A Derivation of the Optimum Continuous Linear Estimator for Systems with Correlated Measurement Noise

J. J. Deyst*
Massachusetts Institute of Technology, Cambridge, Mass.

A derivation is presented for differential equations describing the minimum-variance continuous linear filter for separating signals from additive correlated noise. The filter equations are derived by applying discrete time-optimum estimation formulas and taking appropriate limits. If the measurements available to the filter do not all contain first integrals of white noise, the equations are singular and it is shown that the optimum filter must be preceded by successive differentiations until first integrals of white noise are attained. A method is presented for reducing the order of the equations by the number of input processes available to the filter. Consideration is also given to the filter startup procedure, and it is found that appropriate initial conditions cannot be determined a priori, but are functions of the initial values of the measurements.

I. Introduction

In a number of filtering problems, measurement noise superimposed on signal processes is correlated in time. Bryson and Henrikson[1] have solved the discrete correlated measurement noise problem. In the continuous correlated noise case, the filter equations developed by Kalman and Bucy[6] are singular. Bucy[3] considers the correlated noise filter and obtains filter equations. Bryson and Johansen[2] develop a method of differentiating the measurement processes to attain white noise in the measurements. The Kalman-Bucy white noise filter is then applied to the resulting system to produce the minimum-variance estimator.

The purpose of this paper is to offer an alternative derivation of the minimum-variance continuous correlated measurement noise filter. The derivation is straightforward and may provide insight into the filtering problem. Furthermore, the derivation does not involve differentiation of first integrals of white noise and hence evades questions of existence of such derivatives.

II. Problem Statement

Let $x(t)$ represent the kth-order state vector of a dynamical system satisfying the differential equation†

$$\dot{x}(t) = F(t)x(t) + q(t) \tag{1}$$

where

$x(t) \equiv$ system state vector of dimension k
$F(t) \equiv$ system dynamics matrix $(k \times k)$, continuous with continuous derivative
$q(t) \equiv$ white Gaussian process noise of dimension k

and the process noise statistics are

$$E[q(t)] = 0,\ E[q(t)q^T(t+\tau)] = Q(t)\delta(\tau)$$

$Q(t) \equiv$ non-negative-definite matrix $(k \times k)$ continuous with continuous derivative
$\delta(\tau) \equiv$ Dirac delta function

Measurements available to the filter are composed of linear combinations of elements of $x(t)$. Thus if $m(t)$ is the vector of measurement processes available to the filter then

$$m(t) = H(t)x(t) \tag{2}$$

where $H(t) =$ measurement matrix $(p \times k)$, continuous with continuous first and second derivatives.

Equations (1) and (2) define the basic filtering problem to be solved. If $H(t)$ is a square, nonsingular matrix, then the problem is trivial because $x(t)$ is obtained directly from $m(t)$ by inverting $H(t)$. Of interest, in terms of filter theory, are those cases for which $H(t)$ has more columns than linearly independent rows. Thus the filtering problem is to estimate $x(t)$ from measurements comprised of linear combinations of its elements. Minimum mean square error is chosen as the criterion for optimal estimation. Since all processes are Gaussian, the optimal estimate is the mean of $x(t)$ conditioned on the measurement history;

$$\hat{x}(t) = E[x(t)|m(\tau)],\quad 0 \leq \tau \leq t$$

The problem of estimating signals in the presence of correlated measurement noise is a special case of the system (1), (2). Let $x_1(t)$ be the system state to be estimated and assume it satisfies the differential equation

$$\dot{x}_1(t) = F_1(t)x_1(t) + q_1(t) \tag{3}$$

Similarly, let the correlated measurement noise be generated by a system whose state is $x_2(t)$ where

$$\dot{x}_2(t) = F_2(t)x_2(t) + q_2(t) \tag{4}$$

and $q_1(t)$, $q_2(t)$ are independent Gaussian white noises

$$E[q_1(t)] = 0,\quad E[q_1(t)q_1^T(t+\tau)] = Q_1(t)\delta(\tau)$$
$$E[q_2(t)] = 0,\quad E[q_2(t)q_2^T(t+\tau)] = Q_2(t)\delta(\tau)$$
$$E[q_1(t)q_2^T(t+\tau)] = 0$$

Measurements available to the filter are comprised of linear combinations of elements of $x_1(t)$ corrupted by elements of $x_2(t)$

$$m(t) = H_1(t)x_1(t) + H_2(t)x_2(t) \tag{5}$$

The problem posed by (3–5) may be readily cast into the form

Presented as Paper 69-73 at the AIAA 7th Aerospace Sciences Meeting, New York, January 20–22, 1969; submitted February 5, 1969; revision received June 11, 1969. This research was performed under the sponsorship of the Electronics Research Center of NASA through NASA Grant NGR-22-009-207.

* Assistant Professor, Department of Aeronautics and Astronautics, Measurement Systems Laboratory. Member AIAA.

† Equation (1) is more correctly written as a stochastic differential equation with the forcing term defined as Brownian motion. It is to be understood that $q(t)$ is the formal derivative of Brownian motion. Solutions of Eq. (1) are defined in terms of stochastic integrals.

(1), (2) by defining

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \quad F(t) = \begin{bmatrix} F_1(t) & 0 \\ 0 & F_2(t) \end{bmatrix} \tag{6, 7}$$

$$H(t) = [H_1(t)H_2(t)] \tag{8}$$

$$Q(t) = \begin{bmatrix} Q_1(t) & 0 \\ 0 & Q_2(t) \end{bmatrix} \tag{9}$$

Thus the separation of signals from correlated noise is a special case of estimating a state from linear combinations of its elements. Furthermore, since $x(t)$ contains both system and noise state vectors, the filter will generate estimates of both the signal and noise processes.

III. Discrete Estimation

As a first step in the development, equations for discrete estimation of $x(t)$ will be obtained. Consider a partition of time into increments Δt where

$$\Delta t = t_n - t_{n-1} \tag{10}$$

Solutions of (1) evaluated at points t_n may be written in recursive form as

$$x(t_n) = \Phi(t_n,t_{n-1})x(t_{n-1}) + v(t_n) \tag{11}$$

where Φ is the state transition matrix satisfying

$$(d/dt)\Phi(t,t_i) = F(t)\Phi(t,t_i), \quad \Phi(t_i,t_i) = I \tag{12}$$

and the $v(t_n)$ are Gaussian vector valued random variables with statistics

$$E[v(t_n)] = 0, \quad E[v(t_n)v^T(t_m)] = 0, \quad m \neq n$$

$$E[v(t_n)v^T(t_n)] = V(t_n) = \int_{t_{n-1}}^{t_n} \Phi(t_n,\tau)Q(\tau)\Phi^T(t_n,\tau)d\tau$$

Assuming measurement information is not continuous, but measurements are taken at discrete times t_n, the measurement equation is written as

$$m(t_n) = H(t_n)x(t_n) \tag{13}$$

Recursion formulas for optimally estimating $x(t_n)$ from the discrete measurements $m(t_n)$ were derived by Kalman.[7] Applying these equations to the system (11), (13) obtains

$$\hat{x}(t_n) = \hat{x}'(t_n) + P'(t_n)H^T(t_n)[H(t_n)P'(t_n)H^T(t_n)]^{-1} \times [m(t_n) - H(t_n)\hat{x}'(t_n)] \tag{14}$$

$$\hat{x}'(t_n) = \Phi(t_n,t_{n-1})\hat{x}(t_{n-1}) \tag{15}$$

$$P(t_n) = P'(t_n) - P'(t_n)H^T(t_n)[H(t_n)P'(t_n)H^T(t_n)]^{-1} \times H(t_n)P'(t_n) \tag{16}$$

$$P'(t_n) = \Phi(t_n,t_{n-1})P(t_{n-1})\Phi^T(t_n,t_{n-1}) + V(t_n) \tag{17}$$

$$\hat{x}(0) = E[x(0)], \quad P(0) = \mathrm{cov}[x(0)]$$

where $\hat{x}(t_n)$ is the minimum variance estimate of $x(t_n)$ and $P(t_n)$ is the covariance of errors in the estimate.

IV. Continuous Estimation

For a particular realization of the $x(t)$ process and resulting $m(t)$ process, a partition of time into j steps yields a particular estimate $\hat{x}_j(t)$. If an appropriate sequence of finer and finer partitions is constructed, implying more and more measurements, it can be shown‡ that the sequence of estimates $\hat{x}_j(t)$ converges to a limit $\hat{x}(t)$. The object, in what follows, is to derive equations for determining this limit.

The discrete estimation formulas (14–17) are applied to the task of determining $\hat{x}(t)$. Defining matrices $A(t_n)$ and $B(t_n)$ as

$$A(t_n) = P'(t_n)H^T(t_n)[H(t_n)P'(t_n)H^T(t_n)]^{-1} \tag{18}$$

$$B(t_n) = I - A(t_n)H(t_n) \tag{19}$$

It can be readily shown that

$$H(t_n)A(t_n) = I, \quad B(t_n)A(t_n) = 0, \quad B(t_n)^2 = B(t_n) \tag{20}$$

and (14) may be written as

$$\hat{x}(t_n) = B(t_n)\hat{x}'(t_n) + A(t_n)m(t_n) \tag{21}$$

Multiplying (21) by $B(t_n)$ and using (20) yields

$$B(t_n)\hat{x}(t_n) = B(t_n)\hat{x}'(t_n) \tag{22}$$

Now (22) must hold at each measurement and in particular at t_{n-1}

$$B(t_{n-1})\hat{x}'(t_{n-1}) = B(t_{n-1})\hat{x}(t_{n-1}) \tag{23}$$

Also, since the state transition matrix has a continuous derivative (12), Eq. (15) may be written as

$$\hat{x}'(t_n) = [I + F(\tau)\Phi(\tau,t_{n-1})\Delta t]\hat{x}(t_{n-1}), \quad t_{n-1} < \tau < t_n \tag{24}$$

Multiplying (24) through by $B(t_n)$ and subtracting (23) yields

$$B(t_n)\hat{x}'(t_n) - B(t_{n-1})\hat{x}'(t_{n-1}) = [B(t_n) - B(t_{n-1}) + B(t_n)F(\tau)\Phi(\tau,t_{n-1})\Delta t]\hat{x}(t_{n-1}) \tag{25}$$

It will be assumed, for the time being, that $B(t)$ exists and has a continuous derivative. Then by defining

$$y(t_n) = B(t_n)\hat{x}'(t_n) \tag{26}$$

substituting into (25), dividing through by Δt, and taking the limit as $\Delta t \to 0$ one obtains the differential equation

$$\dot{y}(t) = [\dot{B}(t) + B(t)F(t)]\hat{x}(t) \tag{27}$$

Further, substituting (26) into (21) yields

$$\hat{x}(t) = y(t) + A(t)m(t) \tag{28}$$

and from the definitions (1) and (2), $m(t)$ is continuous, with probability one so $\hat{x}(t)$ and $\dot{y}(t)$ will be continuous if it can be shown that $A(t)$ and $\dot{B}(t)$ are continuous.

Equations (27) and (28) determine the basic form of the correlated measurement noise filter. The remaining problem is to derive expressions for $A(t)$, $B(t)$, and $\dot{B}(t)$ and demonstrate continuity. From (16) it is clear that

$$H(t_{n-1})P(t_{n-1}) = 0 \tag{29}$$

and expressions for $P'(t_n)$ and $H(t_n)$ can be written as

$$P'(t_n) = P(t_{n-1}) + [F(t_{n-1})P(t_{n-1}) + P(t_{n-1})F^T(t_{n-1}) + Q(t_{n-1})]\Delta t + O(\Delta t) \tag{30}$$

$$H(t_n) = H(t_{n-1}) + \dot{H}(t_{n-1})\Delta t + O(\Delta t) \tag{31}$$

where $O(\Delta t)$ represents a matrix of higher-order terms

$$\lim_{\Delta t \to 0} \frac{O(\Delta t)}{\Delta t} = 0$$

Substituting (29, 30, and 31) into (16) and taking limits as Δt approaches zero yields a Riccati equation for $P(t)$

$$\dot{P} = FP + PF^T + Q - [P\dot{H}^T + (PF^T + Q)H^T] \times [HQH^T]^{-1}[\dot{H}P + H(FP + Q)] \tag{32}$$

where it is understood that all matrices are functions of time. Similarly, substitution into (18) obtains an expression for $A(t)$

$$A = [P\dot{H}^T + (PF^T + Q)H^T][HQH^T]^{-1} \tag{33}$$

For the time being it is also assumed that HQH^T is nonsingu-

‡ Proof of convergence depends upon the fact that the sequence $\hat{x}_j(t)$ is a martingale and

$$\lim_{j\to\infty} E[x(t)|m(t_1),m(t_2),\ldots m(t_j)] = E[x(t)|m(\tau)] = \hat{x}(t)$$

$$0 \leq t_1 < t_2 < \ldots < t_j \leq t, \quad 0 \leq \tau \leq t$$

See Wonham[8] or Doob.[5]

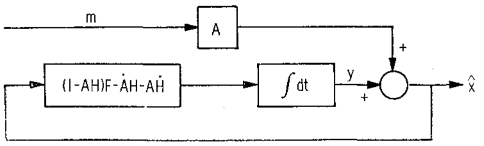

Fig. 1 Optimum filter.

lar. The problem of singular HQH^T matrices will be treated in Sec. V. From (32) and the definitions of F, Q, and H, the covariance matrix has a continuous derivative and hence A has a continuous derivative. Thus the assumptions made in writing (27) and (28) are validated.

The entire set of filter equations can now be written out as follows:

$$\hat{x} = y + Am \tag{34}$$

$$\dot{y} = [(I - AH)F - \dot{A}H - A\dot{H}]\hat{x} \tag{35}$$

$$A = [P\dot{H}^T + (PF^T + Q)H^T][HQH^T]^{-1} \tag{36}$$

$$\dot{P} = FP + PF^T + Q - A[\dot{H}P + H(FP + Q)] \tag{37}$$

Equations (34) and (35) determine the basic filter configuration as shown in Fig. 1. The weighting matrix A is determined from (36), which requires P from the solution of the Riccati equation (37).

Figure 1 also demonstrates that the correlated measurement noise filter has a form significantly different from the Kalman-Bucy filter. In particular, the measurement process appears directly in the estimate, after transformation by the matrix A, whereas in the white noise case, only integrals of the measurements appear in the estimate. This difference in form results in singular filter equations when the white noise filter is applied to correlated measurement noise problems.

V. Singular HQH^T Matrices

The matrix HQH^T must be nonsingular if the methods of Sec. IV are to be applied. Physically this matrix represents the strength of first integrals of white noise in the measurements. Singularity of HQH^T implies that there exists a linear transformation of the measurement vector such that one or more components of the transformed vector contain no first integrals of white noise.

Singularities can occur for two reasons. First, if one (or more) of the rows of H is a linear combination of other rows, then HQH^T is singular. This difficulty is readily eliminated by ignoring the measurement associated with the linearly dependent row, since it is redundant and provides no additional information to the filter. Hence measurements that are linear combinations of other measurements are eliminated until the H matrix contains only linearly independent rows.

The second cause of singularity occurs when one of the measurements is separated from white noise by two or more integrations. The simple example system shown in Fig. 2 will serve to illustrate the difficulty. An estimator for x_1 is desired. Inputs q_1 and q_2 are independent white noises with strengths Q_1 and Q_2. If the state vector is defined as

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

then the H and Q matrices are

$$H = [0 \quad 0 \quad 1], \quad Q = \begin{bmatrix} Q_1 & 0 & 0 \\ 0 & Q_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and $HQH^T = 0$ is a singular matrix of dimension one. Reference to Fig. 2 shows that the measurement x_3 is separated from white noise by two integrals. However, x_3 is the integral of x_2, so x_2 can be constructed perfectly (at least in theory) by differentiating x_3. If this is done, the filter has x_2 available to estimate x_1. The state vector and H and Q matrices then become

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad H = [0 \quad 1], \quad Q = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix}$$

so $HQH^T = Q_2$, which is positive. Thus the filter must be preceded by a differentiator to obtain x_2 as the filter input. In the general case one or more differentiations may be required to eliminate the singularity problem.

The procedures outlined previously do not compromise the optimality of the filter. Eliminating linearly dependent measurements or differentiating measurements does not alter the information content of the measurement data since the original measurements can be reconstructed perfectly.

VI. Reduction of the Filter Dimension

From the definitions of Sec. II the measurement vector is a linear combination of elements of the state

$$m(t) = H(t)x(t)$$

Hence each element of $m(t)$ represents a perfect measurement of the projection of the state in the direction of the corresponding row of $H(t)$. If a coordinate system for the state space is chosen so that some of the basis vectors lie in the directions of the rows of $H(t)$, then the measurements $m(t)$ will eliminate estimation errors in those vector directions. Hence certain of the state variables will always be estimated perfectly and the corresponding rows and columns of the P matrix will be identically zero. The resulting filter calculations are then simplified and the dimension of the filter state is reduced by the dimension of the measurement vector. This technique is illustrated in the example problem of Sec. VIII.

VII. Initial Conditions

In order to completely specify the optimal filter, initial conditions must be obtained for $y(t)$ and $P(t)$. If filtering is begun at time zero$^+$, then, before any measurement data are available,

$$\hat{x}(0) = E[x(0)]$$

$$P(0) = E\{[x(0) - \hat{x}(0)][x(0) - \hat{x}(0)]^T\}$$

From Eq. (14), just after measurement data is available, the estimate is

$$\hat{x}(0^+) = \hat{x}(0) + P(0)H^T(0)[H(0)P(0)H^T(0)]^{-1}[m(0) - H(0)\hat{x}(0)]$$

Similarly the error covariance becomes

$$P(0^+) = P(0) - P(0)H^T(0)[H(0)P(0)H^T(0)]^{-1}H(0)P(0) \tag{38}$$

From (34) the initial condition on $y(t)$ is

$$y(0^+) = \hat{x}(0^+) - A(0)m(0) \tag{39}$$

where

$$A(0) = \{P(0^+)\dot{H}^T(0) + [P(0^+)F^T(0) + Q(0)]H^T(0)\} \times [H(0)Q(0)H^T(0)]^{-1} \tag{40}$$

Thus the initial condition on the filter is dependent upon the initial measurement and cannot be determined a priori.

VIII. Example Problem

As a means of demonstrating the application of the theory, a problem of separating exponentially correlated signal and

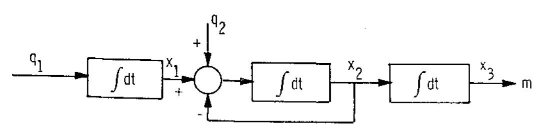

Fig. 2 Triple integral plant.

noise processes is solved. The system is as shown in Fig. 3. A signal process $s(t)$ is to be estimated from the measurement process $m(t)$ which is corrupted by correlated measurement noise $n(t)$. Feedback gains α and β are constant. The total system (signal and noise) is second order and $s(t)$ and $n(t)$ could be chosen as state variables. For the reasons given in Sec. VI, however, it is advantageous to choose $s(t)$ and $m(t)$ as state variables. They span the state space because $n(t)$ can be generated by the linear combination

$$n(t) = m(t) - s(t) \tag{41}$$

The state vector is defined as

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} s(t) \\ m(t) \end{bmatrix} \tag{42}$$

and the two state variables satisfy differential equations

$$\dot{x}_1(t) = -\alpha x_1(t) + q_1(t) \tag{43}$$

$$\dot{x}_2(t) = (\beta - \alpha)x_1(t) - \beta x_2(t) + q_1(t) + q_2(t) \tag{44}$$

Thus if the white noise statistics are stationary so

$$E[q_1(t)] = 0, \quad E[q_1(t)q_1(t+\tau)] = a\delta(\tau)$$
$$E[q_2(t)] = 0, \quad E[q_2(t)q_2(t+\tau)] = b\delta(\tau)$$
$$E[q_1(t)q_2(t+\tau)] = 0$$

then the F, Q, H, and $\dot{H}$ matrices become

$$F = \begin{bmatrix} -\alpha & 0 \\ (\beta - \alpha) & -\beta \end{bmatrix}, \quad Q = \begin{bmatrix} a & a \\ a & a+b \end{bmatrix} \tag{45}$$
$$H = [0 \quad 1], \quad \dot{H} = [0 \quad 0]$$

Now $x_2(t)$ is, by definition, the measurement available to the estimator. If the four elements of the error covariance matrix are defined as

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \tag{46}$$

then, since $x_2(t)$ is known without error,

$$p_{12}(t) = p_{21}(t) = p_{22}(t) = 0 \tag{47}$$

and only p_{11} need be calculated. Applying (45) and (47) to (37) obtains the differential equation

$$\dot{p}_{11} = a - 2\alpha p_{11} - [(\beta - \alpha)p_{11} + a]^2/(a + b) \tag{48}$$

Also, the gain matrix A is calculated from (36) as

$$A = \begin{bmatrix} [(\beta - \alpha)p_{11} + a]/(a+b) \\ 1 \end{bmatrix} \tag{49}$$

and the filter feedback matrix is

$$(I - AH)F - \dot{A}H - A\dot{H} =$$
$$\left[\begin{array}{c|c} \dfrac{(\alpha - \beta)[(\beta - \alpha)p_{11} + a]}{(a+b)} - \alpha & \dfrac{(\beta - \alpha)(\beta p_{11} - \dot{p}_{11}) + a\beta}{(a+b)} \\ \hline 0 & 0 \end{array}\right] \tag{50}$$

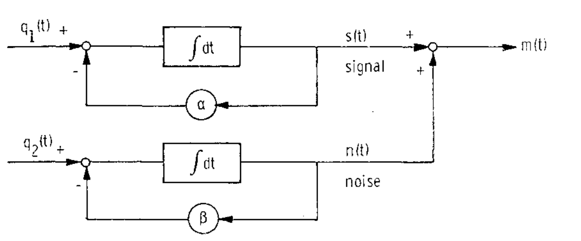

Fig. 3 Example system.

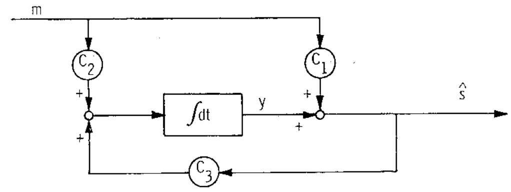

Fig. 4 Optimal filter.

Applying these expressions to the filter block diagram (Fig. 1) and eliminating extraneous signal paths yields the optimal system shown in Fig. 4 with optimal gains C_1, C_2, and C_3 given as

$$C_1 = [(\beta - \alpha)p_{11} + a]/(a + b)$$
$$C_2 = [(\beta - \alpha)(\beta p_{11} - \dot{p}_{11}) + a\beta]/(a + b)$$
$$C_3 = \{(\alpha - \beta)[(\beta - \alpha)p_{11} + a]/(a + b)\} - \alpha$$

The solution of (48) can be obtained by standard methods,[4] obtaining

$$p_{11}(t) = \frac{[(a+b)\omega \cosh\omega t - (\alpha b + \beta a)\sinh\omega t]p_{11}(0^+) + ab\sinh\omega t}{[(\beta-\alpha)^2\sinh\omega t]p_{11}(0^+) + (a+b)\omega\cosh\omega t + (\alpha b + \beta a)\sinh\omega t}$$

where

$$\omega = [\alpha^2 b^2 + ab(\alpha^2 + \beta^2) + \beta^2 a^2]/(a + b)$$

Finally, if the system Fig. 3 is in stationary operation when filtering begins, the initial statistics are

$$E[s(0)] = E[n(0)] = 0$$
$$E[s(0)^2] = a/2\alpha, \quad E[n(0)^2] = b/2\beta, \quad E[s(0)n(0)] = 0$$

and the initial conditions on p_{11} and y are calculated from (38–40) as

$$p_{11}(0^+) = ab/2(a\beta + \alpha b)$$
$$y(0^+) = [ab(\beta - \alpha)/2(a + b)(a\beta + b\alpha)]m(0)$$

IX. Conclusions

A derivation has been presented for the optimum correlated measurement noise filter. Methods were also described for handling singular problems and reducing filter complexity. Appropriate initial conditions were derived and an example problem served to demonstrate application of the results to a filtering problem.

References

[1] Bryson, A. E., Jr. and Henrikson, L. J., "Estimation Using Sampled Data Containing Sequentially Correlated Noise," *Journal of Spacecraft and Rockets*, Vol. 5, No. 6, June 1968, pp. 662–665.

[2] Bryson, A. E. and Johansen, D. E., "Linear Filtering for Time-Varying Systems Using Measurements Containing Colored Noise," *IEEE Transactions on Automatic Control*, Vol. AC-10, No. 1, Jan. 1965.

[3] Bucy, R. S., "Optimal Filtering for Correlated Noise," Rept. 5107-Pr, Sept. 1966, RAND Corp.

[4] Deyst, J. J., "Optimum Continuous Estimation of Nonstationary Random Variables," Rept. T-369, Jan. 1964, M.I.T. Instrumentation Lab.

[5] Doob, J. L., *Stochastic Processes*, Wiley, New York, 1953, pp. 319, 331.

[6] Kalman, R. E. and Bucy, R. S., "New Results in Linear Filtering and Prediction Theory," *Journal of Basic Engineering*, March 1961.

[7] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, March 1960.

[8] Wonham, W. M., lecture notes from a course on stochastic control, Spring 1966, Massachusetts Institute of Technology.