General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)
MODELING ERROR ANALYSIS OF STATIONARY LINEAR DISCRETE-TIME FILTERS

Rajnikant Patel and Mitsuhiko Toda

Ames Research Center
Moffett Field, California 94035

February 1977
The performance of Kalman-type, linear, discrete-time filters in the presence of modeling errors is considered. The discussion is limited to stationary performance, and bounds are obtained for the performance index, the mean-squared error of estimates for suboptimal and optimal (Kalman) filters. The computation of these bounds requires information on only the model matrices and the range of errors for these matrices. Consequently, a designer can easily compare the performance of a suboptimal filter with that of the optimal filter, when only the range of errors in the elements of the model matrices is available.
Abstract. The performance of Kalman-type, linear, discrete-time filters in the presence of modeling errors is considered. The discussion is limited to stationary performance, and bounds are obtained for the performance index, the mean-squared error of estimates for suboptimal and optimal (Kalman) filters. The computation of these bounds requires information on only the model matrices and the range of errors for these matrices. Consequently, a designer can easily compare the performance of a suboptimal filter with that of the optimal filter, when only the range of errors in the elements of the model matrices is available.

1. Introduction. One of the problems arising in the application of the minimum variance optimal filter of Kalman and Bucy [8] is that a design based on imperfect knowledge of the system configuration and noise statistics often results in poor performance. Thus, there has been considerable research on the effect of modeling errors on filter performance [3,5,6,10,11]. In particular, errors in prior information on state statistics and noise covariances [6,10,11] and in system models [3,5] have been considered.

This paper is concerned with providing bounds on the performance for suboptimal as well as optimal discrete-time filters based on information...
about the range of modeling errors. These results are useful from the practical point of view, as a designer often has information on the range of modeling errors rather than a precise knowledge of the modeling errors. In this paper we limit the discussion to stationary conditions and obtain performance bounds for discrete-time filters for two types of errors: noise covariance errors and system configuration errors. Such bounds were obtained for continuous-time filters in an earlier paper [12]. The results reported here and in [12] are quite different from those of earlier work in that we do not limit the discussion to small-scale (differential) analysis as in [5], nor do we assume sign definiteness of the covariance errors as in [11].

This paper is organized as follows: the problem is formulated in Sec. 2; in Sec. 3, two general bounds are obtained. These bounds are analyzed in Sec. 4 to obtain practical expressions when the system configuration is assumed to be known and, in Sec. 5 when the noise covariances are assumed to be known. The results in Sec. 4 and 5 are illustrated by an example in Sec. 6. Section 7 concludes with some remarks on the results and comments about future research.

2. Problem Statement. Consider a time-invariant process described by

\[ x(k + 1) = Ax(k) + Gw(k) \]  \hspace{1cm} (1)

where the \( n \)-dimensional state vector \( x(k + 1) \) is measured by an \( m \)-dimensional vector \( y(k + 1) \):

\[ y(k + 1) = Hx(k + 1) + v(k + 1) \]  \hspace{1cm} (2)
The 1-dimensional process noise \( w(k) \) and the m-dimensional measurement noise \( v(k + 1) \) are assumed to be mutually independent Gaussian noises with zero mean and

\[
E[w(k)w(j)^T] = Q\delta_{k,j}, \quad Q \geq 0^t
\]

\[
E[v(k)v(j)^T] = R\delta_{k,j}, \quad R > 0^t
\]

where \( \delta_{k,j} \) denotes the Kronecker delta. The system matrices \( A, G, \) and \( H \) are assumed to have appropriate dimensions, and \( A \) is assumed to be a convergent matrix.

The optimal estimates \( \hat{x}(k + 1|k + 1) \) that minimizes

\[
J(k + 1) = E \left[ \| \hat{x}(k + 1|k + 1) - x(k + 1) \|^2 \right]^t
\]

with observations \( [y(0), y(1), \ldots, y(k + 1)] \), are given by [8]:

\[
\hat{x}(k + 1|k + 1) = A\hat{x}(k|k) + K(k + 1)[y(k + 1) - HA\hat{x}(k|k)]
\]

where \( K(k + 1) \) is the Kalman gain matrix specified by the recursive relations:

\[
K(k + 1) = P_o(k + 1|k)H^T[H\hat{x}(k|k)H^T + R]^{-1}
\]

\[
P_o(k + 1|k) = AP_o(k|k)A^T + GQG^T
\]

and

\[
P_o(k + 1|k + 1) = [I_n - K(k + 1)H]P_o(k + 1|k)
\]

with \( \hat{x}(0) \) and \( P_o(0) \) provided by prior information on \( x(0) \) and \( I_n \), the \( n \times n \) identity matrix. The minimized index \( J(k + 1) \) is given by

___

\( ^t \)A symmetric matrix \( W \) is denoted as \( W > 0 \) \((W \geq 0)\) when \( W \) is positive definite \((\text{semidefinite})\). Also, \( W > Z \) \((W \geq Z)\) denotes \( W - Z > 0 \) \((W - Z \geq 0)\).

\( \| \mathbf{w} \| = \left[ \sum_{i,j} |W_{ij}|^2 \right]^{1/2} \)

denotes the Euclidean norm of a vector \( \mathbf{w} \) (matrix \( W \)).
Min \( J(k + 1) = \text{tr}[P_0(k + 1|k + 1)] \) \hspace{1cm} (9)

where \( \text{tr}(W) \) denotes the trace of a square matrix \( W \).

We assume that complete information on the system matrices \( A \) and \( H \) and the noise covariances \( Q \) and \( R \) is not available, and that the optimal estimator is replaced by a suboptimal estimator based on a model:

\[
\hat{x}_M(k + 1|k + 1) = A_M\hat{x}_M(k|k) + K_M(k + 1)[y(k + 1) - H_MA_M\hat{x}_M(k|k)] \quad (10)
\]

where \( A_M = A + AA \) and \( H_M = H + AH \) are model representations of \( A \) and \( H \), respectively, and \( AA \) and \( AH \) denote modeling errors. The gain \( K_M(k + 1) \) can be computed from the model matrices using the following relations:

\[
K_M(k + 1) = P_M(k + 1|k)H_M^T[H_MP_M(k + 1|k)H_M^T + R_M]^{-1} \quad (11)
\]

\[
P_M(k + 1|k) = A_MP_M(k|k)A_M^T + GQG^T \quad (12)
\]

and

\[
P_M(k + 1|k + 1) = \left[I_n - K_M(k + 1)H_M\right]P_M(k + 1|k) \quad (13)
\]

where \( Q_M = Q + \Delta Q \) and \( R_M = R + \Delta R \) are the model representations of \( Q \) and \( R \), respectively. The mean-squared error of this estimate is expressed by

\[
E\left[\|\hat{x}_M(k + 1) - x(k + 1)\|^2\right] = \text{tr}[P(k + 1)] \quad (14)
\]

In (14) and in the remainder of this section, the index \( (j|j) \) is denoted by \( (j) \) wherever appropriate. The covariance matrix \( P(k + 1) \) is described by [11]

\[
P(k + 1) = L(k)A_MP(k)A_M^TL(k)^T + \Delta C(k)V(k)A_M^TL(k)^T + L(k)A_MV(k)T\Delta C(k)^T
\]

\[
+ \Delta C(k)U(k)\Delta C(k)^T + [L(k) + K_M(k)\Delta H]GQG^T[L(k) + K_M(k)\Delta H]^T
\]

\[
+ K_M(k)R_KM(k)^T \quad (15)
\]
\[ V(k + 1) = AV(k)A^T L(k) + AU(k)\Delta C(k)^T - GQG^T[L(k) + K_M(k)\Delta H]^T \]  
\[ U(k + 1) = AU(k)A^T + GQG^T \]  

where

\[ L(k) \equiv I_n - K_M(k)H_M \]  
\[ \Delta C(k) \equiv L(k)\Delta A - K_M(k)\Delta HA_M + K_M(k)\Delta AA \]  
\[ P(k) \equiv E[\bar{x}(k)\bar{x}(k)^T] \]  
\[ V(k) \equiv E[x(k)\bar{x}(k)^T] \]  
\[ U(k) \equiv E[x(k)x(k)^T] \]  
\[ \bar{x}(k) \equiv \hat{x}_M(k) - x(k) \]  
\[ \bar{x}(0) \equiv E[x(0)] \]

Since the modeling error matrices \( \Delta A, \Delta H, \Delta Q, \) and \( \Delta R \) are generally not known exactly, (15) to (17) cannot be solved to obtain \( \text{tr}[P(k + 1)] \), the performance of the suboptimal filter. However, a designer usually has estimates of the magnitude of errors in the model matrices, e.g., \( |\Delta A_{ij}| \) in the \((i,j)\)th element of \( A_M \). Therefore, it is reasonable to obtain upper bounds for \( \text{tr}[P(k + 1)] \) based on such estimates. It is also helpful to obtain lower bounds for \( \text{tr}[P_o(k + 1)] \) as functions of the estimates so that the designer can evaluate the performance degradation, \( \text{tr}[P(k + 1)] - \text{tr}[P_o(k + 1)] \), which he should expect with the possible modeling errors. Such bounds are obtained in later sections for stationary conditions.

3. Performance Analysis for Stationary Conditions. The covariances of estimation errors are constant matrices for stationary conditions, i.e., in (6) to (8), \( P_o(k + 1|k) = P_o(k|k - 1) \) and \( P_o(k + 1|k + 1) = P_o(k|k) \).
and, in (15) to (17), \( P(k + 1) = P(k) \), \( V(k + 1) = V(k) \), and
\( U(k + 1) = U(k) \) for stationary conditions. Hence the stationary filtering error covariance matrix for the suboptimal filter is denoted by \( P \) and the stationary filtering and (one-step) prediction error covariances for the optimal filter are denoted by \( P_F \) and \( P_P \), respectively. In this section, general expressions for an upper bound for \( \text{tr}(P) \) and a lower bound for \( \text{tr}(P_F) \) are obtained. These are specialized in later sections to obtain more practical expressions.

In the sequel, the notation \( W \otimes Z \) is used for the Kronecker product of matrices \( W \) and \( Z \). The column string of an \( n \times n \) matrix \( W \), denoted by \( \text{cs}(W) \), is defined by the following \( n^2 \)-dimensional column vector:

\[
\text{cs}(W) = [w_{11} \ldots w_{n1}, w_{12} \ldots w_{n2}, \ldots, w_{1n} \ldots w_{nn}]^T
\]

where \( w_{jk} \) is the \((j,k)\)th element of matrix \( W \). Note that
\[
\text{tr}(W) = [\text{cs}(I_n)]^T \text{cs}(W) \tag{20}
\]
and
\[
\| \text{cs}(W) \| = \| W \| \tag{21}
\]

**Theorem 1**: An upper bound for \( \text{tr}(P) \) is given by
\[
\text{tr}(P) \leq \tau_1 = i^T M^{-1} b + \| (M^{-1})^T i \| \| D \| \tag{22}
\]
where
\[
M \equiv I_{n^2} - L A_M \otimes I A_M \tag{23}
\]
\[
L \equiv I_n - K H_M \tag{24}
\]
\[
i \equiv \text{cs}(I_n)
\]
\[
b \equiv \text{cs}(B), \quad B \equiv LGQ_M^T L^T + K R K_M^T
\]

\[ D = L_{MG}G^{T}K_{M}^{T} + K_{M}AHG_{Q}G^{T}(L + K_{M}AH)^{T} + ACV_{M}^{T}L_{M}^{T} + LA_{V}^{T}A^{T} + ACUAC^{T} - (L + K_{M}AH)G_{Q}G^{T}(L + K_{M}AH)^{T} - K_{M}ARK_{M} \]  
\[ \Delta C = L\Delta A - K_{M}\Delta HA_{M} + K_{M}\Delta H\Delta A \]  

**Remark:** Note that in the expression for \( \nu_{1} \), only \( D \) contains the error matrices \( \Delta A, \Delta H, \Delta Q, \) and \( \Delta R \). In Secs. 4 and 5, the term \( \| D \| \) is analyzed further to obtain more explicit expressions.

**Proof of Theorem 1:** Equation (15) can be written as
\[ P = LA_{M}P(LA_{M})^{T} + B + D \]
and, in Kronecker form, as
\[ \text{cs}(P) = [(LA_{M}) \otimes (LA_{M})]\text{cs}(P) + \text{cs}(B) + \text{cs}(D) \]
Recalling (20), we obtain
\[ \text{tr}(P) = i^{T}\text{cs}(P) = i^{T}M^{-1}b + i^{T}M^{-1}\text{cs}(D) \leq i^{T}M^{-1}b + \| (M^{-1})^{T}i \| \| \text{cs}(D) \| \]
where the Schwartz inequality was used to obtain the inequality. Noting (21), we obtain (22), thereby completing the proof.

Next, we obtain a lower bound for the filtering error covariance matrix of the optimal filter.

**Theorem 2:** A lower bound for \( \text{tr}(PF) \) is given by
\[ \text{tr}(PF) \geq \nu_{2} \equiv \frac{-a_{2} + \sqrt{a_{2}^{2} + 4a_{1}a_{3}}}{2a_{1}} \]  
\[ a_{1} \equiv \| A^{T}H^{T}K_{M}^{-1}H \|_{S}^{\frac{S}{S}} \]  
\[ \| W \|_{S}^{\frac{S}{S}} = \max[\lambda(WW^{T})]^{1/2} \] denotes the spectral norm (maximum singular value) of a matrix \( W \); \( \max \lambda(W)[\min \lambda(W)] \) denotes the maximum [minimum] eigenvalue of a symmetric matrix \( W \).
\[
a_2 \equiv n + \text{tr}(H^T R^{-1} H G Q G^T) - \sum_{i=1}^{n} |\lambda_i(A)|^2
\]  
(30)

and

\[
a_3 \equiv \text{tr}(G Q G^T)
\]  
(31)

Proof of Theorem 2: From (6) to (8), we obtain

\[
I_n + H^T R^{-1} H (A P F A^T + G Q G^T) = P_F^{-1} (A P F A^T + G Q G^T)
\]

Taking the trace of both sides yields

\[
n + \text{tr}[H^T R^{-1} H (A P F A^T + G Q G^T)] = \text{tr}(P_F^{-1} A P F A^T) + \text{tr}(P_F^{-1} G Q G^T)
\]

Applying Lemmas 1 and 2 (see Appendix), we obtain the following inequality:

\[
n + \|A^T H R^{-1} HA\|_S \text{tr}(P_F) + \text{tr}(H^T R^{-1} H G Q G^T) \geq \sum_{i=1}^{n} |\lambda_i(A)|^2 + \frac{\text{tr}(G Q G^T)}{\text{tr}(P_F)}
\]  
(32)

which can be rearranged in the form

\[
a_1[\text{tr}(P_F)]^2 + a_2\text{tr}(P_F) - a_3 \geq 0
\]

Solving the above inequality for \(\text{tr}(P_F)\), we obtain (28), thereby completing the proof.

4. Performance Bounds with Incorrect Noise Covariances. We now consider the case for which \(A\) and \(H\) are known exactly, i.e., \(\Delta A = 0\) and \(\Delta H = 0\), and we obtain bounds for \(\text{tr}(P)\) and \(\text{tr}(P_F)\) in terms of \(\Delta Q\) and \(\Delta R\) explicitly.

Theorem 3: If \(\Delta A = 0\) and \(\Delta H = 0\), then an upper bound for \(\text{tr}(P)\) is given by

\[
\text{tr}(P) \leq \pi_3 = i^T M^{-1} b + \| (M^{-1})^T i \| \left( \| L G \|_S^2 \| \Delta Q \| + \| K_M \|_S^2 \| \Delta R \| \right)
\]  
(33)
Proof of Theorem 3: Since $\Delta A = 0$ and $\Delta H = 0$ in (26),

$$\| D \| = \| LGQG^T + K_M \Delta R \|_M$$

$$\leq \| LG \|_s \| \Delta Q \| + \| K_M \|_s \| \Delta R \|$$

where the last inequality follows from the inequality $\| WZ \| \leq \| W \|_s \| Z \|$ [2, p. 37]. The result in (33) then follows from (22).

Theorem 4: If $\Delta A = 0$ and $\Delta H = 0$, then a lower bound for $\text{tr}(P_F)$ is given by

$$\text{tr}(P_F) \geq \pi_4 \equiv \frac{-\alpha_2 + \sqrt{\alpha_2^2 + 4\alpha_1 \alpha_3}}{2\alpha_1}$$

where

$$\alpha_1 \equiv \frac{\| HA \|_s^2}{\min \lambda(R_M) - \| \Delta R \|}$$

and

$$\alpha_2 \equiv n + \frac{\| HG \|_s^2 \left( \| Q_M \|_s + \| \Delta Q \| \right)}{\min \lambda(R_M) - \| \Delta R \|} - \sum_{i=1}^{n} \lambda_i(A)^2$$

Proof of Theorem 4: In (28), $\pi_2$ decreases monotonically as $\alpha_1$ and $\alpha_2$ increase and as $\alpha_3$ decreases. Hence, to obtain the bound $\pi_u$, we bound $\alpha_1$ and $\alpha_2$ from above and $\alpha_3$ from below in terms of $\| \Delta Q \|$ and $\| \Delta R \|$. This is done by bounding $\| A^T R^{-1} H A \|_s$ and $\text{tr}(H^T R^{-1} HGQG^T)$ from above and $\text{tr}(QQ^T)$ from below.

An upper bound for $\| A^T R^{-1} H A \|_s$ is given by

$$\| A^T R^{-1} H A \|_s \leq \| HA \|_s^2 \| R^{-1} \|_s \leq \frac{\| HA \|_s^2}{\min \lambda(R)} \leq \frac{\| HA \|_s^2}{\min \lambda(R_M) - \| \Delta R \|}$$

(38)
where Weyl's inequality [4, p. 157] is used to obtain the last inequality. An upper bound for $\text{tr}(H^T R^{-1} HQ G^T)$ is given by

$$\text{tr}(H^T R^{-1} HQ G^T) \leq \text{tr}(G^T H R^{-1} H G) \| Q \|_S$$

$$\leq \text{tr}(H G G^T) R^{-1} \| G \|_S \| Q \|_S$$

$$\leq \frac{\| H G \|^2}{\min \lambda(R_M)} \left( \| Q_M \|_S + \| \Delta Q \| \right)$$

(39)

where Lemma 1(i) in the Appendix is used to obtain the first two inequalities. A lower bound for $\text{tr}(G G^T)$ is obtained as follows:

$$\text{tr}(G G^T) = \text{tr}[G(Q_M - \Delta Q) G^T]$$

$$= \text{tr}(G Q_M G^T) - \text{tr}(G^T G \Delta Q)$$

$$\geq \text{tr}(G Q_M G^T) - \| G^T G \Delta Q \|_S$$

$$\geq \text{tr}(G Q_M G^T) - \| G \|_S^2 \| \Delta Q \|$$

(40)

where the first inequality follows since $G^T G \Delta Q$ is an $\times$ matrix.

Using (38) to (40) in (29) to (31) results in (34), thereby completing the proof.

Remark: For $\Delta Q \geq 0$ and $\Delta R \geq 0$ (respectively, $\Delta Q \leq 0$ and $\Delta R \leq 0$), Nishimura [11] has given bounds on the matrix $P$ of the form $P_M \geq P$ (respectively, $P \geq P_M$). These results also hold for the nonstationary case. Note that the results of Theorems 3 and 4 do not require any sign definiteness for $\Delta Q$ and $\Delta R$, although they are limited to the stationary case. However, unlike the analysis in [11], the results of the theorems do not require that the filter gain $K_M$ be obtained by a Kalman filter design ((11) to (13)).

5. Performance Bounds with Incorrect Process Configuration. We now assume that $\Delta Q = 0$ and $\Delta R = 0$, i.e., $Q$ and $R$, are known exactly. Theorems 5 and 6 then give an upper bound for $\text{tr}(P)$ and a lower bound
for $\text{tr}(P_F)$, respectively. The analysis used to obtain the upper bound for $\text{tr}(P)$ is more involved than that in Sec. 4 since (15) is coupled with (16) and (17). To derive an upper bound, we first analyze (17) for the stationary case to obtain the following.

**Lemma:** If $A_M$ is diagonalizable, then, for stationary $U$,

$$\|U\|_S \leq \beta = \frac{1}{2} \|U_M\|_S \exp\left[ -\frac{\kappa^2}{2\sigma} \|I_n - F_M\|_S^2 \left( \|A\|_S^2 + 2\|A\|_S \|A_M\|_S \right) \right]$$

(41)

where

$$F_M \equiv (A_M^T + I_n)^{-1}(A_M^T - I_n)$$

(42)

$$\sigma \equiv \max[\text{Re} \lambda(F_M)] < 0$$ (real part of dominant eigenvalue of $F_M$)

$$\kappa \equiv \|T\|_S \|T^{-1}\|_S$$ (spectral condition number of matrix $T$)

$T$ is a similarity transformation matrix to diagonalize $A_M$

and $U_M$ is the solution of the Lyapunov equation:

$$F_M^T U_M + U_M F_M = -(I_n - F_M^T)GQG^T(I_n - F_M)$$

**Remarks:** (a) It is easy to show that a similarity transformation matrix $T$ that diagonalizes $A_M$ also diagonalizes $F_M$.

(b) The diagonalizability condition in the above lemma can be removed by the following procedure: when a model matrix $A_M'$ is not diagonalizable, we can always find a diagonalizable matrix $A_M$ such that $\|A_M - A_M'\|_S$ is as small as we wish [4, p.111]. The lemma can then be used for such $A_M$.

(c) The (Cayley) transformation in (42) maps the eigenvalues of $A_M$ inside the unit circle to the left-half complex plane, Re($\lambda$) < 0.
Proof of Lemma: For stationary $U$, (17) can be written as

$$U = A_M U A_M^T + \Delta A U A_M^T - A_M T U A + G Q G^T$$

(43)

From (42),

$$A_M = (I_n - F_M^T)^{-1}(I_n + F_M^T)$$

(44)

Pre- and postmultiplying (43) by $(I_n - F_M^T)$ and $(I_n - F_M)$ respectively, and substituting for $A_M$ from (44) yields:

$$F_M^T U + U F_M = -\frac{1}{2} (I_n - F_M^T)(G Q G^T + \Delta A U A_M^T - \Delta A U A_M^T - A_M U A_M^T)(I_n - F_M)$$

Therefore, $U$ can be expressed by the following integral [1, p. 239]:

$$U = \frac{1}{2} U_M - \frac{1}{2} \int_0^\infty [\exp(F_M^T t)](I_n - F_M^T)(\Delta A U A_M^T + A_M U A_M^T$$

$$- \Delta A U A_M^T)(I_n - F_M)[\exp(F_M t)]dt$$

which yields the norm inequality

$$\| U \|_S \leq \frac{1}{2} \left[ \| U_M \|_S + \int_0^\infty \| I_n - F_M \|_S^2(\| \Delta A \|_S^2$$

$$+ 2 \| \Delta A \|_S \| A_M \|_S \) \| \exp(F_M t) \|_S \| U \|_S \right] dt$$

(45)

Applying the Bellman-Gronwall lemma [7, p. 420] yields

$$\| U \|_S \leq \frac{1}{2} \| U_M \|_S \exp \left[ \| I_n - F_M \|_S^2(\| \Delta A \|_S^2$$

$$+ 2 \| \Delta A \|_S \| A_M \|_S \) \int_0^\infty \| \exp(F_M t) \|_S \| U \|_S \right] dt$$

Since $T$ diagonalizes $F_M$, it follows that

$$\exp(F_M^T t) = T \exp(\Lambda t) T^{-1}$$

where $\Lambda$ is a diagonal matrix with the eigenvalues of $F_M$ on its diagonal. Therefore,

$$\| \exp(F_M t) \|_S \leq \| T \|_S \| T^{-1} \|_S \| \exp(\Lambda t) \|_S \leq \kappa \| \exp(\delta t)$$

12
Hence

\[
\int_0^\infty \| \exp(F_M t) \|^2_s dt \leq \int_0^\infty k^2 \exp(2\sigma t) dt = -\frac{\kappa^2}{2\sigma}
\]

Substituting the above result into (45) yields (41).

We can now derive the following upper bound for \( \text{tr}(P) \) from Theorem 1.

**Theorem 5:** If \( \Delta Q = 0 \) and \( \Delta R = 0 \), then an upper bound for \( \text{tr}(P) \) is given by

\[
\text{tr}(P) \leq \pi_5 \equiv \left( \mu_1 + \sqrt{\mu_1^2 + \mu_2^2} \right)^2
\]

where

\[
\mu_1 \equiv \sqrt{\beta} (M^{-1})_1 \| L_M \|_s (\| L \|_s \| \Delta A \| + \| K_M \|_s \| \Delta H \| + \| K_M \|_s \| \Delta A \| + \| \Delta A \|)
\]

and

\[
\mu_2 \equiv i^T M^{-1} b + (M^{-1})_1 \left[ \beta \left( \| L \|_s \| \Delta A \| + \| K_M \|_s \| A_M \|_s \| \Delta H \| + \| K_M \|_s \| \Delta H \| + \| \Delta A \| \right)^2 + \| K_M \|_s \| GQG^T \|_s \| \Delta H \| \right]^2 + 2 \| K_M \|_s \| LQG^T \|_s \| \Delta H \|
\]

**Proof of Theorem 5:** Since \( \Delta Q = 0 \) and \( \Delta R = 0 \) in (26),

\[
\| D \| \leq 2 \| LQG^T \Delta H K_M^T \| + \| K_M \|_s \| \Delta H \| + \| K_M \|_s \| GQG^T \|_s \| \Delta H \| \right)^2 + 2 \| \Delta C V A \|_M L^T \|
\]

\[
+ \| \Delta C V A \|_M L^T \|
\]

\[
\leq 2 \| LQG^T \|_s \| K_M \|_s \| \Delta H \| + \| K_M \|_s \| GQG^T \|_s \| \Delta H \| \right)^2 + 2 \| \Delta C \|_s \| V \|_s \| L_M \|_s + \| \Delta C \|_s \| U \|_s
\]

\[
(47)
\]
From (27),
\[
\| \Delta C \| \leq \| I \|_s \| \Delta A \| + \| K_M \|_s \| A_M \|_s \| \Delta A \| + \| K_M \|_s \| \Delta A \| \| \Delta H \|
\]
(48)

Since
\[
\begin{bmatrix}
    P & V^T \\
    V & U
\end{bmatrix} = E \begin{bmatrix}
    xx^T & xx^T \\
    xx^T & xx^T
\end{bmatrix}
\]
is a covariance matrix, it is positive semidefinite, and with \( P > 0 \),

Lemma 3 (see Appendix) yields
\[
\| V \|_s \leq \| P \|_s^{1/2} \| U \|_s^{1/2} \leq [\text{tr}(P)]^{1/2} \| U \|_s^{1/2}
\]
(49)

Using inequalities (47) to (49) in (22), we get a quadratic inequality in \([\text{tr}(P)]^{1/2}\):
\[
\text{tr}(P) - 2\mu_1[\text{tr}(P)]^{1/2} - \mu_2 \leq 0
\]
which yields (46).

**Theorem 6:** If \( \Delta Q = 0 \) and \( \Delta R = 0 \), then a lower bound for \( \text{tr}(P_F) \)
is given by
\[
\text{tr}(P_F) \geq \pi_6 \equiv \frac{-\gamma_2 + \sqrt{\gamma_2^2 + 4\gamma_1\gamma_3}}{2\gamma_1}
\]
(50)

where
\[
\gamma_1 \equiv \left( \| H_M \|_s + \| \Delta H \| \right)^2 R^{-1} \| \Delta H \|_s \left( \| A_M \|_s + \| \Delta A \| \right)^2
\]
\[
\gamma_2 \equiv n + \left( \| H_M \|_s + \| \Delta H \| \right)^2 R^{-1} \| \Delta A \|_s \text{tr}(QG^T)
\]
\[
- \frac{1}{n} \left[ \text{tr}(A_H) - \sum_{i=1}^{n} |\Delta A_{ii}| \right]^2
\]
and
\[
\gamma_3 \equiv \text{tr}(QG^T)
\]

where \( \Delta A_{ii} \) denotes the \((i,i)\)th term of \( \Delta A \).
Proof of Theorem 6: As in the proof of Theorem 4, we bound $a_1$ and $a_2$ from above and $a_3$ from below in (28) to obtain $\pi_6$. Therefore,

$$a_1 = \|A^T H R^{-1} H A\| \leq \|H\| \|R^{-1}\| \|A\|$$

$$\leq \left(\|H_N\| + \|\Delta H\|\right)^2 \|R^{-1}\| \left(\|A_N\| + \|\Delta A\|\right)^2 = \gamma_1$$

$$a_2 = n + \text{tr}(H^T R^{-1} H G G^T) - \sum_{i=1}^{n} |\lambda_i(A)|^2$$

$$\leq n + \|H^T R^{-1} H\| \text{tr}(G G^T) - \frac{1}{n} [\text{tr}(A)]^2$$

$$\leq n + \left(\|H_N\| + \|\Delta H\|\right)^2 \|R^{-1}\| \text{tr}(G G^T)$$

$$- \frac{1}{n} [\text{tr}(A_N) - \text{tr}(\Delta A)]^2$$

where Lemma 1 (i) and Lemma 2 are used to obtain the first inequality for $a_2$. Since

$$[\text{tr}(A_N) - \text{tr}(\Delta A)]^2 \geq \left[|\text{tr}(A_N)| - |\text{tr}(\Delta A)|\right]^2$$

and

$$|\text{tr}(\Delta A)| \leq \sum_{i=1}^{n} |\Delta A_{ii}|$$

it follows that $a_2 \leq \gamma_2$. Finally, we note that $a_3 = \gamma_3$ and that in (28) $\pi_2$ decreases monotonically as $a_1$ and $a_2$ increase. Consequently, the above bounds suffice to obtain the bound $\pi_6$.

6. Example. To illustrate the results presented in the preceding sections, we consider a process with the following state space description:

$$x(k + 1) = \begin{bmatrix} a_{M1} & 0 \\ a_{M2} & a_{M3} \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} w(k)$$

$$y(k + 1) = [1 \ 0] x(k + 1) + v(k + 1)$$
The numerical values for a model are specified as $a_{M1} = -1/3$, $a_{M2} = 1/10$, and $a_{M3} = -1/4$. The modeled noise variances of the zero mean white noises are

$$q_M = E[w(k)^2]_{\text{model}} = 10; \quad r_M = E[v(k)^2]_{\text{model}} = 5$$

The parameters $a_{M2}$ and $a_{M3}$ are assumed to be correct and the modeling error bounds for the parameters $a_{M1}$, $q_M$, and $r_M$ are given by

$$|\Delta a_{M1}| \leq 0.1, \quad |\Delta q| \leq 1, \quad \text{and} \quad |\Delta r| \leq 0.5$$

For the numerical values specified, Theorems 3 to 6 yield the following bounds:

Theorem 3:
- Case 1: $\pi_3 = 3.567$ (3.519)
- Case 2: $\pi_3 = 3.748$ (3.644)

Theorem 4:
- Case 1: $\pi_4 = 2.206$ (3.295)
- Case 2: $\pi_4 = 2.432$ (3.171)

Theorem 5: $\pi_5 = 4.246$ (3.450)

Theorem 6: $\pi_6 = 2.506$ (3.390)

For comparison, the least upper bounds for $\text{tr}(P)$ and the greatest lower bounds for $\text{tr}(P_F)$ are given in parentheses. Cases 1 and 2 for Theorems 3 and 4 denote the following:

Case 1: $|\Delta q| \leq 1, \quad |\Delta r| = 0$

Case 2: $|\Delta r| \leq 0.5, \quad |\Delta q| = 0$
7. Conclusions. The performance of Kalman-type, stationary, linear, discrete-time filters in the presence of modeling errors has been analyzed. The mean-square error of the estimates was used as the performance measure; modeling errors in the system configuration (ΔA and ΔH) and in the noise covariances (ΔQ and ΔR) were considered. Upper bounds for the performance measure of suboptimal filters with modeling errors are given in Theorems 1, 3, and 5 and lower bounds for the optimal filters without errors are given in Theorems 2, 4, and 6. The bounds in Theorems 3 to 6 require knowledge of only the model matrices and the range of errors of these matrices. Consequently, these bounds are useful in practice, as a designer often has information on the range of modeling errors rather than on the exact values of the error matrices.

It has been implicitly assumed in the derivation of the bounds that stability of the system and the filter is preserved in the presence of modeling errors ΔA and ΔH and that sign definiteness of Q and R is preserved in the presence of modeling errors ΔQ and ΔR. It should also be noted that the bounds obtained in this paper may be conservative for some systems, e.g., those with very small stability margins, and it may be desirable to obtain tighter bounds for specific cases.
8. Appendix: Inequalities for Positive Semidefinite Matrices. Some of the inequalities used in the proofs of the theorems presented in the text are proved below.

**Lemma 1:**
(i) If $A > 0$ and $B > 0$, then
\[
\text{tr}(AB) \leq \|A\|_s \text{tr}(B) \leq \text{tr}(A) \text{tr}(B)
\]
(ii) If $A > 0$ and $B > 0$, then
\[
\text{tr}(A^{-1}B) \geq \frac{\text{tr}(B)}{\|A\|_s} \geq \frac{\text{tr}(B)}{\text{tr}(A)}
\]

**Proof of Lemma 1:**
(i) Since $A > 0$ and $B > 0$, $A^{1/2}$ and $B^{1/2}$ exist; hence
\[
\text{tr}(AB) = \text{tr}(A^{1/2}B^{1/2}B^{1/2}A^{1/2})
\]
and since $A > 0$, $\|A\|_s \leq \text{tr}(A)$, thereby completing the proof.

(ii) To show the second inequality, we write
\[
\|A\|_s \text{tr}(A^{-1}B) = \|A^{1/2}\|_s^2 \text{tr}(A^{-1/2}B^{1/2}B^{1/2}A^{-1/2})
\]
and since $A > 0$, $\|A\|_s \geq \text{tr}(A)$, thereby completing the proof.

**Lemma 2:** If $A > 0$ and $\phi$ is an $n \times n$ matrix,
\[
\text{tr}(A^{-1/2}\phi A^{1/2}) \geq \sum_{i=1}^n |\lambda_i(\phi)|^2 \geq \frac{1}{n} [\text{tr}(\phi)]^2
\]

**Proof of Lemma 2:**
\[
\text{tr}(A^{-1/2}\phi A^{1/2}) = \text{tr}(A^{-1/2}\phi A^{1/2})^T = \|A^{-1/2}\phi A^{1/2}\|^2
\]
Since $\|W\|^2 \geq \sum_{i=1}^n |\lambda_i(W)|^2$ for any $n \times n$ matrix $W$ and
\[
\lambda_i(A^{-1/2}\phi A^{1/2}) = \lambda_i(\phi),
\]
Using Cauchy's inequality [9, p. 42], we have
\[
\sum_{i=1}^{n} |\lambda_i(\phi)|^2 \geq \frac{1}{n} \left( \sum_{i=1}^{n} |\lambda_i(\phi)| \right)^2
\]

The result of the lemma then follows since
\[
\sum_{i=1}^{n} |\lambda_i(\phi)| \geq \sum_{i=1}^{n} \lambda_i(\phi) = \text{tr}(\phi)
\]

**Lemma 3:** If \( A \geq 0 \) is partitioned as
\[
\begin{bmatrix}
A_1 & A_3^T \\
A_3 & A_2
\end{bmatrix}
\]
where \( A_1 > 0 \), then
\[
\| A_1 \|_s \| A_2 \|_s \geq \| A_3 \|_s^2
\]

**Proof of Lemma 3:** Since \( A_1^{-1} \) exists:
\[
A = \begin{bmatrix}
I_n & 0 \\
A_3A_1^{-1} & I_n
\end{bmatrix}\begin{bmatrix}
A_1 & 0 \\
0 & A_2 - A_3A_1^{-1}A_3^T
\end{bmatrix}\begin{bmatrix}
I_n & A_1^{-1}A_3^T \\
0 & I_n
\end{bmatrix}
\]

Therefore, \( A \geq 0 \) implies that \( A_2 - A_3A_1^{-1}A_3^T \geq 0 \). Using Weyl's inequality for eigenvalues [4, p. 157], we have
\[
\max \lambda(A_2) - \max \lambda(A_3A_1^{-1}A_3^T) \geq \min \lambda(A_2 - A_3A_1^{-1}A_3^T) \geq 0
\]
i.e.,
\[
\| A_2 \|_s \geq \max \lambda(A_3A_1^{-1}A_3^T)
\]
\[
= \max \lambda(A_1^{-1/2}A_3^TA_3A_1^{-1/2})
\]
\[
= \| A_1^{-1/2}A_3^TA_3A_1^{-1/2} \|_s
\]
Therefore,
\[ \|A_1\|_s \|A_2\|_s \geq \|A_1^{1/2}\|_s \|A_1^{-1/2}A_3 A_3^T A_1^{-1/2}\|_s \|A_1^{1/2}\|_s \]
\[ \geq \|A_3^T A_3\|_s = \|A_3\|^2 \]
thereby completing the proof.

REFERENCES


