

NASA Technical Memorandum 4020

Development of a Technique for Estimating Noise Covariances Using Multiple Observers

W. Thomas Bundick
*Langley Research Center
Hampton, Virginia*



National Aeronautics
and Space Administration

Scientific and Technical
Information Division

1988

Summary

There are times when knowledge of the covariance matrices of the plant and sensor noises is needed, such as in the design of Kalman filter. When this knowledge is not available a priori, estimates of the covariances may be useful. Friedland has developed a technique for estimating the unknown variances by using the sample correlation of the innovations from multiple observers, but he developed a detailed solution for only a particular example. This report extends the work of Friedland by developing a general solution for the estimates of the covariance matrices, specializing the solution to the case of diagonal matrices, developing the statistics (mean and standard deviation) of these estimates, and demonstrating the solution with two examples.

For a particular system the absolute accuracy, or standard deviation, of the estimates and the accuracy of the estimates relative to each other are functions of the number of observers used and the observer gains. This is evident in the examples where very good estimates, that is, estimates with small standard deviations, are obtained for some variances while other estimates have a much larger standard deviation. For the two examples demonstrated, the estimates are shown numerically to be unbiased.

Though the solution assumes that the noise variances are constant, the technique can be used to estimate variances that are time varying if the change is sufficiently slow. Results are presented for a time-varying example in which a step-function increase is introduced into the plant and sensor noise variances. The algorithm does estimate the new values for the variances, but there is, of course, a transient period during which the estimate approaches the new solutions. The length of this transient period depends on the number of samples of the innovations that are averaged to compute the sample correlation.

Introduction

There are many times when knowledge of the variances of the plant and sensor noises in a linear system is useful information. For example, the determination of the gains of a Kalman filter requires knowledge of the covariance matrices of the plant and sensor noises. Inaccuracies in these matrices result in a suboptimal filter and thus in suboptimal estimates of the system state. Another instance where knowledge of the variances is beneficial is during the testing of gyros, accelerometers, and other sensors (ref. 1). Knowledge of the noise variance is used in some parameter estimation problems, such as in the estimation of aircraft stability derivatives. As another example, it has been shown (refs. 2 through 4) that the ability to detect and isolate control element failures in aircraft systems is adversely affected by wind turbulence, which represents the plant noise in this system. Knowledge of the magnitude (variance) of this turbulence could be used potentially in an adaptive failure detection system to improve its performance.

Sometimes knowledge of the noise variances is obtained via direct measurement, such as measurement of the noise at a sensor output. When a priori knowledge of the noise intensity is not available, estimates of the variances may be useful if the estimates are sufficiently accurate. The ability to obtain the estimates in a timely manner is also important in a real-time application, such as in an adaptive system. Friedland (ref. 1) has developed a technique for estimating the unknown variances by using the sample statistics of the innovations from multiple observers. Other approaches have been investigated by Mehra (refs. 5 and 6), Belanger (ref. 7), Godbole (ref. 8), Myers and Tapley (ref. 9), and Dee et al. (ref. 10).

Friedland presents his technique in reference 1 and demonstrates it by applying it to an example of a laser gyro. However, he presents a detailed solution for only this example. The present report extends the work of Friedland by developing a general solution for the estimates and by developing expressions for the statistics (mean and variance) of these estimates. To illustrate the technique it is applied to Friedland's example and to the example of a damped Schuler loop used by Mehra (ref. 5).

Symbols

		a_{ij}, b_{ij}	elements of matrices A and B , respectively, used in the appendix
A	hypothetical matrix		
A_i	$= \mathbf{I} \otimes \mathbf{I} - \mathbf{\Lambda}_i \otimes \mathbf{\Lambda}_i$ (see eq. (19))	B	hypothetical matrix
a_i	row <i>i</i> of matrix A	B_i	$= \chi_i \otimes \chi_i$ (see eq. (20))

C_m	result of intermediate calculation defined by equation (76)	w	number of samples averaged to compute sample correlation of residuals
$e_i(k)$	error in estimate (prediction) of $\mathbf{x}(k)$ produced by i th observer	$\mathbf{w}(k)$	plant noise vector
F	premultiplying matrix that removes redundant elements from $S(\Upsilon_i)$	$\mathbf{x}(k)$	state vector
G	postmultiplying matrix defined by equation (33) to remove unused columns of $F(\mathbf{T}\mathbf{A}_i^{-1}\mathbf{B}_i + \mathbf{U})$	$\hat{\mathbf{x}}_i(k k)$	estimate of state vector at sample k produced by i th observer given k observations
g_{ij}	elements of G	$\hat{\mathbf{x}}_i(k+1 k)$	estimate (prediction) of state vector at sample $k+1$ produced by i th observer given k observations
H	observation matrix	$\mathbf{y}(k)$	vector of observations
I	identity matrix	Z	result of intermediate calculation defined by equation (79)
i, j, k, l	indices	Γ	disturbance input matrix
K_i	gain matrix for i th observer	$\gamma_i(k)$	residuals from i th observer
M	matrix of blocks M_i (see eq. (37))	$\gamma_m^\alpha(k)$	α component of residual from m th observer
M_i	result of intermediate calculation defined by equation (34)	Δ_{mn}^4	matrix of second moments of sample correlations of residuals from m th and n th observers
N	number of samples	$\delta(j, k)$	Kronecker delta function
n	number of states	Λ_i	$= \Phi(\mathbf{I} - \mathbf{K}_i\mathbf{H})$ (see eq. (9))
P_i	steady-state covariance matrix of estimation (prediction) error for i th observer	λ_j	eigenvalue of Λ_i
$P_i(k)$	covariance matrix of estimation (prediction) error for i th observer	Ξ	covariance matrix of combined plant and sensor noise vector $\xi(k)$
P_{mn}	steady-state covariance matrix of estimation (prediction) errors $e_m(k)$ and $e_n(k)$ of m th and n th observers	$\underline{\xi}$	vector of plant and sensor noise variances, that is, of diagonal elements of \mathbf{Q} and \mathbf{R} (see eq. (35))
$P_{mn}(k)$	covariance matrix of estimation (prediction) errors $e_m(k)$ and $e_n(k)$ of m th and n th observers	$\xi(k)$	combined plant and sensor noise vector (see eq. (11))
p	number of observers	$\underline{\xi}_j$	elements of $\underline{\xi}$
Q	covariance matrix of plant noise	$\Sigma_{\xi\xi}$	covariance matrix of estimate $\hat{\xi}$ of plant and sensor noise variances
q	dimension of plant noise vector	σ_{gh}^4	element of matrix of second moments of sample correlations of residuals (see eq. (54))
R	covariance matrix of sensor noise	$\hat{\Upsilon}$	stacked elements of sample correlation of residuals from p observers with redundant elements removed
r	dimension of sensor noise vector	Υ_i	correlation of residuals of i th observer (see eq. (15))
T	$= \mathbf{H} \otimes \mathbf{H}$	$\hat{\Upsilon}_i$	sample correlation of residuals of i th observer (see eq. (27))
U	premultiplying matrix defined by equation (24)	Φ	state transition matrix
u_{ij}	elements of U		
$v(k)$	sensor noise vector		

χ_i	$= [\Gamma, -\Phi K_i]$ (see eq. (10))	#	generalized inverse of the matrix (see eq. (40))
Subscripts:		$\hat{}$	estimated value
i, m, n	i th, m th, and n th observer, respectively	Operators:	
Superscripts:		$E\{ \}$	expected value
T	matrix transpose	$S()$	stacking operator defined by equation (17)
$\alpha, \beta, \delta, \theta, \mu, \nu$	denotes component of vector; or if in pairs, denotes element of matrix	\otimes	Kronecker product defined by equation (16)

Theoretical Development

The System

The first step in the development of a solution for the estimates of the variances of the plant and sensor noises is to describe the system and to formulate the problem in a mathematical framework. To this end consider a discrete, linear, time-invariant system described by the state space equations

$$\mathbf{x}(k+1) = \Phi \mathbf{x}(k) + \Gamma \mathbf{w}(k) \quad (1)$$

$$\mathbf{y}(k) = \mathbf{H} \mathbf{x}(k) + \mathbf{v}(k) \quad (2)$$

where $\mathbf{x}(k)$ is an n dimensional state vector, $\mathbf{y}(k)$ is an r dimensional vector of observations, and the vector noise sequences $\mathbf{w}(k)$ and $\mathbf{v}(k)$ are white and Gaussian with a mean of zero. The vectors $\mathbf{w}(k)$ and $\mathbf{v}(k)$ have the dimensions q and r , respectively. Assume that

$$\left. \begin{aligned} E\{\mathbf{w}(k)\} &= E\{\mathbf{v}(k)\} = 0 \\ E\{\mathbf{w}(j) \mathbf{w}^T(k)\} &= \mathbf{Q} \delta(j, k) \\ E\{\mathbf{v}(j) \mathbf{v}^T(k)\} &= \mathbf{R} \delta(j, k) \\ E\{\mathbf{v}(j) \mathbf{w}^T(k)\} &= 0 \end{aligned} \right\} \quad (3)$$

where the covariance matrices \mathbf{Q} and \mathbf{R} are constant but unknown, $E\{ \}$ denotes the expected value, and $\delta(j, k)$ is the Kronecker delta function. The matrix Φ is the state transition matrix, Γ is the disturbance input matrix, and \mathbf{H} is the observation matrix. The matrices Φ , Γ , and \mathbf{H} are assumed to be constant and known.

The problem is to estimate the unknown noise covariance matrices \mathbf{Q} and \mathbf{R} . The solution to this problem will be developed in the following steps: (1) define a set of observers for the system, (2) develop a (linear) relationship between the covariance of the observer estimation error and the covariances of the plant and measurement noises, (3) develop a (linear) relationship between the correlation of the observer residuals and the covariance of the observer estimation error, (4) from the results of steps (2) and (3) develop an expression for the correlation of the observer residuals in terms of the covariances of the plant and measurement noises, (5) develop an expression for the sample correlation of the observer residuals, (6) use the sample correlation of the residuals as an estimate of the true residual correlation to calculate an estimate of the plant and measurement noise covariances utilizing the solution found in step (4), and (7) develop expressions for the mean and the standard deviation of the estimates of the covariances.

The Observers

For the system described by equations (1) and (2), consider the i th observer described by

$$\hat{\mathbf{x}}_i(k+1|k) = \Phi \hat{\mathbf{x}}_i(k|k) \quad (4)$$

$$\widehat{\mathbf{x}}_i(k|k) = \widehat{\mathbf{x}}_i(k|k-1) + \mathbf{K}_i \boldsymbol{\gamma}_i(k) \quad (5)$$

where \mathbf{K}_i is the gain matrix for the i th observer, and $\boldsymbol{\gamma}_i(k)$, the vector sequence of observer residuals, is given as

$$\boldsymbol{\gamma}_i(k) = \mathbf{y}(k) - \mathbf{H} \widehat{\mathbf{x}}_i(k|k-1) \quad (6)$$

Define the observer error as the difference between the actual state of the plant and the observer's estimate (prediction) of the state as follows:

$$\mathbf{e}_i(k) = \mathbf{x}(k) - \widehat{\mathbf{x}}_i(k|k-1) \quad (7)$$

From equations (1), (2), and (4) through (7),

$$\begin{aligned} \mathbf{e}_i(k+1) &= \mathbf{x}(k+1) - \widehat{\mathbf{x}}_i(k+1|k) \\ &= \boldsymbol{\Phi} \mathbf{x}(k) + \boldsymbol{\Gamma} \mathbf{w}(k) - \boldsymbol{\Phi} \{ \widehat{\mathbf{x}}_i(k|k-1) + \mathbf{K}_i [\mathbf{H} \mathbf{x}(k) \\ &\quad + \mathbf{v}(k) - \mathbf{H} \widehat{\mathbf{x}}_i(k|k-1)] \} \\ &= \boldsymbol{\Phi}(\mathbf{I} - \mathbf{K}_i \mathbf{H}) \mathbf{e}_i(k) + \boldsymbol{\Gamma} \mathbf{w}(k) - \boldsymbol{\Phi} \mathbf{K}_i \mathbf{v}(k) \\ &= \boldsymbol{\Lambda}_i \mathbf{e}_i(k) + \boldsymbol{\chi}_i \boldsymbol{\xi}(k) \end{aligned} \quad (8)$$

where a closed-loop observer matrix $\boldsymbol{\Lambda}_i$, a new disturbance input matrix $\boldsymbol{\chi}_i$, and a combined plant/sensor noise vector $\boldsymbol{\xi}(k)$ are defined by

$$\boldsymbol{\Lambda}_i = \boldsymbol{\Phi}(\mathbf{I} - \mathbf{K}_i \mathbf{H}) \quad (9)$$

$$\boldsymbol{\chi}_i = [\boldsymbol{\Gamma}, -\boldsymbol{\Phi} \mathbf{K}_i] \quad (10)$$

$$\boldsymbol{\xi}(k) = \begin{bmatrix} \mathbf{w}(k) \\ \mathbf{v}(k) \end{bmatrix} \quad (11)$$

The covariance $\mathbf{P}_i(k)$ of the estimation (prediction) error is

$$\mathbf{P}_i(k) = \mathbf{E} \{ \mathbf{e}_i(k) \mathbf{e}_i^T(k) \} \quad (12)$$

If \mathbf{K}_i is chosen such that $\boldsymbol{\Lambda}_i$ is asymptotically stable (that is, eigenvalues λ_j of $\boldsymbol{\Lambda}_i$ such that $|\lambda_j| < 1$), then the steady-state covariance \mathbf{P}_i is the solution of

$$\mathbf{P}_i - \boldsymbol{\Lambda}_i \mathbf{P}_i \boldsymbol{\Lambda}_i^T = \boldsymbol{\chi}_i \boldsymbol{\Xi} \boldsymbol{\chi}_i^T \quad (13)$$

where

$$\begin{aligned} \boldsymbol{\Xi} &= \mathbf{E} \{ \boldsymbol{\xi}(k) \boldsymbol{\xi}^T(k) \} \\ &= \begin{bmatrix} \mathbf{Q} & | & 0 \\ \hline 0 & | & \mathbf{R} \end{bmatrix} \end{aligned} \quad (14)$$

Such stabilizing gains \mathbf{K}_i can be found (e.g., by using the Kalman filter algorithm) if the system (eqs. (1) and (2)) is detectable. Equation (13) is a linear relationship between the covariance of the estimation error and the covariances of the plant and sensor noises. Thus, step (2) is complete.

Correlation of the Residuals

The correlation of the residuals of the i th observer is given by

$$\begin{aligned}
 \Upsilon_i &= E \left\{ \gamma_i(k) \gamma_i^T(k) \right\} \\
 &= E \left\{ [\mathbf{y}(k) - \mathbf{H} \hat{\mathbf{x}}_i(k|k-1)] [\mathbf{y}(k) - \mathbf{H} \hat{\mathbf{x}}_i(k|k-1)]^T \right\} \\
 &= E \left\{ [\mathbf{H} \mathbf{e}_i(k) + \mathbf{v}(k)] [\mathbf{H} \mathbf{e}_i(k) + \mathbf{v}(k)]^T \right\} \\
 &= \mathbf{H} \mathbf{P}_i \mathbf{H}^T + \mathbf{R}
 \end{aligned} \tag{15}$$

Equation (15) relates the correlation Υ_i of the residuals to the observer estimation error \mathbf{P}_i as desired in step (3). The residual correlation Υ_i now can be linearly related to the plant and sensor noise covariance Ξ by eliminating \mathbf{P}_i from equations (13) and (15) as follows:

As shown by the example in the appendix, the Kronecker product (ref. 11, p. 22) can be defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix} \tag{16}$$

Define the stacking operator $S(\cdot)$ such that if \mathbf{a}_i are the rows of the $m \times n$ matrix \mathbf{A} , then $S(\mathbf{A})$ is an $mn \times 1$ column vector as follows (see the appendix for an example):

$$S(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1^T \\ \cdot \\ \mathbf{a}_2^T \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{a}_m^T \end{bmatrix} \tag{17}$$

From using this notation, the solution of equation (13) for \mathbf{P}_i in terms of Ξ is given by (ref. 12, p. 89)

$$\begin{aligned}
 [\mathbf{I} \otimes \mathbf{I} - \Lambda_i \otimes \Lambda_i] S(\mathbf{P}_i) &= S(\chi_i \Xi \chi_i^T) \\
 &= [\chi_i \otimes \chi_i] S(\Xi)
 \end{aligned} \tag{18}$$

Let

$$\mathbf{A}_i = \mathbf{I} \otimes \mathbf{I} - \Lambda_i \otimes \Lambda_i \tag{19}$$

$$\mathbf{B}_i = \chi_i \otimes \chi_i \tag{20}$$

Then,

$$\mathbf{A}_i S(\mathbf{P}_i) = \mathbf{B}_i S(\Xi)$$

$$S(\mathbf{P}_i) = \mathbf{A}_i^{-1} \mathbf{B}_i S(\Xi) \tag{21}$$

A solution to equation (18) is assured if the eigenvalues of \mathbf{A}_i have a magnitude less than unity (ref. 13, p. 181, prob. 2.6-2). Rewriting equation (15) gives

$$\begin{aligned} S(\mathbf{Y}_i) &= S(\mathbf{H}\mathbf{P}_i\mathbf{H}^T) + S(\mathbf{R}) \\ &= (\mathbf{H} \otimes \mathbf{H}) S(\mathbf{P}_i) + S(\mathbf{R}) \\ &= \mathbf{T} S(\mathbf{P}_i) + S(\mathbf{R}) \end{aligned} \quad (22)$$

where

$$\mathbf{T} = \mathbf{H} \otimes \mathbf{H}$$

But $S(\mathbf{R})$ can be written as

$$S(\mathbf{R}) = \mathbf{U} S(\mathbf{E}) \quad (23)$$

where

$$\begin{aligned} u_{ij} &= 1 && \begin{cases} 1 \leq l \leq r; \\ (l-1)r+1 \leq i \leq lr; \\ j = i + q(q+r) + lq \end{cases} \\ &= 0 && \text{(elsewhere)} \end{aligned} \quad (24)$$

Combining equations (21), (22), and (23) gives

$$\begin{aligned} S(\mathbf{Y}_i) &= \mathbf{T}\mathbf{A}_i^{-1}\mathbf{B}_i S(\mathbf{E}) + \mathbf{U} S(\mathbf{E}) \\ &= (\mathbf{T}\mathbf{A}_i^{-1}\mathbf{B}_i + \mathbf{U}) S(\mathbf{E}) \end{aligned} \quad (25)$$

But the matrix \mathbf{Y}_i is symmetric; thus, elements $\mathbf{Y}_i(m, n)$ are redundant for $m > n$. To reduce the size of the arrays and to avoid potential problems with taking the matrix inverse, remove these elements from $S(\mathbf{Y}_i)$ and remove the corresponding rows of $\mathbf{T}\mathbf{A}_i^{-1}\mathbf{B}_i + \mathbf{U}$ as follows: Define a matrix \mathbf{F} of dimension $[(r^2 + r)/2] \times r^2$ as the $r^2 \times r^2$ identity matrix with the following rows removed: $r+1; 2r+1, 2r+2; 3r+1, 3r+2, 3r+3; \dots; (r-1)r+1, \dots, r^2-1$. Equation (25) now becomes

$$\mathbf{F}S(\mathbf{Y}_i) = \mathbf{F}(\mathbf{T}\mathbf{A}_i^{-1}\mathbf{B}_i + \mathbf{U}) S(\mathbf{E}) \quad (26)$$

We now have a linear relationship (eq. (26)) between the correlation \mathbf{Y}_i of the residuals and the covariance \mathbf{E} of the noises, thus completing step (4).

Sample Correlation

Let $\hat{\mathbf{Y}}_i(k)$ be the sample correlation of the residuals γ_i computed from

$$\begin{aligned} \hat{\mathbf{Y}}_i(k) &= \frac{1}{k} \sum_{j=1}^k [\gamma_i(j) \gamma_i^T(j)] \\ &= \frac{k-1}{k} \hat{\mathbf{Y}}_i(k-1) + \frac{1}{k} [\gamma_i(k) \gamma_i^T(k)] \end{aligned} \quad (27)$$

In the software implementation of equation (27), only elements $\hat{\mathbf{Y}}_i(m, n)$ for $m \geq n$ are computed since the matrix $\hat{\mathbf{Y}}(k)$ is symmetric.

The sample correlation $\hat{\mathbf{Y}}_i(k)$ is an estimate of \mathbf{Y}_i as required in step (5), and $S(\hat{\mathbf{Y}}_i(k))$ could be used in equation (26) to solve for $\hat{\mathbf{E}}$, an estimate of \mathbf{E} . However, $S(\mathbf{E})$ is $(q+r)^2 \times 1$, whereas $\mathbf{F} S(\hat{\mathbf{Y}}_i(k))$ is only

$[(r^2 + r)/2] \times 1$; thus, there are $(r^2 + r)/2$ equations in $(q + r)^2$ unknowns. Actually, $2qr$ elements of Ξ are always zero, so there are only $q^2 + r^2$ unknowns.

This difficulty can be overcome by using multiple observers, with a different gain K_i for each observer and thus for different residuals $\gamma_i(k)$. Then,

$$\begin{bmatrix} S(\hat{Y}_1(k)) \\ \text{-----} \\ S(\hat{Y}_2(k)) \\ \text{-----} \\ \vdots \\ S(\hat{Y}_p(k)) \end{bmatrix} = \begin{bmatrix} \mathbf{TA}_1^{-1}\mathbf{B}_1 + \mathbf{U} \\ \text{-----} \\ \mathbf{TA}_2^{-1}\mathbf{B}_2 + \mathbf{U} \\ \text{-----} \\ \vdots \\ \mathbf{TA}_p^{-1}\mathbf{B}_p + \mathbf{U} \end{bmatrix} S(\hat{\Xi}) \quad (28)$$

The problem here is that a necessary condition for a unique solution of equation (28) is that

$$p \geq \frac{q^2 + r^2}{(r^2 + r)/2} \quad (29)$$

Thus, p can be quite large, and the number of observers required can be larger than practical. As an illustration, for Mehra's example (ref. 5) which will be discussed later, $q = 3$, $r = 2$, and thus $p \geq 5$. By taking advantage of the fact that \mathbf{Q} and \mathbf{R} are symmetric, the number of unknowns can be further reduced to $(q^2 + q + r^2 + r)/2$, but this matter will not be pursued here.

Diagonal Covariances

Although equation (28) can be solved for the estimate $\hat{\Xi}$ of the covariances, the number of unknowns in Ξ (and thus in $\hat{\Xi}$) can be reduced considerably if the problem is restricted by making the frequent assumption that the covariances \mathbf{Q} and \mathbf{R} are diagonal; that is, the components of $\mathbf{w}(k)$ are uncorrelated with each other, and similarly for $\mathbf{v}(k)$. Then, Ξ is also diagonal, the number of unknown variances is reduced to $q + r$, and the number p of observers required for a solution may be reduced also. In this case,

$$E\{\mathbf{w}(k) \mathbf{w}^T(j)\} = \begin{bmatrix} q_1 & & & \\ & q_2 & 0 & \\ & & \ddots & \\ & 0 & & q_q \end{bmatrix} \delta(j, k) \quad (30)$$

$$E\{\mathbf{v}(k) \mathbf{v}^T(j)\} = \begin{bmatrix} r_1 & & & \\ & r_2 & 0 & \\ & & \ddots & \\ & 0 & & r_r \end{bmatrix} \delta(j, k) \quad (31)$$

such that Ξ becomes

$$\Xi = \begin{bmatrix} \mathbf{Q} & | & 0 \\ \text{-----} & & \\ 0 & | & \mathbf{R} \end{bmatrix} = \begin{bmatrix} \xi_1 & & & \\ & \xi_2 & 0 & \\ & & \ddots & \\ 0 & & & \xi_{q+r} \end{bmatrix} \quad (32)$$

Use of equation (32) in equation (26) shows that only $q + r$ columns of $\mathbf{F}(\mathbf{T}\mathbf{A}_i^{-1}\mathbf{B}_i + \mathbf{U})$ are multiplying nonzero elements of $\mathbf{S}(\hat{\Xi})$. To remove the assumed zero elements from $\mathbf{S}(\hat{\Xi})$ (off-diagonal elements of $\hat{\Xi}$) and to eliminate the corresponding unused $(q + r)^2 - (q + r)$ columns of $\mathbf{F}(\mathbf{T}\mathbf{A}_i^{-1}\mathbf{B}_i + \mathbf{U})$, define a new matrix \mathbf{G} with dimensions $(q + r)^2 \times (q + r)$ and elements g_{ij} such that for $1 \leq j \leq q + r$ and $1 \leq i \leq (q + r)^2$,

$$g_{ij} = \begin{cases} 1 & \text{(for } 1 \leq j \leq q + r; i = (j - 1)(q + r) + j) \\ 0 & \text{(elsewhere)} \end{cases} \quad (33)$$

See the appendix for an illustration of \mathbf{G} . Define a new matrix \mathbf{M}_i by

$$\mathbf{M}_i = \mathbf{F}(\mathbf{T}\mathbf{A}_i^{-1}\mathbf{B}_i + \mathbf{U})\mathbf{G} \quad (34)$$

(Eliminate all columns except columns j , $j = 1, q + r + 2, 2(q + r) + 3, 3(q + r) + 4, \dots, (q + r - 1)(q + r) + q + r = (q + r)^2$.) Similarly, define a new vector of unknown variances $\underline{\xi}$ as

$$\underline{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_q \\ \xi_{q+1} \\ \vdots \\ \xi_{q+r} \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_q \\ r_1 \\ \vdots \\ r_r \end{bmatrix} = \mathbf{G}^T \mathbf{S}(\Xi) \quad (35)$$

Corresponding to equation (26), the relationship between the correlation of the residuals and the covariances of the noises (step (4)) for the case of diagonal covariance matrices becomes

$$\mathbf{F} \mathbf{S}(\Upsilon_i) = \mathbf{M}_i \underline{\xi} \quad (36)$$

Equation (36) is a set of $(r^2 + r)/2$ equations in $q + r$ unknowns.

If $(r^2 + r)/2 < q + r$, then p observers can be used, as in equation (28), such that $p(r^2 + r)/2 \geq q + r$. Using the sample covariances $\hat{\Upsilon}_i(k)$ as estimates of the residual covariances $\Upsilon_i(k)$, equation (37) can be used to compute an estimate $\hat{\underline{\xi}}$ of the plant and sensor variance vector $\underline{\xi}$. Thus,

$$\hat{\underline{\Upsilon}} = \begin{bmatrix} \mathbf{F} \mathbf{S}(\hat{\Upsilon}_1(k)) \\ \text{-----} \\ \mathbf{F} \mathbf{S}(\hat{\Upsilon}_2(k)) \\ \text{-----} \\ \vdots \\ \text{-----} \\ \mathbf{F} \mathbf{S}(\hat{\Upsilon}_p(k)) \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1 \\ \text{-----} \\ \mathbf{M}_2 \\ \text{-----} \\ \vdots \\ \text{-----} \\ \mathbf{M}_p \end{bmatrix} \hat{\underline{\xi}} = \mathbf{M} \hat{\underline{\xi}} \quad (37)$$

If $\text{rank}(\mathbf{M}) = p(r^2 + r)/2$ and $p(r^2 + r)/2 = q + r$, then the solution is

$$\hat{\underline{\xi}} = \mathbf{M}^{-1} \hat{\underline{\Upsilon}} \quad (38)$$

Otherwise, a generalized inverse $\mathbf{M}^\#$ must be used such that

$$\hat{\underline{\xi}} = \mathbf{M}^\# \hat{\underline{\Upsilon}} \quad (39)$$

If \mathbf{M} is full rank, that is, $\text{rank}(\mathbf{M}) = q + r$, a least-squares solution for $\hat{\underline{\xi}}$ is given by equation (39) when $\mathbf{M}^\#$ is given by

$$\mathbf{M}^\# = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \quad (40)$$

Thus, we have developed a general solution for the estimates $\hat{\underline{\xi}}$ of the unknown elements of a diagonal noise matrix $\underline{\Xi}$ in terms of the sample correlation $\hat{\underline{\Gamma}}$ of the residuals of the observers, thus completing step (6).

Statistics of the estimated covariance. We next develop an expression for the mean and the standard deviation of the estimate of the plant and sensor noise variances for use in evaluating the accuracy of the estimation process. Before doing that, we need to develop some preliminary relationships. Mehra (ref. 5) has shown that the innovation sequence $\gamma(k)$ from a suboptimal filter is a Gaussian process. This sequence has a mean of zero in the steady state. Consider the following:

$$\begin{aligned} E\{\gamma_i(k)\} &= E\{\mathbf{y}(k) - \mathbf{H} \hat{\mathbf{x}}_i(k|k-1)\} \\ &= E\{\mathbf{H} \mathbf{x}(k) + \mathbf{v}(k) - \mathbf{H} \hat{\mathbf{x}}_i(k|k-1)\} \\ &= \mathbf{H} E\{\mathbf{e}_i(k|k-1)\} + E\{\mathbf{v}(k)\} \end{aligned} \quad (41)$$

In equations (3) it was assumed that $E\{\mathbf{v}(k)\} = 0$. The error $\mathbf{e}_i(k)$ in the estimate $\hat{\mathbf{x}}_i(k|k-1)$ of the state was given by equation (8) to be

$$\mathbf{e}_i(k) = \Phi(\mathbf{I} - \mathbf{K}_i \mathbf{H}) \mathbf{e}_i(k-1) + \Gamma \mathbf{w}(k-1) - \Phi \mathbf{K}_i \mathbf{v}(k-1) \quad (42)$$

Equation (42) has a solution

$$\begin{aligned} \mathbf{e}_i(k) &= [\Phi(\mathbf{I} - \mathbf{K}_i \mathbf{H})]^k \mathbf{e}_i(0) - \sum_{j=1}^k [\Phi(\mathbf{I} - \mathbf{K}_i \mathbf{H})]^{j-1} \Phi \mathbf{K}_i \mathbf{v}(k-j) \\ &\quad + \sum_{j=1}^k [\Phi(\mathbf{I} - \mathbf{K}_i \mathbf{H})]^{j-1} \Gamma \mathbf{w}(k-j) \end{aligned} \quad (43)$$

such that

$$\begin{aligned} E\{\mathbf{e}_i(k)\} &= [\Phi(\mathbf{I} - \mathbf{K}_i \mathbf{H})]^k E\{\mathbf{e}_i(0)\} \\ &\quad - \sum_{j=1}^k [\Phi(\mathbf{I} - \mathbf{K}_i \mathbf{H})]^{j-1} \Phi \mathbf{K}_i E\{\mathbf{v}(k-j)\} \\ &\quad + \sum_{j=1}^k [\Phi(\mathbf{I} - \mathbf{K}_i \mathbf{H})]^{j-1} \Gamma E\{\mathbf{w}(k-j)\} \end{aligned} \quad (44)$$

From equations (3), $E\{\mathbf{v}(k-j)\} = 0$ and $E\{\mathbf{w}(k-j)\} = 0$. Furthermore, if the filter is asymptotically stable, then in the steady state $[\Phi(\mathbf{I} - \mathbf{K}_i \mathbf{H})]^k$ decays to zero, as k approaches infinity; thus, $E\{\mathbf{e}_i(k)\} = 0$. Therefore, the innovations $\gamma_i(k)$ have a mean of zero in the steady state.

Before proceeding with the development of the statistics of the estimates of the noise variances, consider the second moment of the errors $\mathbf{e}_m(k+1)$ and $\mathbf{e}_n(k+1)$ in the prediction estimates $\hat{\mathbf{x}}_m(k+1|k)$ and $\hat{\mathbf{x}}_n(k+1|k)$ of the state produced by the m th and n th filters, respectively. Thus,

$$\begin{aligned} \mathbf{P}_{mn}(k+1) &= E\{\mathbf{e}_m(k+1) \mathbf{e}_n^T(k+1)\} \\ &= \Phi(\mathbf{I} - \mathbf{K}_m \mathbf{H}) E\{\mathbf{e}_m(k) \mathbf{e}_n^T(k)\} (\mathbf{I} - \mathbf{K}_n \mathbf{H})^T \Phi^T \end{aligned}$$

$$\begin{aligned}
& + \Gamma E\{\mathbf{w}(k) \mathbf{w}^T(k)\} \Gamma^T + \Phi \mathbf{K}_m E\{\mathbf{v}(k) \mathbf{v}^T(k)\} \mathbf{K}_n^T \Phi^T \\
& = \Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H}) \mathbf{P}_{mn}(k) (\mathbf{I} - \mathbf{K}_n \mathbf{H})^T \Phi^T + \Gamma \mathbf{Q} \Gamma^T \\
& \quad + \Phi \mathbf{K}_m \mathbf{R} \mathbf{K}_n^T \Phi^T
\end{aligned} \tag{45}$$

In the steady state, equation (45) for the second moment of the prediction error becomes the discrete Lyapunov equation, which can be solved to obtain the error covariance \mathbf{P}_{mn} . Thus,

$$\mathbf{P}_{mn} = \Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H}) \mathbf{P}_{mn} (\mathbf{I} - \mathbf{K}_n \mathbf{H})^T \Phi^T + \Gamma \mathbf{Q} \Gamma^T + \Phi \mathbf{K}_m \mathbf{R} \mathbf{K}_n^T \Phi^T \tag{46}$$

Mean of the estimates. Now let us proceed with the task of determining the mean and the standard deviation of the estimates of the variances of the plant and sensor noises. The mean of the vector $\hat{\xi}$ of estimated variances is given by

$$\begin{aligned}
E\{\hat{\xi}\} & = \mathbf{M}^\# E\{\hat{\mathbf{Y}}\} \\
& = \mathbf{M}^\# E \left\{ \begin{bmatrix} \mathbf{F} \mathbf{S}(\hat{\mathbf{Y}}_1) \\ \vdots \\ \mathbf{F} \mathbf{S}(\hat{\mathbf{Y}}_p) \end{bmatrix} \right\} \\
& = \mathbf{M}^\# \begin{bmatrix} \mathbf{F} \mathbf{S}(E\{\hat{\mathbf{Y}}_1\}) \\ \vdots \\ \mathbf{F} \mathbf{S}(E\{\hat{\mathbf{Y}}_p\}) \end{bmatrix}
\end{aligned} \tag{47}$$

where $\hat{\mathbf{Y}}_i$ is the sample correlation of the residuals of the i th filter defined by equation (27). The expected value of $\hat{\mathbf{Y}}_i$ can be found from

$$\begin{aligned}
E\{\hat{\mathbf{Y}}_i\} & = E \left\{ \frac{1}{w} \sum_{l=1}^w \gamma_i(l) \gamma_i^T(l) \right\} \\
& = \frac{1}{w} \sum_{l=1}^w E \left\{ [\mathbf{y}(l) - \mathbf{H} \hat{\mathbf{x}}_i(l|l-1)] [\mathbf{y}(l) - \mathbf{H} \hat{\mathbf{x}}_i(l|l-1)]^T \right\} \\
& = \frac{1}{w} \sum_{l=1}^w E \left\{ [\mathbf{H} \mathbf{e}_i(l) + \mathbf{v}(l)] [\mathbf{H} \mathbf{e}_i(l) + \mathbf{v}(l)]^T \right\} \\
E\{\hat{\mathbf{Y}}_i\} & = \frac{1}{w} \sum_{l=1}^w [\mathbf{H} \mathbf{P}_{ii}(l) \mathbf{H}^T + \mathbf{R}]
\end{aligned} \tag{48}$$

In the steady state this becomes

$$E\{\hat{\mathbf{Y}}_i\} = \mathbf{H} \mathbf{P}_{ii} \mathbf{H}^T + \mathbf{R} \tag{49}$$

Therefore, the expected value of the estimate of the variances is

$$E\{\hat{\underline{\xi}}_i\} = \mathbf{M}^\# \begin{bmatrix} \mathbf{F S} (\mathbf{H P}_{11} \mathbf{H}^T + \mathbf{R}) \\ \vdots \\ \mathbf{F S} (\mathbf{H P}_{pp} \mathbf{H}^T + \mathbf{R}) \end{bmatrix} \quad (50)$$

Equation (50) is an expression for the mean of the estimated variance, the first statistic to be developed in step (7).

Mean square of the estimates. The mean square, or second moment, of the estimates of the variances of the plant and sensor noises can be found similarly, but the computations are somewhat more tedious. Thus,

$$E\{\hat{\underline{\xi}} \hat{\underline{\xi}}^T\} = \mathbf{M}^\# E\{\hat{\underline{\mathbf{Y}}} \hat{\underline{\mathbf{Y}}}^T\} (\mathbf{M}^\#)^T \quad (51)$$

The second moment of the sample correlation of the residuals can be found as follows:

$$E\{\hat{\underline{\mathbf{Y}}} \hat{\underline{\mathbf{Y}}}^T\} = E \left\{ \begin{bmatrix} \mathbf{F S} (\hat{\mathbf{Y}}_1) \\ \vdots \\ \mathbf{F S} (\hat{\mathbf{Y}}_p) \end{bmatrix} \begin{bmatrix} \mathbf{F S} (\hat{\mathbf{Y}}_1) & \cdots & \mathbf{F S} (\hat{\mathbf{Y}}_p) \end{bmatrix} \right\}$$

$$= \begin{bmatrix} | & & | \\ \hline & \Delta_{mn}^4 & \\ \hline | & & | \end{bmatrix} \quad (52)$$

where

$$\begin{aligned} \Delta_{mn}^4 &= E \left\{ \mathbf{F S} (\hat{\mathbf{Y}}_m) \mathbf{S}^T (\hat{\mathbf{Y}}_n) \mathbf{F}^T \right\} \\ &= E \left\{ \mathbf{F S} \left(\frac{1}{w} \sum_{l=1}^w \gamma_m(l) \gamma_m^T(l) \right) \mathbf{S}^T \left(\frac{1}{w} \sum_{k=1}^w \gamma_n(k) \gamma_n^T(k) \right) \mathbf{F}^T \right\} \\ &= [\sigma_{gh}^4] \end{aligned} \quad (53)$$

In terms of notation, Δ_{mn}^4 is the block in row m and column n of blocks when the matrix $E\{\hat{\underline{\mathbf{Y}}} \hat{\underline{\mathbf{Y}}}^T\}$ is partitioned with regard to observers. The variable σ_{gh}^4 is the element in row g and column h of $E\{\hat{\underline{\mathbf{Y}}} \hat{\underline{\mathbf{Y}}}^T\}$ and is expressed by

$$\sigma_{gh}^4 = \frac{1}{w} \sum_{l=1}^w \sum_{k=1}^w E \left\{ \gamma_m^\alpha(l) \gamma_m^\beta(l) \gamma_n^\theta(k) \gamma_n^\delta(k) \right\} \quad (54)$$

where

$$\left. \begin{aligned} g &= \frac{\alpha - 1}{2} (2r + 2 - \alpha) + \beta + 1 - \alpha + \frac{m(r^2 + r)}{2} \\ (1 \leq \alpha \leq r; \alpha \leq \beta \leq r) \\ h &= \frac{\theta - 1}{2} (2r + 2 - \theta) + \delta + 1 - \theta + \frac{n(r^2 + r)}{2} \\ (1 \leq \theta \leq r; \theta \leq \delta \leq r) \end{aligned} \right\} \quad (55)$$

and $\gamma_m^\alpha(l)$ denotes the α component of the vector $\gamma_m(l)$.

Since the random variables $\gamma_i^\nu(j)$ are Gaussian with a mean of zero, the expected value of the product of four variables as in equation (54) is given by (ref. 14, p. 274, prob. 8-14)

$$\begin{aligned} \sigma_{gh}^4 &= \frac{1}{w^2} \sum_{l=1}^w \sum_{k=1}^w \left[E \left\{ \gamma_m^\alpha(l) \gamma_m^\beta(l) \right\} E \left\{ \gamma_n^\theta(k) \gamma_n^\delta(k) \right\} \right. \\ &\quad + E \left\{ \gamma_m^\alpha(l) \gamma_n^\theta(k) \right\} E \left\{ \gamma_m^\beta(l) \gamma_n^\delta(k) \right\} \\ &\quad \left. + E \left\{ \gamma_m^\alpha(l) \gamma_n^\delta(k) \right\} E \left\{ \gamma_m^\beta(l) \gamma_n^\theta(k) \right\} \right] \end{aligned} \quad (56)$$

To this point we have found the mean square of the estimate $\hat{\xi}$ of the variances of the plant and sensor noises in terms of the second moments of the sample correlations $\hat{\mathbf{Y}}$ of the filter residuals (eq. (51)). For each element σ_{gh}^4 of the second-moment matrix of the sample correlations of the residuals, we found an expression in terms of the second moments of the residuals γ_m^α (eq. (56)).

Let us turn now to consider the general term $E \left\{ \gamma_m^\mu(l) \gamma_n^\nu(k) \right\}$ of the matrix of second moments of the residuals. First, consider

$$E \left\{ \gamma_m(l) \gamma_n^T(k) \right\} = E \left\{ [\mathbf{H} \mathbf{e}_m(l) + \mathbf{v}(l)] [\mathbf{H} \mathbf{e}_n(k) + \mathbf{v}(k)]^T \right\} \quad (57)$$

Equation (57) becomes

$$\begin{aligned} E \left\{ \gamma_m(l) \gamma_n^T(k) \right\} &= \mathbf{H} E \left\{ \mathbf{e}_m(l) \mathbf{e}_n^T(k) \right\} \mathbf{H}^T + E \left\{ \mathbf{v}(l) \mathbf{e}_n^T(k) \right\} \mathbf{H}^T \\ &\quad + \mathbf{H} E \left\{ \mathbf{e}_m(l) \mathbf{v}^T(k) \right\} + E \left\{ \mathbf{v}(l) \mathbf{v}^T(k) \right\} \end{aligned} \quad (58)$$

From equations (3),

$$E \left\{ \mathbf{v}(l) \mathbf{v}^T(k) \right\} = \mathbf{R} \delta(l, k)$$

We will examine the remaining terms of equation (58) for three different conditions: $l < k$, $l = k$, and $l > k$. First, assume that $l > k$. Note that the state estimation error can be written as

$$\begin{aligned} \mathbf{e}_m(l) &= [\Phi(\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{l-k} \mathbf{e}_m(k) - \sum_{i=1}^{l-k} [\Phi(\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{i-1} \Phi \mathbf{K}_m \mathbf{v}(l-i) \\ &\quad + \sum_{i=1}^{l-k} [\Phi(\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{i-1} \Gamma \mathbf{w}(l-i) \end{aligned} \quad (59)$$

Substituting equation (59) into (58) gives

$$\begin{aligned} \mathbf{E} \left\{ \mathbf{e}_m(l) \mathbf{e}_n^T(k) \right\} &= [\Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{l-k} \mathbf{E} \left\{ \mathbf{e}_m(k) \mathbf{e}_n^T(k) \right\} \\ &\quad - \sum_{i=1}^{l-k} [\Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{i-1} \Phi \mathbf{K}_m \mathbf{E} \left\{ \mathbf{v}(l-i) \mathbf{e}_n^T(k) \right\} \\ &\quad + \sum_{i=1}^{l-k} [\Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{i-1} \Gamma \mathbf{E} \left\{ \mathbf{w}(l-i) \mathbf{e}_n^T(k) \right\} \end{aligned} \quad (60)$$

But from equations (45) and (46) we get

$$\mathbf{E} \left\{ \mathbf{e}_m(k) \mathbf{e}_n(k) \right\} = \mathbf{P}_{mn}(k) \quad (61)$$

and since the prediction estimation error does not depend on current or future noise samples, the result is

$$\left. \begin{aligned} \mathbf{E} \left\{ \mathbf{v}(l-i) \mathbf{e}_n^T(k) \right\} &= 0 \\ \mathbf{E} \left\{ \mathbf{w}(l-i) \mathbf{e}_n^T(k) \right\} &= 0 \end{aligned} \right\} \quad (l-i \geq k) \quad (62)$$

Therefore, the first term on the right side of equation (58) becomes

$$\mathbf{H} \mathbf{E} \left\{ \mathbf{e}_m(l) \mathbf{e}_n^T(k) \right\} \mathbf{H}^T = \mathbf{H} [\Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{l-k} \mathbf{P}_{mn}(k) \mathbf{H}^T \quad (l > k) \quad (63)$$

the second term of equation (58) becomes

$$\mathbf{E} \left\{ \mathbf{v}(l) \mathbf{e}_n^T(k) \right\} \mathbf{H}^T = 0 \quad (l > k) \quad (64)$$

and the third term of equation (58) becomes

$$\begin{aligned} \mathbf{H} \mathbf{E} \left\{ \mathbf{e}_m(l) \mathbf{v}^T(k) \right\} &= \mathbf{H} [\Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{l-k} \mathbf{E} \left\{ \mathbf{e}_m(k) \mathbf{v}^T(k) \right\} \\ &\quad - \mathbf{H} \sum_{i=1}^{l-k} [\Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{i-1} \Phi \mathbf{K}_m \mathbf{E} \left\{ \mathbf{v}(l-i) \mathbf{v}^T(k) \right\} \\ &\quad + \mathbf{H} \sum_{i=1}^{l-k} [\Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{i-1} \Gamma \mathbf{E} \left\{ \mathbf{w}(l-i) \mathbf{v}^T(k) \right\} \\ &= -\mathbf{H} [\Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{l-k-1} \Phi \mathbf{K}_m \mathbf{R} \quad (l > k) \end{aligned} \quad (65)$$

Now assume that $l < k$. As in the development of equations (60) through (65), it can be shown that

$$\mathbf{E} \left\{ \mathbf{e}_m(l) \mathbf{e}_n^T(k) \right\} = \mathbf{P}_{mn}(l) \left[(\mathbf{I} - \mathbf{K}_n \mathbf{H})^T \Phi^T \right]^{k-l} \quad (l < k) \quad (66)$$

$$\mathbf{E} \left\{ \mathbf{v}(l) \mathbf{e}_n^T(k) \right\} = -\mathbf{R} \mathbf{K}_n^T \Phi^T \left[(\mathbf{I} - \mathbf{K}_n \mathbf{H})^T \Phi^T \right]^{k-l-1} \quad (l < k) \quad (67)$$

Now assume that $l = k$. In this case,

$$\mathbf{E} \left\{ \mathbf{e}_m(k) \mathbf{e}_n^T(k) \right\} = \mathbf{P}_{mn}(k) \quad (68)$$

$$\mathbb{E} \left\{ \mathbf{v}(k) \mathbf{e}_n^T(k) \right\} = 0 \quad (69)$$

$$\mathbb{E} \left\{ \mathbf{e}_m(k) \mathbf{v}^T(k) \right\} = 0 \quad (70)$$

$$\mathbb{E} \left\{ \mathbf{v}(k) \mathbf{v}^T(k) \right\} = \mathbf{R} \quad (71)$$

Combining equations (58) and (63) through (71) gives the correlation matrix for the filter residuals as

$$\mathbb{E} \left\{ \gamma_m(l) \gamma_n^T(k) \right\} = \begin{cases} \mathbf{H} [\Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H})]^{l-k} \mathbf{P}_{mn}(k) \mathbf{H}^T \\ \quad - \mathbf{H} \{ \Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H}) \}^{l-k-1} \Phi \mathbf{K}_m \mathbf{R} & (l > k) \\ \mathbf{H} \mathbf{P}_{mn}(k) \mathbf{H}^T + \mathbf{R} & (l = k) \\ \mathbf{H} \mathbf{P}_{mn}(l) [(\mathbf{I} - \mathbf{K}_n \mathbf{H})^T \Phi^T]^{k-l} \mathbf{H}^T \\ \quad - \mathbf{R} \mathbf{K}_n^T \Phi^T [(\mathbf{I} - \mathbf{K}_n \mathbf{H})^T \Phi^T]^{k-l-1} \mathbf{H}^T & (l < k) \end{cases} \quad (72)$$

Now, we restate the general term σ_{gh}^4 (given previously in eq. (56)) in the matrix of second moments of the sample correlations of the residuals:

$$\begin{aligned} \sigma_{gh}^4 &= \frac{1}{w^2} \sum_{l=1}^w \sum_{k=1}^w \left[\mathbb{E} \left\{ \gamma_m^\alpha(l) \gamma_m^\beta(l) \right\} \mathbb{E} \left\{ \gamma_n^\theta(k) \gamma_n^\delta(k) \right\} \right. \\ &\quad + \mathbb{E} \left\{ \gamma_m^\alpha(l) \gamma_n^\theta(k) \right\} \mathbb{E} \left\{ \gamma_m^\beta(l) \gamma_n^\delta(k) \right\} \\ &\quad \left. + \mathbb{E} \left\{ \gamma_m^\alpha(l) \gamma_n^\delta(k) \right\} \mathbb{E} \left\{ \gamma_m^\beta(l) \gamma_n^\theta(k) \right\} \right] \end{aligned}$$

Each term $\mathbb{E} \left\{ \gamma_m^\mu(l) \gamma_n^\nu(k) \right\}$ is given by the entry in the μ th row and the ν th column of the residual correlation matrix in equations (72). The general term then becomes in the steady state

$$\begin{aligned} \sigma_{gh}^4 &= [\mathbf{H} \mathbf{P}_{mn} \mathbf{H}^T + \mathbf{R}]^{\alpha\beta} [\mathbf{H} \mathbf{P}_{mn} \mathbf{H}^T + \mathbf{R}]^{\theta\delta} \\ &\quad + \frac{1}{w} \left\{ [\mathbf{H} \mathbf{P}_{mn} \mathbf{H}^T + \mathbf{R}]^{\alpha\theta} [\mathbf{H} \mathbf{P}_{mn} \mathbf{H}^T + \mathbf{R}]^{\beta\delta} \right. \\ &\quad \left. + [\mathbf{H} \mathbf{P}_{mn} \mathbf{H}^T + \mathbf{R}]^{\alpha\delta} [\mathbf{H} \mathbf{P}_{mn} \mathbf{H}^T + \mathbf{R}]^{\beta\theta} \right\} \\ &\quad + \frac{1}{w^2} \sum_{l=1}^{w-1} \sum_{k=l+1}^w \left\{ [\mathbf{C}_n^T (\chi_n^T)^{k-l-1} \mathbf{H}^T]^{\alpha\theta} [\mathbf{C}_n^T (\chi_n^T)^{k-l-1} \mathbf{H}^T]^{\beta\delta} \right. \\ &\quad \left. + [\mathbf{C}_n^T (\chi_n^T)^{k-l-1} \mathbf{H}^T]^{\alpha\delta} [\mathbf{C}_n^T (\chi_n^T)^{k-l-1} \mathbf{H}^T]^{\beta\theta} \right\} \\ &\quad + \frac{1}{w^2} \sum_{l=2}^w \sum_{k=1}^{l-1} \left\{ [\mathbf{H} \chi_m^{l-k-1} \mathbf{C}_m]^{\alpha\theta} [\mathbf{H} \chi_m^{l-k-1} \mathbf{C}_m]^{\beta\delta} \right. \\ &\quad \left. + [\mathbf{H} \chi_m^{l-k-1} \mathbf{C}_m]^{\alpha\delta} [\mathbf{H} \chi_m^{l-k-1} \mathbf{C}_m]^{\beta\theta} \right\} \end{aligned} \quad (73)$$

where

$$\chi_m = \Phi (\mathbf{I} - \mathbf{K}_m \mathbf{H}) \quad (74)$$

$$\chi_n^T = (\mathbf{I} - \mathbf{K}_n \mathbf{H})^T \Phi^T \quad (75)$$

$$\mathbf{C}_m = \Phi \left[(\mathbf{I} - \mathbf{K}_m \mathbf{H}) \mathbf{P}_{mn} \mathbf{H}^T - \mathbf{K}_m \mathbf{R} \right] \quad (76)$$

$$\mathbf{C}_n^T = \left[\mathbf{H} \mathbf{P}_{mn} (\mathbf{I} - \mathbf{K}_n \mathbf{H})^T - \mathbf{R} \mathbf{K}_n^T \right] \Phi^T \quad (77)$$

and $[\mathbf{Z}]^{\mu\nu}$ denotes the entry from row μ and column ν of matrix \mathbf{Z} . In equation (73) the steady-state form \mathbf{P}_{mn} of the error covariance was used because the sample correlation (eq. (27)) is computed in the usual application of this algorithm over a window of data of length w which begins after the filter has reached steady state. The double summations in equation (73) can be simplified to single summations as follows:

$$\begin{aligned} & \frac{1}{w^2} \sum_{l=2}^w \sum_{k=1}^{l-1} \left\{ \left[\mathbf{H} \chi_m^{l-k-1} \mathbf{C}_m \right]^{\alpha\theta} \left[\mathbf{H} \chi_m^{l-k-1} \mathbf{C}_m \right]^{\beta\delta} \right. \\ & \left. + \left[\mathbf{H} \chi_m^{l-k-1} \mathbf{C}_m \right]^{\alpha\delta} \left[\mathbf{H} \chi_m^{l-k-1} \mathbf{C}_m \right]^{\beta\theta} \right\} \\ & = \frac{1}{w^2} \sum_{j=1}^{w-1} (w-j) \left\{ [\mathbf{Z}]^{\alpha\theta} [\mathbf{Z}]^{\beta\delta} + [\mathbf{Z}]^{\alpha\delta} [\mathbf{Z}]^{\beta\theta} \right\} \end{aligned} \quad (78)$$

where

$$\mathbf{Z} = \mathbf{H} \chi_m^{j-1} \mathbf{C}_m \quad (79)$$

Equations (46), (50) through (53), (56), and (72) through (78) can be used to calculate the mean vector and second-moment matrix of the vector $\underline{\hat{\xi}}$ of estimates of the plant and sensor noise variances \mathbf{Q}_{ii} and \mathbf{R}_{jj} , respectively, assuming that \mathbf{Q} and \mathbf{R} are diagonal. The covariance matrix of the estimate $\underline{\hat{\xi}}$ is easily found from

$$\begin{aligned} \Sigma_{\xi\xi} &= \mathbf{E} \left\{ \left[\underline{\hat{\xi}} - \mathbf{E} \left\{ \underline{\hat{\xi}} \right\} \right] \left[\underline{\hat{\xi}} - \mathbf{E} \left\{ \underline{\hat{\xi}} \right\} \right]^T \right\} \\ &= \mathbf{E} \left\{ \underline{\hat{\xi}} \underline{\hat{\xi}}^T \right\} - \mathbf{E} \left\{ \underline{\hat{\xi}} \right\} \mathbf{E} \left\{ \underline{\hat{\xi}} \right\}^T \\ &= \mathbf{M}^\# \mathbf{E} \left\{ \underline{\hat{\mathbf{Y}}} \underline{\hat{\mathbf{Y}}}^T \right\} \left(\mathbf{M}^\# \right)^T - \left(\mathbf{M}^\# \mathbf{E} \left\{ \underline{\hat{\mathbf{Y}}} \right\} \right) \left(\mathbf{M}^\# \mathbf{E} \left\{ \underline{\hat{\mathbf{Y}}} \right\} \right)^T \end{aligned} \quad (80)$$

The standard deviations of the estimates of the noise variances are just the square roots of the diagonal terms of $\Sigma_{\xi\xi}$. Step (7) of the development of the general solution is now complete.

Examples

To demonstrate the use and performance of the technique for estimating noise covariances, the procedure was applied to two examples: one example was a ring laser gyro used by Friedland (ref. 1), and the second example was a damped Schuler loop used by Mehra (ref. 5). For both examples the covariance matrices of the plant and sensor noises were diagonal. For each example, equations (46) and (50) were used to calculate the theoretical mean $\mathbf{E} \left\{ \underline{\hat{\xi}} \right\}$ of the estimate of the variance $\underline{\hat{\xi}}$. Equations (51) through (53), (73), and (78) were used to calculate the second moment $\mathbf{E} \left\{ \underline{\hat{\xi}} \underline{\hat{\xi}}^T \right\}$ of the estimate, and equation (80) was used to calculate the variance $\Sigma_{\xi\xi}$ and hence the standard deviation of the estimate. These statistics were computed as functions of w , the number of samples averaged to obtain the sample correlation of the residuals.

In addition, each example system was implemented in digital simulation together with the variance estimation algorithm (eq. (39)). The sample statistics (mean and standard deviation) of the estimates were

computed using a sample of 100 runs of the simulation, each with a different seed for the random number generator. In the simulation the sample correlation of the residuals (eq. (27)) was calculated using a first-order recursive low-pass filter of the form

$$\hat{\mathbf{Y}}_i(k+1) = [1 - (1/w)] \hat{\mathbf{Y}}_i(k) + (1/w) \gamma_i(k) \gamma_i^T(k) \quad (81)$$

This approach used less computer memory than storing a window of residuals (window length denoted by w). Values of w of 200, 400, 800, 1600, and 4000 samples were used.

Friedland's Example

The laser gyro example (ref. 1) was a two-state system with angular position and angular rate as the states. The discrete-time version of the system for a sample time of 15 sec resulted in a state transition matrix as follows:

$$\Phi = \begin{bmatrix} 1 & 15 \\ 0 & 1 \end{bmatrix} \quad (82)$$

The observed quantity was the angular position, and thus the observation matrix was given by

$$\mathbf{H} = [1 \quad 0] \quad (83)$$

The noise input matrix was assumed to be

$$\Gamma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (84)$$

The problem was to solve for estimates $\hat{\xi}$ of the unknown diagonal elements of the plant and measurement noise covariance matrices \mathbf{Q} and \mathbf{R} . However, to implement the system in simulation, the following true values for the noise covariances \mathbf{Q} and \mathbf{R} were assumed:

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 0.25 \times 10^{-4} \end{bmatrix} \quad (85)$$

$$\mathbf{R} = 0.25 \quad (86)$$

For this example, the dimensions were $r = 1$ and $q = 2$, and

$$p(r^2 + r)/2 \geq q + r \quad (87)$$

or

$$p \geq 3$$

Therefore, three constant gain observers were employed with Friedland's values (ref. 1) for the observer gains. These gains were obtained by Friedland using the Kalman filter algorithm and assuming a different set of values for the noise covariances for each observer.

The theoretical means of the estimates are plotted in figure 1(a) together with the sample means. As expected for an unbiased estimator, the theoretical means equal the true values, and the sample means show good agreement with the theoretical and true values. The theoretical and sample standard deviations are plotted in figure 1(b). The theoretical and sample values agree fairly well for \mathbf{Q}_{22} and \mathbf{R}_{11} . For \mathbf{Q}_{11} the sample standard deviations are slightly larger than the theoretical values but follow the same general trend. The theoretical standard deviations are replotted in figure 1(c) in terms of percent of the true value. As in figure 1(b) note that the standard deviations are approximately proportional to the inverse of the square root of the number of samples averaged to obtain the sample correlation of the residuals. Note also that with these gains the algorithm is least accurate in estimating \mathbf{Q}_{22} and most accurate in estimating \mathbf{Q}_{11} , which can be

estimated with a standard deviation of less than 20 percent of its true value when 300 or more samples are averaged.

In four simulation runs Friedland obtained estimates (with $w = 100$) ranging from 0.88 to 1.27 for \mathbf{Q}_{11} , from 0.142×10^{-4} to 0.35×10^{-4} for \mathbf{Q}_{22} , and from 0.17 to 0.36 for \mathbf{R}_{11} (ref. 1). These estimates all fall within one theoretical standard deviation of the true values.

Mehra's Example

Mehra's example had two states for the loop dynamics and three augmented states to account for the correlated noise resulting in a combined five-state system. The system matrices were as follows:

$$\Phi = \begin{bmatrix} 0.75 & -1.74 & -0.3 & 0 & -0.15 \\ 0.09 & 0.91 & -0.0015 & 0 & -0.008 \\ 0 & 0 & 0.95 & 0 & 0 \\ 0 & 0 & 0 & 0.55 & 0 \\ 0 & 0 & 0 & 0 & 0.905 \end{bmatrix} \quad (88)$$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix} \quad (89)$$

$$\Gamma = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 24.64 & 0 & 0 \\ 0 & 0.835 & 0 \\ 0 & 0 & 1.83 \end{bmatrix} \quad (90)$$

For purposes of simulation, the assumed values for the noise covariance matrices were as follows:

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (91)$$

$$\mathbf{R} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (92)$$

For this example, $r = 2$ and $q = 3$, and

$$p(r^2 + r)/2 \geq q + r \quad (93)$$

or

$$p \geq 2$$

Both two-observer and three-observer estimators were evaluated. The following observer gains were used:

$$\mathbf{K}_1 = \begin{bmatrix} 0.953 & 0.772 \\ 0.28 \text{ E}-2 & 0.338 \\ -3.5 & -1.49 \\ -0.176 \text{ E}-3 & 0.35 \\ 0.0319 & -0.77 \end{bmatrix} \quad (94)$$

$$\mathbf{K}_2 = \begin{bmatrix} 0.99 & 1.3 \\ 0.5 \text{ E}-2 & 0.26 \\ -2.9 & -0.77 \\ -0.3 \text{ E}-3 & 0.21 \\ 0.04 & -0.816 \end{bmatrix} \quad (95)$$

$$\mathbf{K}_3 = \begin{bmatrix} 1.03 & 0.5 \\ 0.5 \text{ E}-2 & 0.18 \\ -1 & -1.25 \\ -0.5 \text{ E}-3 & 0.22 \\ 0.05 & -0.54 \end{bmatrix} \quad (96)$$

These gains were not chosen through a systematic process to optimize the estimator performance. However, they were selected through trial and error to be at least in the vicinity of a local minimum for the theoretical standard deviation of the estimates. Gains \mathbf{K}_1 and \mathbf{K}_2 were used for the two-observer case.

The theoretical means of the estimates for the three-observer case are plotted in figure 2(a) together with the sample means, and as expected the theoretical means equal the true values. Again, the sample means show good agreement with the theoretical and true values. The theoretical and sample standard deviations are plotted in figure 2(b) and are in good agreement. The algorithm does an excellent job of estimating \mathbf{Q}_{11} , but not as good of estimating \mathbf{R}_{11} . The relative accuracies of the estimates can be changed by adjusting the observer gains. In the two-observer case with these gain matrices, the standard deviation of the estimates of \mathbf{Q}_{11} and \mathbf{Q}_{22} were not greatly different from the three-observer case, but the estimates of \mathbf{Q}_{33} and \mathbf{R}_{11} were much worse with standard deviations that were larger by a factor of 7 to 8.

Figure 3 shows time history plots of the estimated values of $\hat{\xi}$ of the noise variances for $w = 200$ (fig. 3(a)) and $w = 4000$ (fig. 3(b)) for the three-observer estimator. The estimates with $w = 4000$ are, of course, much smoother. Note in figure 3(b) the extremely smooth curve for $\hat{\xi}_1 = \hat{\mathbf{Q}}(1,1)$ as predicted by the standard deviation of approximately 0.04 for this estimate. On the other hand, the estimate $\hat{\xi}_4 = \hat{\mathbf{R}}(1,1)$ is much noisier commensurate with an expected standard deviation of approximately 0.7. Note that the estimate $\hat{\mathbf{R}}(1,1)$ appears to take on negative values at times. (The time history plots of $\hat{\mathbf{R}}(1,1)$ in fig. 3 are limited between 0 and 2.) These negative values are caused by the values of the sample correlation $\hat{\mathbf{T}}$ of the residuals not being accurate estimates of the true correlation \mathbf{T} , particularly with $w = 200$.

Although the algorithm was derived on the assumption of constant plant and sensor noise variances, the technique can be used to estimate slowly varying noise variances, that is, slowly varying in comparison to the time constant of the averaging filter, or the number of samples averaged to compute the sample correlation of the residuals. An idea of the transient response of the algorithm can be obtained from figure 4, which contains time history plots for the case where the original noise variances are increased in a step-function manner by a factor of 3 at sample number 5000. Following the increase the estimates are seen to approach the true values slowly, except for $\hat{\xi}_3 = \hat{\mathbf{Q}}(3,3)$. Even though $\hat{\mathbf{Q}}(3,3)$ does not approach its true value during the period plotted, its value after the step increase is larger than in figure 3, where the true variances were held constant. Of course, the transient response would have been faster for smaller values of w , even though the estimate would have been "noisier".

This ability to estimate slowly changing variances is important for practical applications, where the plant and measurement noise variances may not be strictly constant. This capability may be most important in real-time applications in which time does not permit the data to be examined and possibly broken down into segments of nearly constant noise variances for subsequent processing.

Concluding Remarks

A general solution has been developed for extending the technique proposed by Friedland for estimating the plant and sensor noise covariance matrices of a discrete, time-invariant linear system using multiple observers.

Furthermore, solutions were developed for the mean and the covariance of the estimates. The general solution was specialized to the case in which the noise covariance matrices are diagonal, and this algorithm was demonstrated on two examples.

In the examples the accuracy of the estimates as measured by their theoretical standard deviations, as well as by their sample means and sample standard deviations, was very good for some of the components of the covariance, but not so good for others. The accuracy depends on the number of observers used in the estimator, the number of terms averaged in computing the sample correlation of the observer residuals, the observer gains, and the relationship between the observations and the system states. The accuracy of the estimator can and should be investigated for each application using the expression for the theoretical variance of the estimate. Optimization of the gains to minimize the error in the estimates was not investigated. It should be noted that for large systems the technique requires considerable computer memory, since some of the matrices get quite large.

The development of the solution assumed that the noise variances were constant. However, one of the examples demonstrated that the technique can be used to estimate slowly changing noise variances if the change is sufficiently slow compared with the time over which the products of the residuals are averaged to compute the sample correlation.

An area for future research is the investigation of a procedure for choosing the observer gains to optimize the accuracy of the estimator. Also, further work to compare this technique with others in terms of accuracy of the estimates and practicality of implementation of the algorithms would be beneficial.

NASA Langley Research Center
Hampton, VA 23665-5225
November 10, 1987

Appendix

Some Matrix Operations

Kronecker Product

Consider an $m \times n$ matrix \mathbf{A} and an arbitrary matrix \mathbf{B} , as shown in equation (16). Define the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ (ref. 11, p. 22) as

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}$$

For example, suppose \mathbf{A} is 2×3 and \mathbf{B} is 2×2 . Then,

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} & a_{13}b_{11} & a_{13}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} & a_{13}b_{21} & a_{13}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} & a_{23}b_{11} & a_{23}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} & a_{23}b_{21} & a_{23}b_{22} \end{bmatrix}$$

Stacking Operator

Define the stacking operator $S(\cdot)$ such that if \mathbf{a}_i represents the rows of the $m \times n$ matrix \mathbf{A} , then $S(\mathbf{A})$ is an $mn \times 1$ column vector as follows (see eq. (17)):

$$S(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1^T \\ \text{---} \\ \mathbf{a}_2^T \\ \text{---} \\ \cdot \\ \cdot \\ \text{---} \\ \mathbf{a}_m^T \end{bmatrix}$$

For example, suppose \mathbf{A} is 2×3 . Then,

$$S(\mathbf{A}) = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \end{bmatrix}$$

G Matrix

Suppose that \mathbf{Q} and \mathbf{R} are both 2×2 and diagonal. Then, the stack of Ξ is given by

$$S(\Xi) = \begin{bmatrix} q_{11} \\ 0 \\ 0 \\ 0 \\ 0 \\ q_{22} \\ 0 \\ 0 \\ 0 \\ 0 \\ r_{11} \\ 0 \\ 0 \\ 0 \\ 0 \\ r_{22} \end{bmatrix} = \begin{bmatrix} \xi_1 \\ 0 \\ 0 \\ 0 \\ 0 \\ \xi_2 \\ 0 \\ 0 \\ 0 \\ 0 \\ \xi_3 \\ 0 \\ 0 \\ 0 \\ 0 \\ \xi_4 \end{bmatrix}$$

and the vector ξ is

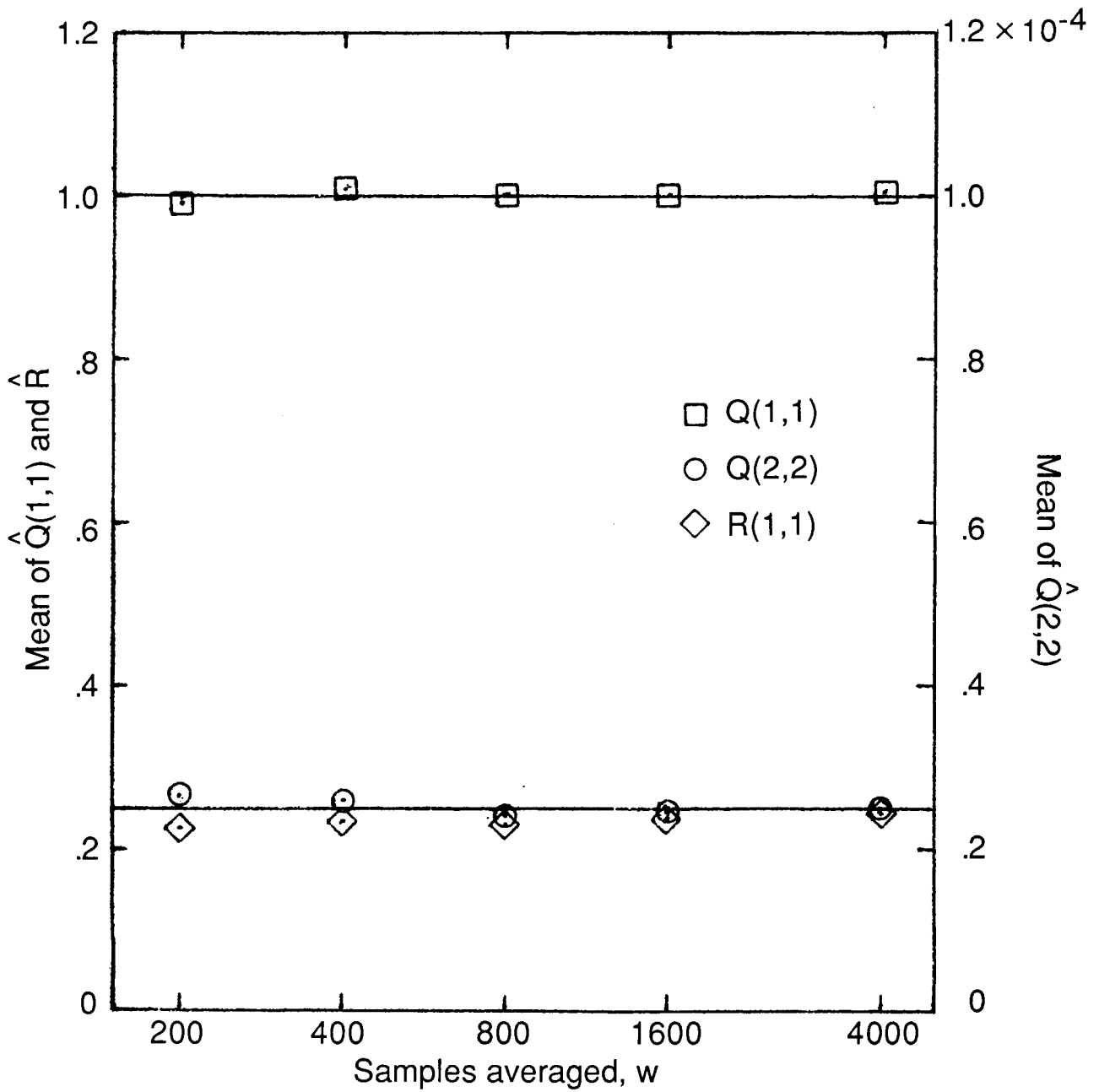
$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix}$$

For this case the matrix \mathbf{G} is defined by the following:

$$\mathbf{G}^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

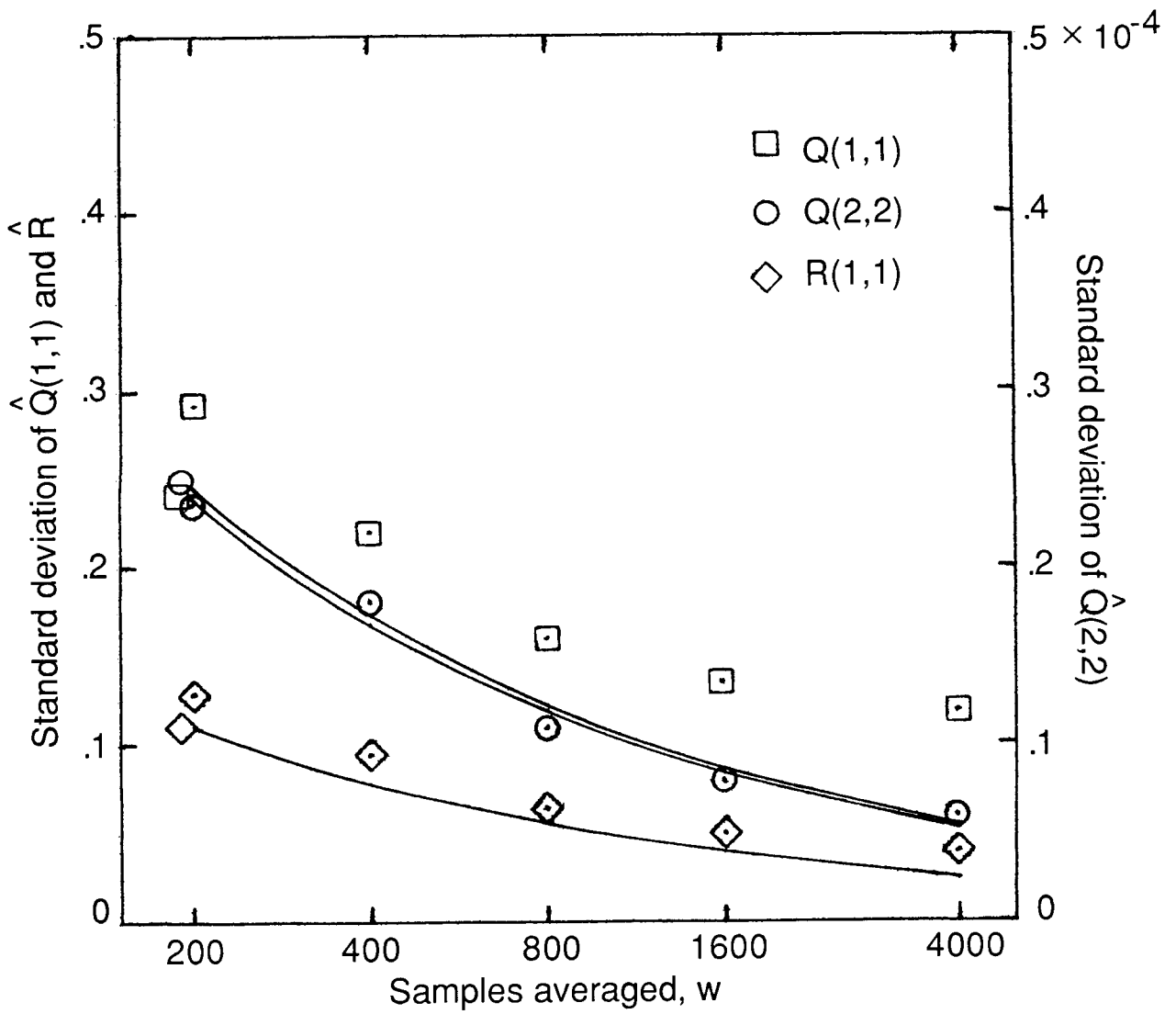
References

1. Friedland, Bernard: Estimating Noise Variances by Using Multiple Observers. *IEEE Trans. Aerosp. & Electron. Systems*, vol. 18, no. 4, July 1982, pp. 442-448.
2. Bundick, W. Thomas: *A Preliminary Evaluation of a Failure Detection Filter for Detecting and Identifying Control Element Failures in a Transport Aircraft*. NASA TM-87576, 1985.
3. Bundick, W. Thomas: *A Preliminary Evaluation of the Generalized Likelihood Ratio for Detecting and Identifying Control Element Failures in a Transport Aircraft*. NASA TM-87620, 1985.
4. Bonnice, W.; Wagner, E.; Hall, S.; and Motyka, P.: *The Evaluation of the OSGLR Algorithm for Restructurable Control*. NASA CR-178083, 1986.
5. Mehra, Raman K.: On the Identification of Variances and Adaptive Kalman Filtering. *IEEE Trans. Autom. Control*, vol. AC-15, no. 2, Apr. 1970, pp. 175-184.
6. Mehra, Raman K.: Approaches to Adaptive Filtering. *IEEE Trans. Autom. Control*, vol. AC-17, no. 5, Oct. 1972, pp. 693-698.
7. Belanger, Pierre R.: Estimation of Noise Covariance Matrices for a Linear Time-Varying Stochastic Process. *Proceedings of the IFAC 5th World Congress, Part 4*, International Federation of Automatic Control, c.1971, pp. 38.3 1-38.3 7.
8. Godbole, S. S.: Kalman Filtering With No A Priori Information About Noise—White Noise Case: Identification of Covariances. *IEEE Trans. Autom. Control*, vol. AC-19, no. 5, Oct. 1974, pp. 561-563.
9. Myers, Kenneth A.; and Tapley, Byron D.: Adaptive Sequential Estimation With Unknown Noise Statistics. *IEEE Trans. Autom. Control*, vol. AC-21, no. 4, Aug. 1976, pp. 520-523.
10. Dee, Dick P.; Cohn, Stephen E.; Dalcher, Amnon; and Ghil, Michael: An Efficient Algorithm for Estimating Noise Covariances in Distributed Systems. *IEEE Trans. Autom. Control*, vol. AC-30, no. 11, Nov. 1985, pp. 1057-1065.
11. Davis, Philip J.: *Circulant Matrices*. John Wiley & Sons, Inc., c.1979.
12. MacDuffee, C. C.: *The Theory of Matrices*, Corrected Reprint of First Edition. Chelsea Publ. Co., 1946.
13. Kailath, Thomas: *Linear Systems*. Prentice-Hall, Inc., c.1980.
14. Papoulis, Athanasios: *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Book Co., Inc., c.1965.



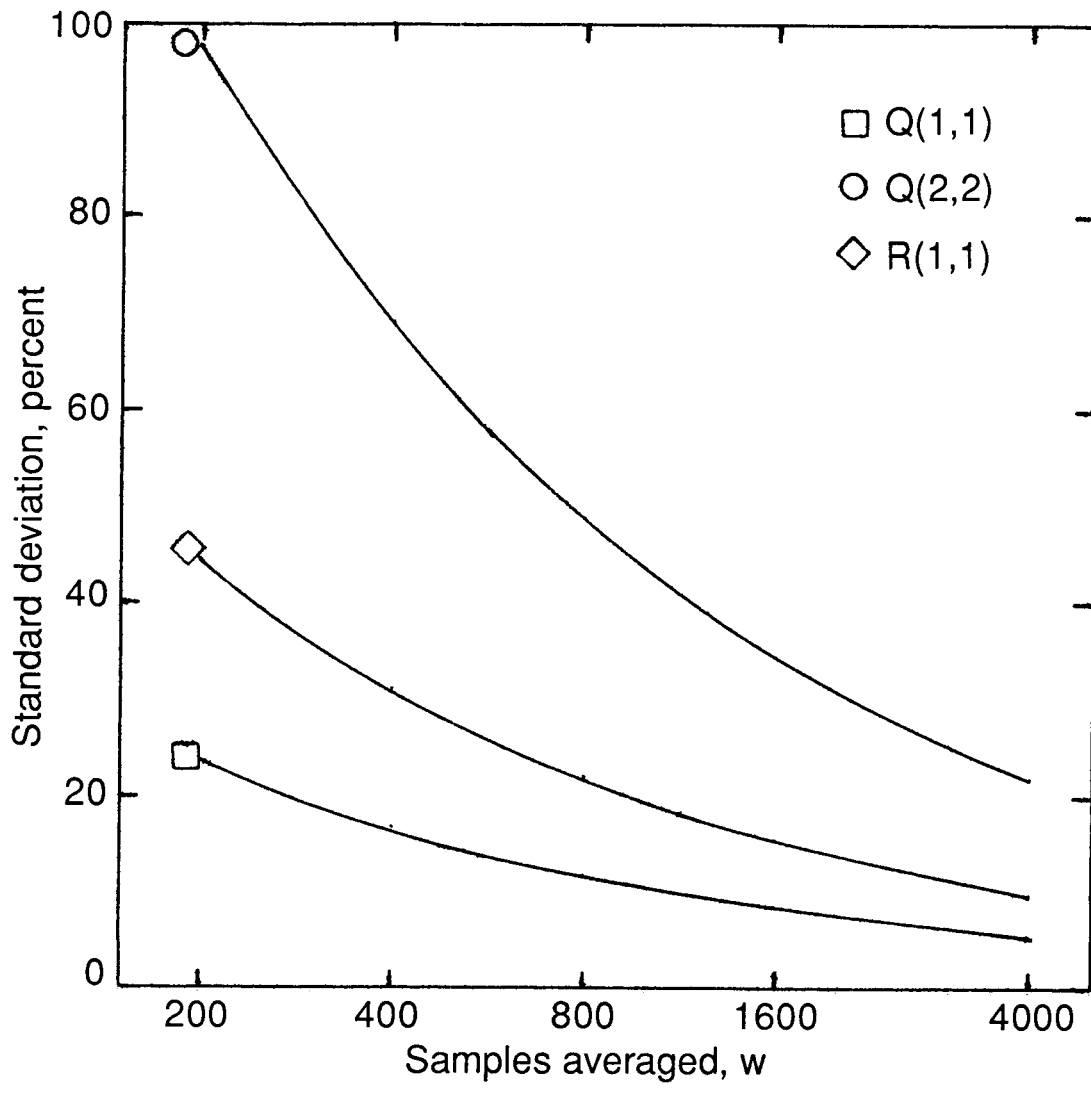
(a) Mean.

Figure 1. Theoretical (solid lines) and sample (points) statistics of estimates of noise variances for Friedland's example.



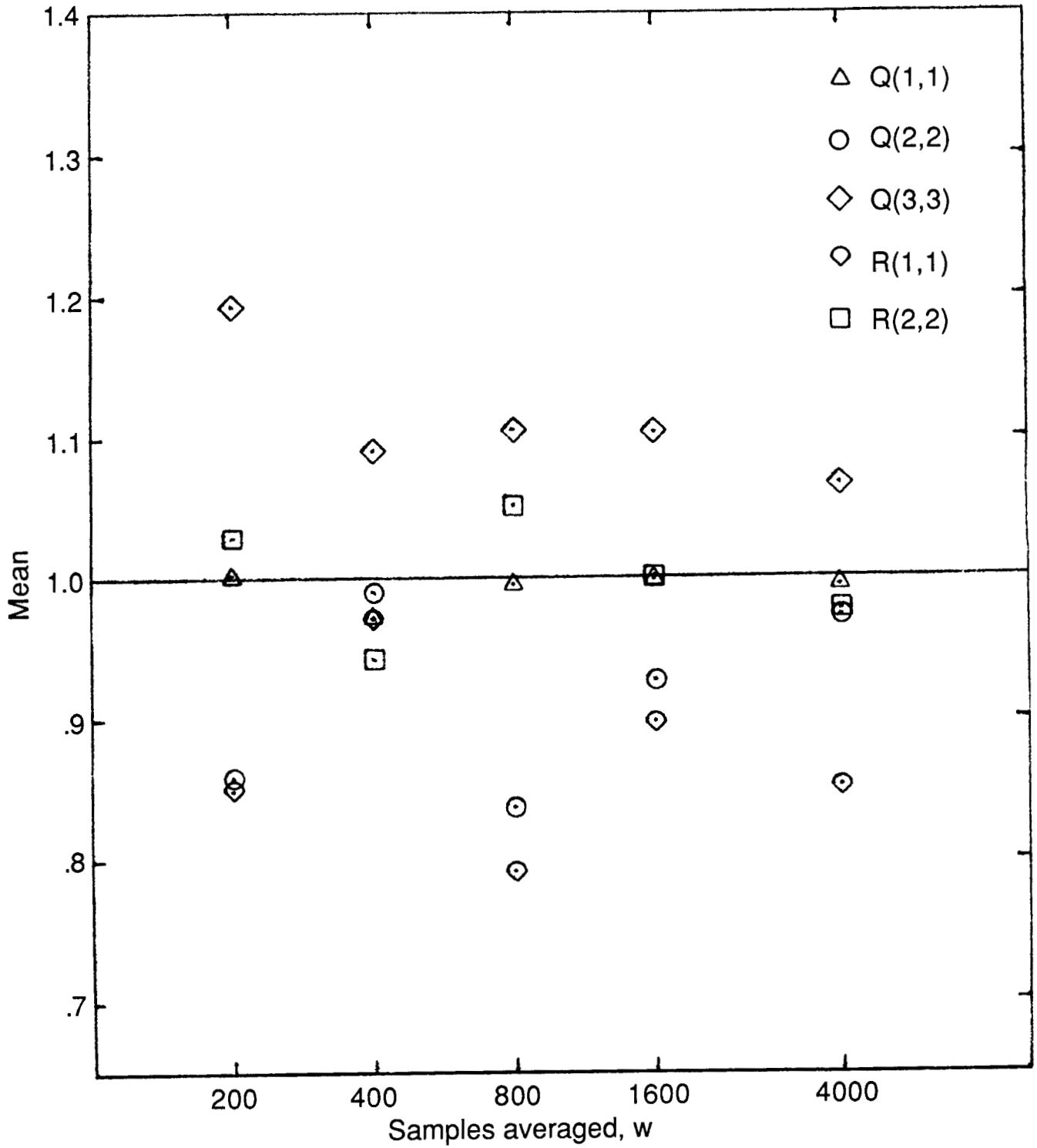
(b) Standard deviation.

Figure 1. Continued.



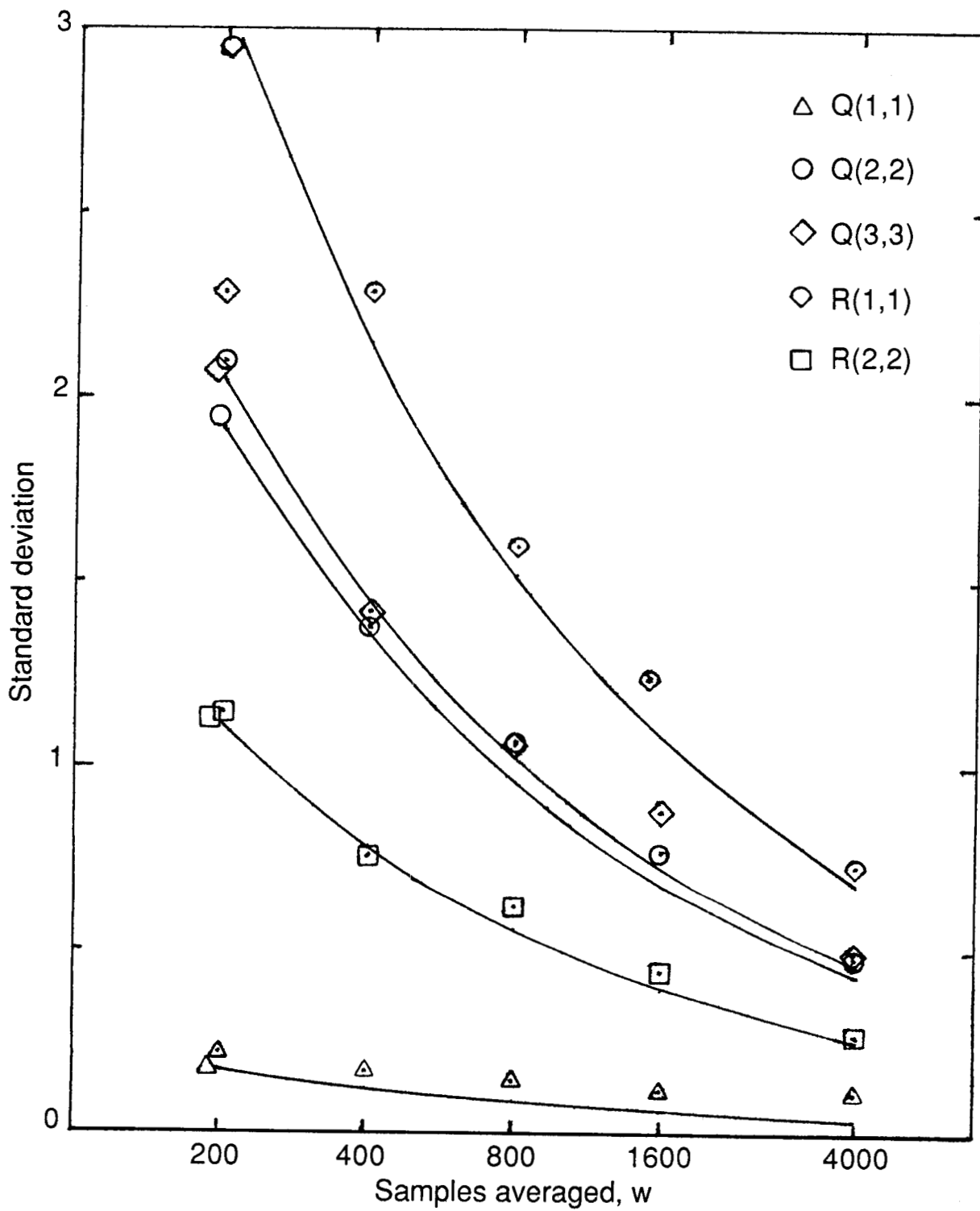
(c) Standard deviation in percent of true value.

Figure 1. Concluded.



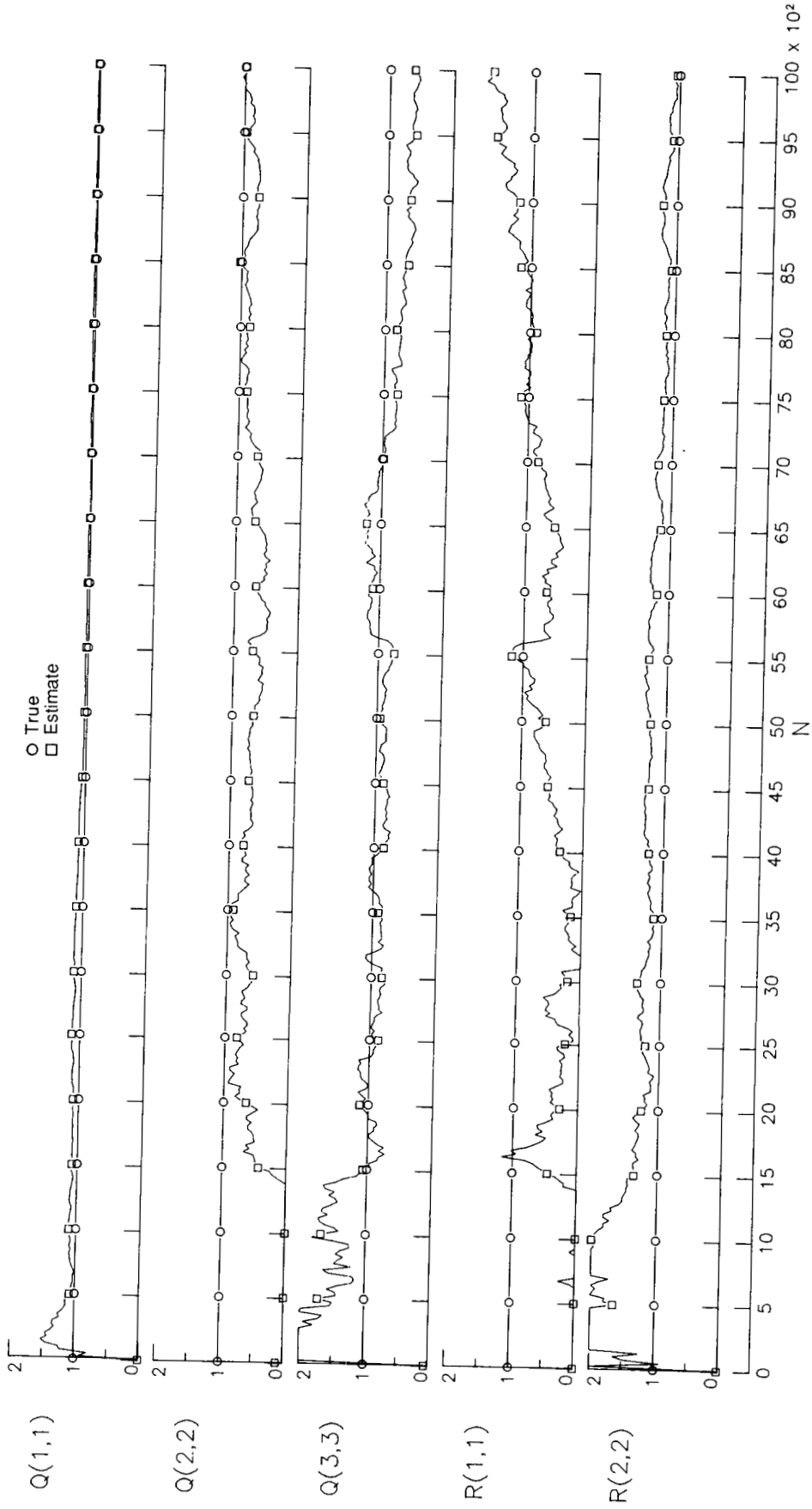
(a) Mean.

Figure 2. Theoretical (solid line) and sample (points) statistics of estimates of noise variances for Mehra's example.



(b) Standard deviation.

Figure 2. Concluded.



(b) $w = 4000$.

Figure 3. Concluded.

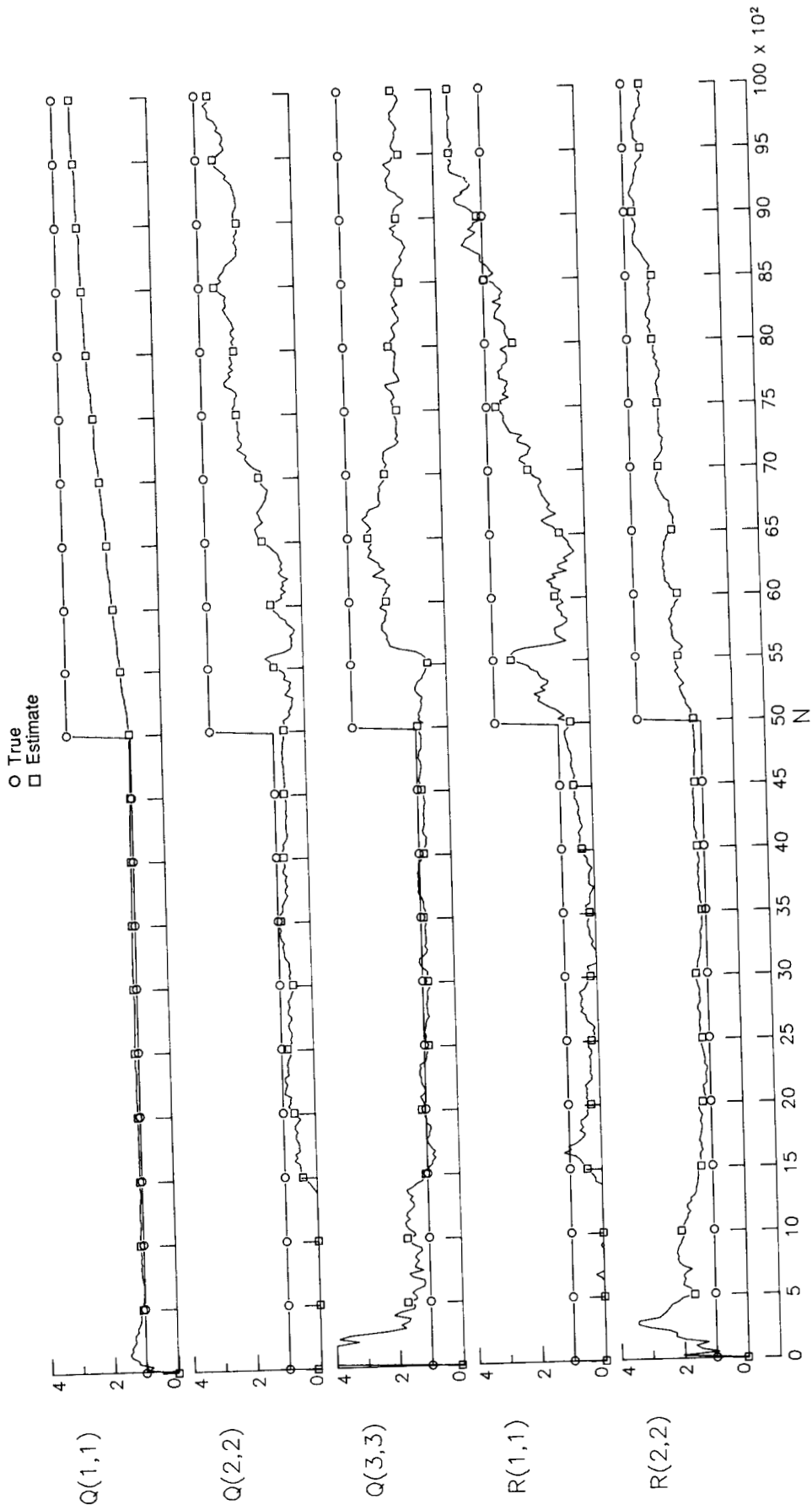


Figure 4. Time history of estimates for Mehra's example with step increase in noise variances.



Report Documentation Page

1. Report No. NASA TM-4020	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Development of a Technique for Estimating Noise Covariances Using Multiple Observers		5. Report Date February 1988	
		6. Performing Organization Code	
7. Author(s) W. Thomas Bundick		8. Performing Organization Report No. L-16349	
		10. Work Unit No. 505-61-01-07	
9. Performing Organization Name and Address NASA Langley Research Center Hampton, VA 23665-5225		11. Contract or Grant No.	
		13. Type of Report and Period Covered Technical Memorandum	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546-0001		14. Sponsoring Agency Code	
		15. Supplementary Notes Programming support was provided by Jessie Yeager of Planning Research Corporation, Hampton, Virginia.	
16. Abstract Friedland's technique for estimating the unknown noise variances of a linear system using multiple observers has been extended by developing a general solution for the estimates of the variances, developing the statistics (mean and standard deviation) of these estimates, and demonstrating the solution on two examples.			
17. Key Words (Suggested by Authors(s)) Linear systems Covariance matrix Estimation		18. Distribution Statement Unclassified—Unlimited Subject Category 08	
19. Security Classif.(of this report) Unclassified	20. Security Classif.(of this page) Unclassified	21. No. of Pages 31	22. Price A03