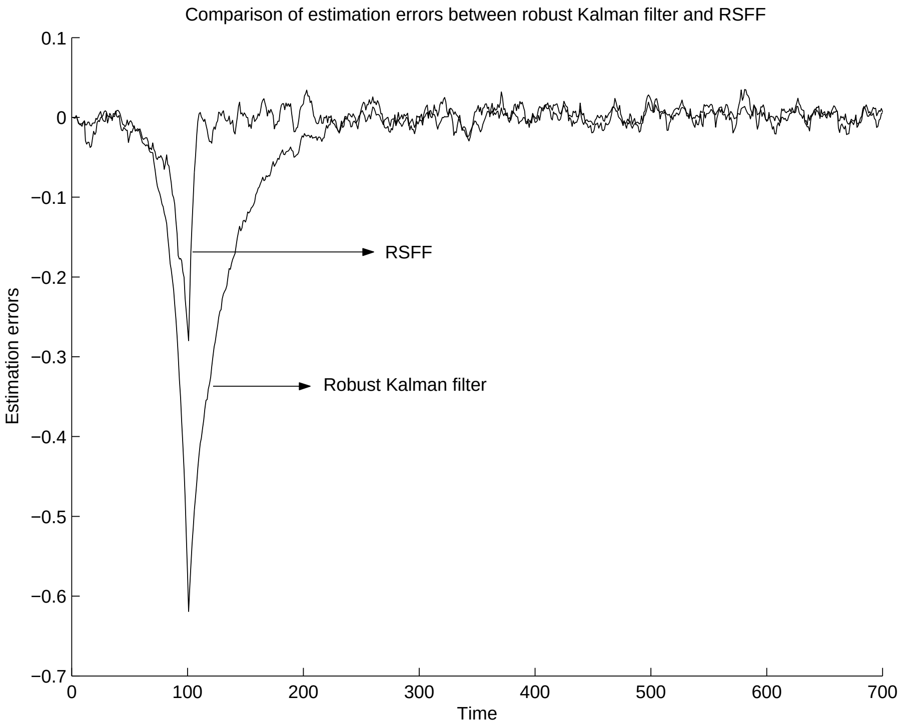

FIGURE 4. Comparison between robust Kalman filter and RSFF

When the IIR filter has a slow response, the proposed RSFF could be a good substitute to achieve a fast response with robustness.

Acknowledgment. This research work was supported by the Basic Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (ID:2010-0024835). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] R. E. Kalman and R. S. Bucy, A new approach to linear filtering and prediction problems, *Trans. on ASME J. of Basic Engr.*, vol.82, pp.35-45, 1960.

[2] K. M. Nagpal and P. P. Khargonekar, Filtering and smoothing in an H_∞ setting, *IEEE Trans. on Automat. Contr.*, vol.36, pp.152-166, 1991.

[3] M. A. Alcorta, M. Basin, J. J. Maldonado and S. G. Anguiano, Sub-optimal risk-sensitive filtering for third degree polynomial stochastic systems, *ICIC Express Letters*, vol.2, no.4, pp.371-377, 2008.

[4] M. A. Alcortata-Garcia, M. Basin and Y. G. A. Sanchez, Risk-sensitive approach to optimal filtering and control for linear stochastic systems, *International Journal of Innovative Computing, Information and Control*, vol.5, no.6, pp.1599-1614, 2009.

[5] B. Song, S. Xu and Y. Zou, Non-fragile H_∞ filtering for uncertain stochastic time-delay systems, *International Journal of Innovative Computing, Information and Control*, vol.5, no.8, pp.2257-2266, 2009.

[6] Y. Liu and W. Wand, Fuzzy H_∞ filtering for nonlinear stochastic systems with missing measurements, *ICIC Express Letters*, vol.3, no.3(B), pp.739-744, 2009.

[7] T. W. Yang, J. H. Lee, K. Y. Lee and K. M. Sung, On robust Kalman filtering with forgetting factor for sequential speech analysis, *Signal Processing*, vol.63, pp.151-156, 1997.

[8] J. B. Burl, *Linear Optimal Control*, Addison-Wesley, CA, USA, 1998.

[9] K. V. Ling and K. W. Lim, Receding horizon recursive state estimation, *IEEE Trans. on Automat. Contr.*, vol.44, no.9, pp.1750-1753, 1999.

[10] A. H. Jazwinski, Limited memory optimal filtering, *IEEE Trans. on Automat. Contr.*, vol.13, pp.558-563, 1968.

[11] W. H. Kwon, P. S. Kim and P. Park, A receding horizon Kalman FIR filter for discrete time-invariant systems, *IEEE Trans. on Automat. Contr.*, vol.44, no.9, pp.1787-1791, 1999.

[12] W. H. Kwon, P. S. Kim and S. Han, A receding horizon unbiased FIR filter for discrete-time state space models, *Automatica*, vol.38, no.3, pp.545-551, 2002.

[13] H. Heffes, The effect of erroneous models on the Kalman filter response models, *IEEE Trans. on Automat. Contr.*, vol.11, no.3, pp.541-543, 1966.

[14] C. F. Price, An analysis of the divergence problem in the Kalman filter, *IEEE Trans. on Automat. Contr.*, vol.12, pp.699-702, 1968.

[15] R. J. Fitzgerald, Divergence of the Kalman filter, *IEEE Trans. on Automat. Contr.*, vol.16, pp.736-747, 1971.

[16] M. Toda and R. V. Patel, Bounds on estimation errors of discrete-time filters under modeling uncertainty, *IEEE Trans. on Automat. Contr.*, vol.25, pp.1115-1121, 1980.

[17] S. Sangsuk-Iam and T. E. Bullock, Analysis of discrete-time Kalman filtering under incorrect noise covariances, *IEEE Trans. on Automat. Contr.*, vol.35, no.12, pp.1304-1309, 1990.

[18] L. Xie, Y. C. Soh and C. E. de Souza, Robust Kalman filtering for uncertain discrete-time systems, *IEEE Trans. on Automat. Contr.*, vol.39, pp.1310-1313, 1994.

[19] C. E. de Souza and L. Xie, Robust H_∞ filtering, *Control and Dynamic Systems*, vol.65, pp.323-377, 1994.

[20] A. Alessandri, M. Baglietto and G. Battistelli, Receding-horizon estimation for discrete-time linear systems, *IEEE Trans. on Automat. Contr.*, vol.48, no.3, pp.473-478, 2003.

[21] A. Alessandri, M. Baglietto and G. Battistelli, On estimation error bounds for receding horizon filters using quadrature boundedness, *IEEE Trans. on Automat. Contr.*, vol.49, no.8, pp.1350-1355, 2004.

[22] C. V. Rao, J. B. Rawlings and J. H. Lee, Contrained linear state estimation – A moving horizon approach, *Automatica*, vol.37, pp.1619-1628, 2001.

[23] S. H. Han, P. S. Kim and W. H. Kwon, Receding horizon FIR filter with estimated horizon initial state and its application to aircraft engine systems, *Proc. of IEEE Int. Conf. Contr. Applications*, HI, USA, pp.33-38, 1999.

[24] B. Hassibi, A. H. Sayed and T. Kailath, *Indefinite Quadratic Estimation and Control (SIAM Studies in Applied and Numerical Mathematics)*, Society for Industrial Mathematics, 1999.

[25] S. Han and W. H. Kwon, L_2-E FIR smoothers for deterministic discrete-time state space signal models, *IEEE Trans. on Automat. Contr.*, vol.52, no.5, pp.927-932, 2007.

[26] B. K. Kwon, S. Han, O. K. Kwon and W. H. Kwon, Minimum variance FIR smoothers for discrete-time state space models, *IEEE Signal Processing Letters*, vol.14, no.8, pp.557-560, 2007.