Benchmark Tests of Functionality for Control Related Software Products

IEEE Boston Section of the Control System Society *

July 19, 1993

Contents

* A particular software product may not be able to handle each one of these tests. Please complete every test that applies and submit results in hard copy form (75 copies) for distribution at the September 1993 IEEE Meeting. For ease in cross-comparisons, please maintain the same order of test cases in reporting the results. If there are any questions about these tests, please contact Dr. Thomas Kerr at **TeK Associates**, 11 Paul Revere Rd., Lexington, MA 02173-6632 or call (617) 862-8680.

1 Notation and Motivation

For a time-invariant linear continuous-time state-variable representation in terms of the matrix triple (A, B, C) as

$$\frac{dx}{dt} = A\,x(t) + B\,u(t) \tag{1}$$

$$y(t) = C\,x(t)\,, \tag{2}$$

with corresponding system transfer function matrix

$$\mathcal{H}(s) = C\,(sI - A)^{-1}\,B\,, \tag{3}$$

the equivalent discrete-time reformulation in terms of the matrix triple (L, M, N) is

$$x(k+1) = L\,x(k) + M\,u(k) \tag{4}$$

$$y(k) = N\,x(k)\,, \tag{5}$$

with corresponding system transfer function matrix

$$\mathcal{H}(z) = N\,(zI - L)^{-1}\,M\,, \tag{6}$$

where s in Eq. 3 is the Laplace transform variable and z in Eq. 6 is the Z-transform variable and use of $x(\cdot)$ in Eqs. 1, 2, 4, and 5 is just notation for the system state variables with $u(\cdot)$ as the input and $y(\cdot)$ as the measured output.

An initial investigation of posing the continuous-time problem as an *exact* formulation in discrete-time proceeds as follows. The form of the solution to the differential equation of Eq. 1 is:

$$x(t) = e^{A(t-s)}\, x(s) + \int_s^t e^{A(t-\tau)}\, B\, u(\tau)\, d\tau\ . \tag{7}$$

In particular, for the upper and lower limits of the above integral being

$$t = (k+1)\,\Delta\ , \tag{8}$$

$$s = k\,\Delta\ , \tag{9}$$

with

$$\Delta = \text{constant incremental step} - \text{size}\ , \tag{10}$$

the solution of Eq. 7 corresponds to the following recursive iteration in discrete-time:

$$x(k+1) = \left[e^{A\,\Delta}\right]\, x(k) + \int_{k\,\Delta}^{(k+1)\,\Delta} e^{A(k\,\Delta-\tau)}\, B\, u(\tau)\, d\tau\ , \tag{11}$$

which, under the further assumption that $u(\tau)$ is essentially constant [1] over the time-step from any $k\Delta$ to any other $(k+1)\Delta$, yields:

$$x(k+1) = \left[e^{A\,\Delta}\right]\, x(k) + \left[\int_{k\,\Delta}^{(k+1)\,\Delta} e^{A\,((k+1)\,\Delta-\tau)}\, B\, d\tau\right]\, u(k)\ . \tag{12}$$

Upon making the change of variable

$$\tau = \tau' + k\,\Delta \tag{13}$$

and substituting into the integral in Eq. 12, yields:

$$\begin{aligned} x(k+1) &= \left[e^{A\,\Delta}\right]\, x(k) + \left[\int_0^{\Delta} e^{A\,\Delta}\, e^{-A\,\tau'}\, d\tau'\right]\, B\, u(k) \\ &= \left[e^{A\,\Delta}\right]\, x(k) + \left[e^{A\,\Delta}\right]\left[\int_0^{\Delta} e^{-A\,\tau'}\, d\tau'\right]\, B\, u(k)\ . \end{aligned} \tag{14}$$

The expression in Eq. 14 is the most general form of the discrete-time formulation that corresponds *exactly* to the continuous-time formulation of Eq. 7 except for the minor error incurred in assuming $u(\tau)$ to be essentially constant over each small step-size Δ, thus allowing it to be taken outside of the integral. In the case where the continuous-time input $u(t)$ is independent, white, Gaussian process noise of continuous-time

[1] This assumption is sometimes enforced through use of a zero-order hold on the input $u(t)$.

covariance intensity level, Q, to have *exact* adherence without any approximation incurred, the discrete-time formulation should be

$$x(k+1) = \left[e^{A\,\Delta}\right]\; x(k) + u'(k)\;, \tag{15}$$

where

$$u'(k) = \text{zero} - \text{mean Gaussian white noise}\;, \tag{16}$$

having discrete-time covariance intensity level [15, p. 270]:

$$Q_d = E\left[u'(k)(u'(j))^T\right] = e^{A\,\Delta}\left[\int_0^{\Delta} e^{-A\,\tau}\; B\; Q\; B^T\; e^{-A^T\,\tau}\, d\tau\right] e^{A^T\,\Delta}\;\; \delta_{kj}\;, \tag{17}$$

where the above Kronecker delta is defined as

$$\delta_{kj} = \begin{cases} 1 & \text{if k=j} \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

The above Q_d in Eq. 17 is the appropriate discrete-time process noise covariance level to use to have *exact* agreement between the discrete-time mechanization of Eq. 15 and the continuous-time formulation of Eqs. 1 or 7. A well-known approximation for Q_d' (due to Kalman) which is sometimes used is to take

$$Q_d' = \Delta\, Q\;; \tag{19}$$

however, the deleterious effect of invoking this approximation is uncalibrated and it can easily be seen to be an obviously unsatisfactory representation of off-diagonal terms. The effects due to incorrect off-diagonal cross-correlation terms can be significant for many applications.

2 Transition Matrix Calculation

For linear systems of the following form:

$$\dot{x}(t) = A_i(t)\, x(t)\;, \tag{20}$$

calculate the appropriate transition matrix $\Phi_i(t)$ (when applicable, can use matrix exponential or Pade approximate approach if you so desire for greater accuracy):

1. **Case 1:** Find the appropriate transition matrix for the following time-invariant system:

$$A_1 \;=\; \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ & \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix}, \tag{21}$$

   for time step $\Delta = 0.405$.

2. **Case 2:** Find the appropriate transition matrix for the following time-invariant system:

$$A_2 = \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{bmatrix}, \tag{22}$$

for time step $\Delta = 0.405$.

3. **Case 3:** Find the appropriate transition matrix $\Phi(t,s)$ [2] for the following time-varying linear system:

$$A_3(t) = \begin{bmatrix} -1 + a(\cos t)^2 & 1 - a \sin t \cos t \\ -1 - a \sin t \cos t & -1 + a(\sin t)^2 \end{bmatrix}, \tag{23}$$

with the parameter $a = 1.5$. Calculate what the corresponding transition matrix is at time=10 (units).

2.1 ANSWERS:

Certain matrices known as "idempotent" matrices have the unusual property that when multiplied times itself again yields itself as the result:

$$A\,A = A. \tag{24}$$

The non-trivial system matrices of Test **Cases 1** and **2** exhibit this property. The present application in software verification is a neat application of idempotent matrices being used to construct test matrices for verifying the transition matrix algorithmic implementations that are used for computer computation of e^{Ft}. The utility of these test matrices is that the resulting analytically derived expression for e^{Ft} is conveniently in closed-form for $F = A$. Hence the output performance of a general e^{Ft} subroutine implementation can ultimately be gauged by how close it comes to achieving the known ideal exact solution.

Using the representation of a matrix exponential, defined in terms of its Taylor series, but evaluated with an idempotent matrix A having the property of Eq. 24 being substituted along with time-step Δ; the matrix Taylor series expansion of $e^{A\Delta}$

[2] Satisfying the equation $\frac{\partial}{\partial t}\Phi(t,s) = A(t)\Phi(t,s)$ with boundary condition $\Phi(s,s) = I_{n \times n}$ (an identity matrix).

now yields

$$
\begin{aligned}
e^{A\Delta} &= \textstyle\sum_{k=0}^{\infty} \frac{A^k}{k!}\Delta^k \\
&= I + \tfrac{A}{1!}\Delta + \tfrac{A^2}{2!}\Delta^2 + \tfrac{A^3}{3!}\Delta^3 + \cdots \\
&= I + A(\tfrac{\Delta}{1!} + \tfrac{\Delta^2}{2!} + \tfrac{\Delta^3}{3!} + \cdots) \\
&= I + A(1 + \tfrac{\Delta}{1!} + \tfrac{\Delta^2}{2!} + \tfrac{\Delta^3}{3!} + \cdots - 1) \\
&= I + A(e^{\Delta} - 1)\ ,
\end{aligned} \tag{25}
$$

as explained in [1, Sec. IV]. Thus, the closed-form exact expression for the transition matrix corresponding to idempotent system matrices is as depicted in the last line of Eq. 25 as a finite two step operation involving just a scalar multiplication of a matrix and a single matrix addition (as compared to an infinite series that must be truncated in the case of standard software implementations for the case of more general matrices).

For clarity, motivation is now offered for how these idempotent matrices were obtained. Consider the problem of seeking to solve the following algebraic equation for $x(n \times 1)$, given $y(m \times 1)$ and $C(m \times n)$:

$$\underline{y} = C\ \underline{x}. \tag{26}$$

Assuming that the rank of C is the same as the rank of the *augmented matrix* $[C|\underline{y}]$, it is reasonably well-known (see [2, Appendix A, Section A.1] and [11, p. 417]) that a solution to Eq. 26 is of the form

$$\underline{x} = C^{\dagger}\underline{y} + (I_n - C^{\dagger}C)\underline{w} \tag{27}$$

for arbitrary $\underline{w}$ and that the term within parenthesis in Eq. 27 is idempotent (where $C^{\dagger}$ in Eq. 27 is the Moore-Penrose pseudoinverse [13]). In forming two counterexamples in [2], [7], the following two matrices and their respective pseudoinverses were obtained (as derived in [2]):

$$C_1 = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \quad ; \quad C_1^{\dagger} = \frac{1}{25}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \tag{28}$$

and

$$C_2 = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \quad ; \quad C_2^{\dagger} = \begin{bmatrix} -\frac{1}{3} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}. \tag{29}$$

Therefore via Eq. 27, the following two matrices are idempotent

$$\begin{aligned} A_1 &\triangleq (I - C_1^\dagger C_1) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{25}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \\ &= \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix} \end{aligned} \tag{30}$$

and

$$\begin{aligned} A_2 &\triangleq (I - C_2^\dagger C_2) \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} -\frac{1}{3} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{bmatrix}, \end{aligned} \tag{31}$$

both of which check as being idempotent by satisfying Eq. 24 as an identity. In considering the step-size to use in the evaluation of the final line of Eq. 25, convenience in using just a scalar multiplying factor of one half times the matrix in Eq. 25 would dictate using

$$\Delta = 0.405 \tag{32}$$

since from Burlington's mathematical tables [12]

$$(\epsilon^{\Delta} - 1) = (\epsilon^{0.405} - 1) = (1.50 - 1) = 0.50. \tag{33}$$

Therefore, the two evaluations corresponding to invoking Eq. 25 are:

1. **Case 1 Answer:**

$$\begin{aligned} e^{A_1\Delta} &= I + \tfrac{1}{2}A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \tfrac{1}{2}\begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix} \\ &= \begin{bmatrix} 1.40 & -0.20 \\ -0.20 & 1.10 \end{bmatrix} \end{aligned} \tag{34}$$

2. **Case 2 Answer:**

$$e^{A_2\Delta} = I + \frac{1}{2}A_2 = \begin{bmatrix} 1.16\overline{6} & -0.16\overline{6} & 0.16\overline{6} \\ -0.16\overline{6} & 1.16\overline{6} & -0.16\overline{6} \\ 0.16\overline{6} & -0.16\overline{6} & 1.16\overline{6} \end{bmatrix}. \tag{35}$$

3. **Case 3 Answer:** As can be analytically confirmed to satisfy the canonical equation repeated in the footnote on page 4, the closed-form expression for the corresponding transition matrix is:

$$\begin{aligned}\Phi(t,0) &= \begin{bmatrix} e^{(a-1)t}\cos t & e^{-t}\sin t \\ -e^{(a-1)t}\sin t & e^{-t}\cos t \end{bmatrix} \\ &= \begin{bmatrix} e^{(1.5-1)t}\cos t & e^{-t}\sin t \\ -e^{(1.5-1)t}\sin t & e^{-t}\cos t \end{bmatrix},\end{aligned} \tag{36}$$

so for $t = 10$, the requested result is

$$\begin{aligned}\Phi(10,0) &= \begin{bmatrix} e^{5}\cos(10) & e^{-10}\sin(10) \\ -e^{5}\sin(10) & e^{-10}\cos(10) \end{bmatrix} \\ &= \begin{bmatrix} -124.5293 & -2.4699\times 10^{-5} \\ 80.7399 & -3.8094\times 10^{-5} \end{bmatrix}.\end{aligned} \tag{37}$$

The results of Eq. 34 and 35 are the now *known* closed-form solutions to an $e^{A\Delta}$ evaluation of the matrices of Eqs. 21 and 22, respectively, with $\Delta = 0.405$ and Eq. 37 offers the corresponding transition matrix spanning a time interval from $t = 0$ to $t = 10$ units.

The transition matrix calculation for numerically converting the continuous-time n-state model description to discrete-time, has historically adaptively tailored the number of terms retained in the defining Taylor series (first two lines of Eq. 25) by using either too coarse a norm (see [16]) or an invalid norm [1, pp. 938-939]. A tighter bound for this purpose has been derived from considerations of both column-sum and row-sum norms in [23] and, additionally, it is prudent to also set an upper limit on the total number of terms from the Taylor series expansion allowed to be used in calculating the transition matrix so that the computation can't run away (otherwise it could incur numerous OVERFLOW's due to the effect of accumulated roundoff).

3 Exact Calculation of Discrete-Time Process (Plant) Noise Covariance Intensity Matrix

For time-invariant linear systems, the discrete-time equivalent of continuous-time white noise is well known to satisfy an equation of the following form for the associated covariance intensity matrix:

$$Q_{di} = e^{A_i \Delta_i} \left[\int_0^{\Delta_i} e^{-A_i \tau} B_i Q_i B_i^T e^{-A_i^T \tau} d\tau \right] e^{A_i^T \Delta_i} , \tag{38}$$

where $\Delta_1 = \Delta_2 = 0.405$ (units). Please evaluate the above matrix expression for Q_{di} (important for accurate Monte-Carlo simulations) for the same parameters as in **Cases 1** and **2** in Section 2 above but, respectively, with $B_1 = I_{2x2}$ and $Q_1 = I_{2x2}$ for **Case 1** and with $B_2 = I_{3x3}$ and $Q_2 = I_{3x3}$ for **Case 2** (so that appropriate dimensions will be conformable under the indicated standard matrix multiplications).

3.1 ANSWERS:

Using the result of Eq. 25 for idempotent matrices within the more general expression of Eqs. 17 or 38, allows this expression for the required discrete-time process noise covariance to be evaluated analytically in closed-form as:

$$
\begin{aligned}
\mathrm{Q_d} &= \left[I + A(e^{\Delta} - 1)\right] \int_0^{\Delta} \left[I + A(e^{-\tau} - 1)\right] BQB^T \left[I + A^T(e^{-\tau} - 1)\right] d\tau \left[I + A^T(e^{\Delta} - 1)\right] \\
&= \left[I + A(e^{\Delta} - 1)\right] \\
&\quad \int_0^{\Delta} \left[BQB^T + (ABQB^T + BQB^TA^T)(e^{-\tau} - 1) + ABQB^TA^T(e^{-2\tau} - 2e^{-\tau} + 1)\right] d\tau \\
&\quad \left[I + A^T(e^{\Delta} - 1)\right] \\
&= \left[I + A(e^{\Delta} - 1)\right] \\
&\quad \left[BQB^T\Delta + (ABQB^T + BQB^TA^T)(1 - e^{-\Delta} - \Delta) + ABQB^TA^T(-\frac{3}{2} - \frac{1}{2}e^{-2\Delta} + 2e^{-\Delta} + \Delta)\right] \\
&\quad \left[I + A^T(e^{\Delta} - 1)\right] .
\end{aligned} \tag{39}
$$

This is a new result [3] that is also useful as a confirming check for software implementations of Eq. 17.

1. **Case 1 Answer:**

$$Q_{d1} = \begin{bmatrix} 0.5793 & -0.0755 \\ -0.0755 & 0.5793 \end{bmatrix} , \tag{40}$$

[3] Along a different line, something similar to Eq. 39 can be computed for numerically evaluating Q_d for *any* constant matrix A, not just for idempotent matrices, by (1) expanding $e^{-A\tau}$ into its matrix Taylor series, (2) by performing the indicated multiplications of the two series within the integrand, (3) by subsequently performing term-by-term integration, and then (4) by retaining enough terms of the final series to be used to provide sufficient accuracy in actual numerical calculations.

2. **Case 2 Answer:**

$$Q_{d2} = \begin{bmatrix} 0.4780 & -0.0730 & 0.0730 \\ -0.0730 & 0.4780 & -0.0730 \\ 0.0730 & -0.0730 & 0.4780 \end{bmatrix} . \tag{41}$$

4 Differential Equation: Time-Varying Parameters

For the linear system of the form

$$\dot{x}(t) = \begin{bmatrix} -1 + a(\cos t)^2 & 1 - a \sin t \cos t \\ -1 - a \sin t \cos t & -1 + a(\sin t)^2 \end{bmatrix} x(t) , \tag{42}$$

with the parameter $a = 1.5$ and initial condition $x(0) = [5, 15]^T$, calculate the solution $x(10)$ at time=10 (units).

4.1 ANSWER:

From the transition matrix obtained in **Case 3** of Section 2, the answer here is:

$$\begin{aligned} x(10) = \Phi(10,0)x(0) &= \begin{bmatrix} e^5 \cos(10) & e^{-10} \sin(10) \\ -e^5 \sin(10) & e^{-10} \cos(10) \end{bmatrix} \begin{bmatrix} 5 \\ 15 \end{bmatrix} \\ &= \begin{bmatrix} -124.5293 & -2.4699 \times 10^{-5} \\ 80.7399 & -3.8094 \times 10^{-5} \end{bmatrix} \begin{bmatrix} 5 \\ 15 \end{bmatrix} \\ &= \begin{bmatrix} -622.6469 \\ 403.6989 \end{bmatrix} . \end{aligned} \tag{43}$$

ADDITIONAL INSIGHTS FROM THIS TIME-VARYING SYSTEM:

In analogy to what is done for time-invariant linear systems, the forming of a so-called *characteristic equation* corresponding to the time-varying system matrix of

Eq. 42 can be attempted and would be of the form:

$$
\begin{aligned}
0 &= \det|\lambda I - A_3(t)| = \det \begin{bmatrix} (\lambda+1) - a(\cos t)^2 & -1 + a\sin t\cos t \\ 1 + a\sin t\cos t & (\lambda+1) - a(\sin t)^2 \end{bmatrix} \\
&= (\lambda+1)^2 - a[(\cos t)^2 + (\sin t)^2](\lambda+1) \\
&\quad + a^2(\cos t)^2(\sin t)^2 - [-1 + a^2(\cos t)^2(\sin t)^2] \\
&= (\lambda+1)^2 - a(\lambda+1) + 1 \\
&= \lambda^2 + (2-a)\lambda + (2-a)\ ,
\end{aligned} \tag{44}
$$

which has so-called *eigenvalues* that are constant as the solution to the above quadratic equation as

$$\lambda = \frac{a - 2 \pm \sqrt{a^2-4}}{2}\ , \tag{45}$$

so these constant eigenvalues are complex numbers having negative real parts for

$$-2 < a < 2\ , \tag{46}$$

but the solution of Eq. 36 is observed to be an unstable unbounded growing exponential for

$$1 < a < 2\ , \tag{47}$$

and in particular for $a = 1.5$. This easily verifiable behavior contradicts some (widely propagated) notions by many control theorists of an earlier era that time-varying *eigenvalues* (sic), if all confined to the left half plane and by not moving around *too much*, correspond to a stable system. This example has eigenvalues that are constant in the left half plane (so can't move around too much or at all even) yet is blatantly unstable thus exposing the earlier notion as a folk theorem without substantiation.

5 Kalman Filter and Predecessor Monte-Carlo Simulator: Stationary Time-Invariant Systems

5.1 Overview of the structure of Kalman filtering

A Kalman filter (see Fig. 1) is an efficient and convenient computational scheme for providing the optimal estimate of the system state and an associated measure of the goodness of that estimate (the variance or covariance). In order to implement a KF, the actual continuous-time system must be adequately characterized by a linear (or linearized) ordinary differential equation model represented in state space at time t in terms of a vector $x(t)$, and having associated initial conditions specified, and availing sensor output measurements $z(t)$ (functions of the state plus additive measurement noise). It is mandatory that the KF itself actually contain within it an analytical mathematical model of the system and sensors in order to perform its computations (designated as a model-based estimator), and it must possess a statistical characterization of the covariance intensity level of the additive white Gaussian measurement and process noises present as well to enable an implementation.

Before the KF code can be validated as performing properly, or in case of known errors, before the source can be pin-pointed, first the inputs to the KF must be validated as being exactly what was intended. To this end, we first turn our attention to validating the Monte-Carlo simulator. A state-variable based Monte-Carlo simulator, of the form depicted in Fig. 2, should be developed/supplied to support testing the performance and adequacy of KF trackers.

5.2 Four test cases for validating simulator and filter

The overall structure of the simulator is depicted in Fig. 2. Using the input parameters of Test Case 1, as depicted in Table 1, the intermediate outputs provided by the software implementation should be verified to be correct. The specific features of the software implementation that can be confirmed using Test Case 1 are detailed in the second column from the left in Table 2.

Using the parameters of Test Case 2, as depicted in Table 1. The specific features of a software implementation that can be confirmed using Test Case 2 are detailed in the third column from the left in Table 2. Test Case 2 has an easy to determine closed-form expression for the transition matrix, for Q_d, for the steady-state Lyapunov equation, and for the ideal output power spectrum.

The specific features of the software implementation that can be confirmed using Test Case 3 are detailed in the fourth column from the left in Table 2. Actual extremely regular essentially *deterministic* sample functions obtained for the underlying known unstable system can conveniently be used to check at a high level that the output is exactly correct. Besides confirming the outputs of the simulator with an easily recognizable expected answer (as contrasted to Test Cases 1, 2, and 4, which provide random noise corrupted sample functions that can be confirmed at the aggregate level only from statistical properties that are a byproduct of downstream KF tracking or spectral estimation), this Test Case 3 also allows a programmer to calibrate (and correct) their plot routines and his scale conversion for output plots, if necessary.

Using the parameters of Test Case 4, as depicted in Table 1, the intermediate outputs provided by a software implementation can be verified to be correct. The specific features of a software implementation that can be confirmed using Test Case 4 are detailed in the fifth column from the left in Table 2. The main purpose of this last test case is to be able to handle the situation of providing prescribed multi-input/multi-

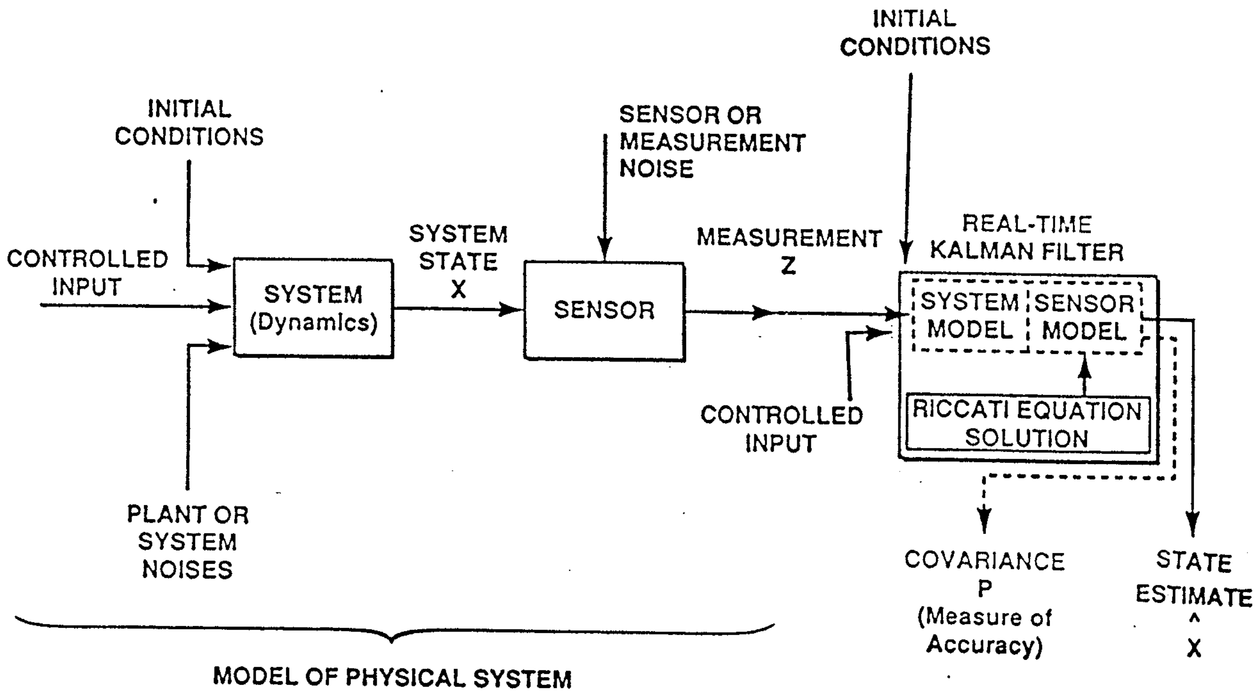


Figure 1: Overview Functional Block Diagram of the Internal Structure of a Kalman Filter

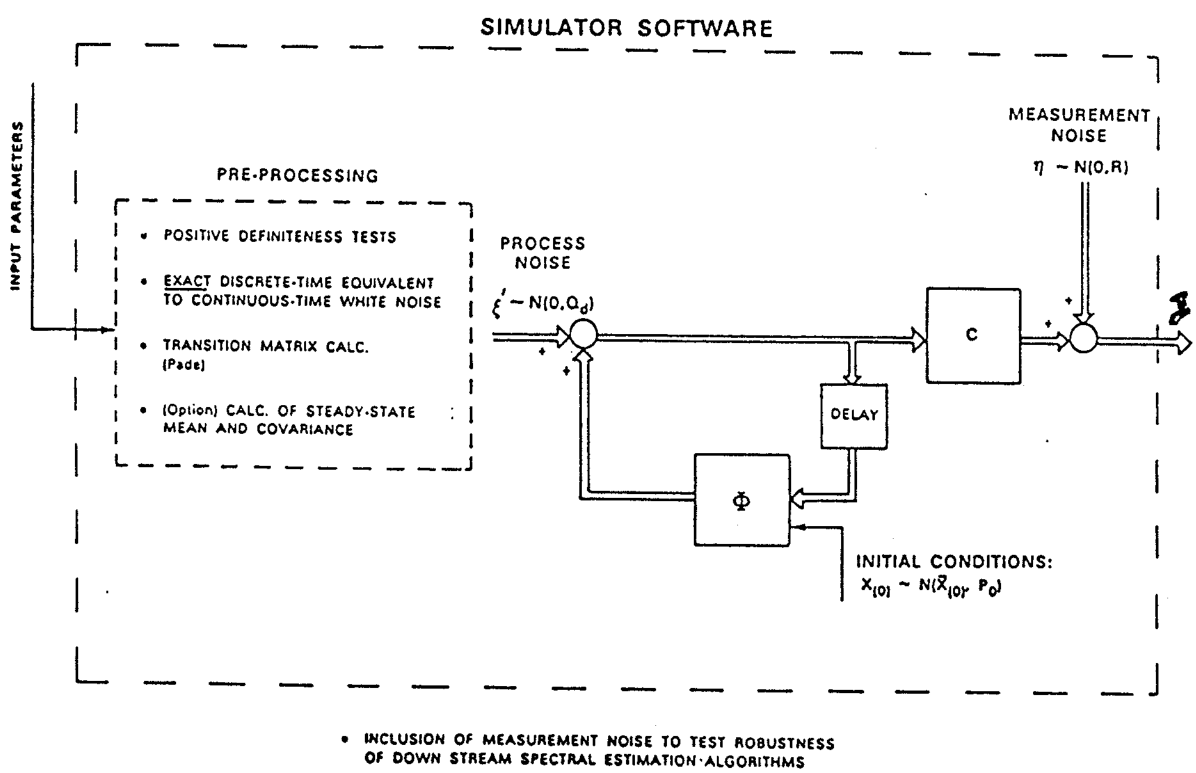


Figure 2: State-Variable Markov-Based Monte-Carlo Simulator

Table 1: Summary of Parameters of Test Case Models Used in Validation Tests of Primary Software Modules

STANDARDIZATION OF MODEL CASES FOR SIMULATION TESTING

Case No.	Test Case 1	Test Case 2	Test Case 3	Test Case 4
Step Size DEL (Δ)	0.405	0.5	0.5	1
System Matrix A	$\begin{bmatrix} 1/3 & -1/3 & 1/3 \\ -1/3 & 1/3 & -1/3 \\ 1/3 & -1/3 & 1/3 \end{bmatrix}$	$\begin{bmatrix} -5 & -1 \\ 6 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$	—
Transition Matrix $e^{F\Delta}$	$\begin{bmatrix} 1.1\overline{66} & -0.1\overline{66} & 0.1\overline{66} \\ -0.1\overline{66} & 1.1\overline{66} & -0.1\overline{66} \\ 0.1\overline{66} & -0.1\overline{66} & 1.1\overline{66} \end{bmatrix}$ as calculated	$\begin{bmatrix} -0.0664 & -0.1447 \\ 0.8685 & 0.6574 \end{bmatrix}$ as calculated	$\begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}$ as calculated	$\begin{bmatrix} 0.34-j0.22 & -0.75 \\ 0.65 & 0.55 \end{bmatrix}$ as entered
NDIM	NDIM = 3	NDIM = 2	NDIM = 2	NDIM = 2
MDIM	MDIM = 2	MDIM = 2	MDIM = 2	MDIM = 2
Process Noise Covariance Intensity Matrix Q	continuous time version $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	continuous time version $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	continuous time version $\begin{bmatrix} 10^{-[illegible]} & 0 \\ 0 & 10^{-[illegible]} \end{bmatrix}$	Discrete time version $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Observation Matrix C	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Measurement Noise Covariance Intensity Matrix R	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 10^{-[illegible]} & 0 \\ 0 & 10^{-[illegible]} \end{bmatrix}$	$\begin{bmatrix} 10^{-[illegible]} & 0 \\ 0 & 10^{-[illegible]} \end{bmatrix}$
Initial mean $\overline{X}_o$	$\begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$	$\begin{bmatrix} 0 & 0 \end{bmatrix}^T$	$\begin{bmatrix} 10 & 4 \end{bmatrix}^T$	$\begin{bmatrix} 0 & 0 \end{bmatrix}^T$
Initial Covariance P_o	$\begin{bmatrix} 6 & 2 & 1 \\ 2 & 8 & 3 \\ 1 & 3 & 12 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 10^{-[illegible]} & 0 \\ 0 & 10^{-[illegible]} \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Table 2: Simulator Testability Coverage Matrix

FUNCTION	CASE 1	CASE 2	CASE 3	CASE 4
Transition Matrix Computation: Pade (Ward's Algorithm)	√	√		
Transition Matrix Computation: Pade (Kleinman's Algorithm)	√	√		
Q_d Computation: Discrete-Time Equivalent of Continuous-Time White Noise	√	√		
Steady-State Computation of Initial Condition Mean		√		
Steady-State Computation of Initial Condition Covariance (Lyapunov Equation Solution)		√		
Verification of SVD-based Positive Definiteness Test for Nondiagonal Matrices	√			
Verification of Abbreviated Positive Definiteness Test for Diagonal Matrices	√	√	√	√
Checked Process Noise Calculations as Output from Random Number Generator	√	√		
Checked Measurement Noise Calculations as Output from Random Number Generator	√	√		
Checked Recursive Calculation of all Constituent Components of Entire Random Process Over Several Iterations	√	√		
Checked Proper Handling of PRN Seed	√	√		
Verification of Stable Sample Functions Indicative of Stationary Process		√		√
Verification of Unstable Sample Functions Indicative of Nonstationary Process	√		√	
Obvious Aggregate High Level At-A-Glance Confirmation From Output that all Functions Work Properly in Concert			√	
Confirmation of Identical Results When Complex Version of Software Enabled	√	√	√	
Eventual Confirmation of Proper Sample Function Statistics from Downstream Spectral Estimation Software Module Outputs		√		√

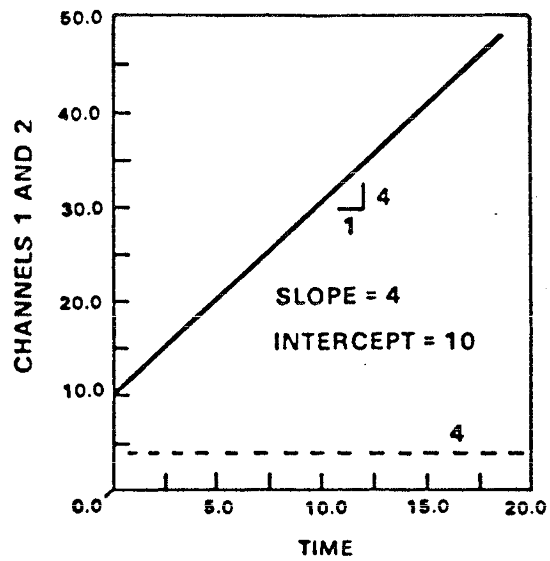
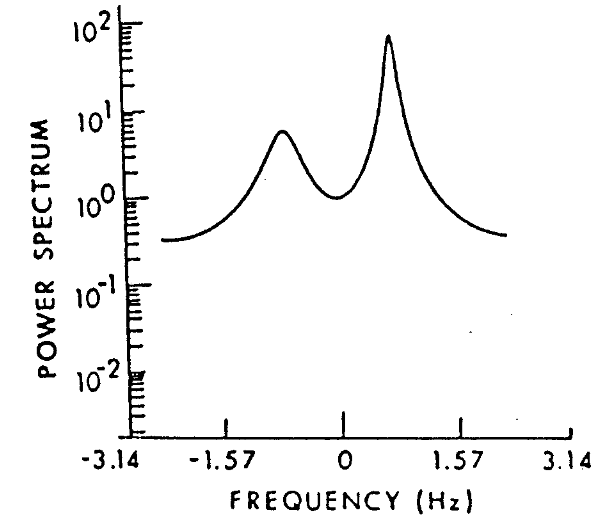


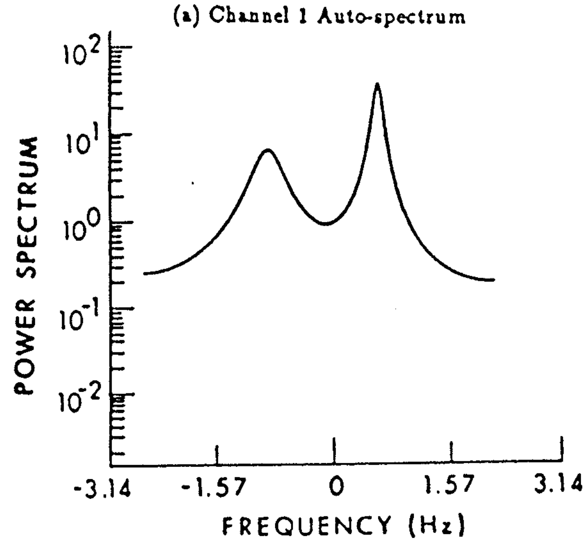
Figure 3: **ANSWER** for Test Case 3: DEGENERATE TEST CASE (with all noises present but cranked down to be miniscule) OF LINEAR RAMP OF KNOWN SLOPE AND INTERCEPT YIELDING CONFIRMING OUTPUT

Table 3: Standard Kalman Filter Implementation/Mechanization Equations

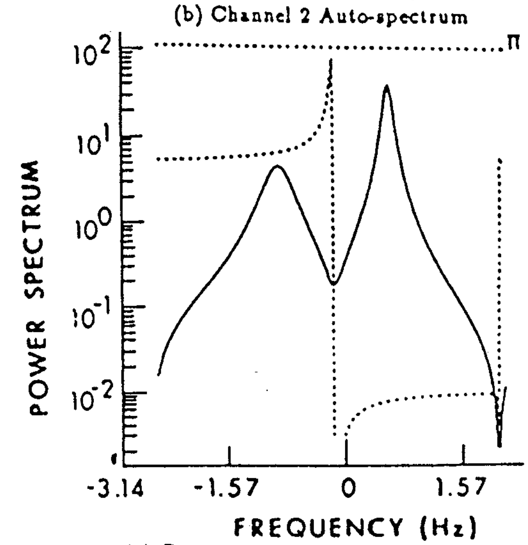
	PROPAGATE STEP	UPDATE STEP
COVARIANCE	$P_{k\|k-1} = \Phi(k,k-1)P_{k-1\|k-1}\Phi^T(k,k-1) + Q_k$	$P_{k\|k} = [I - K_k C_k]P_{k\|k-1}[I - K_k C_k]^T + K_k R K_k^T$
FILTER GAIN	$K_k = P_{k\|k-1}C_k^T[C_k P_{k\|k-1} C_k^T + R]^{-1}$	
FILTER	$\hat{x}_{k\|k-1} = \Phi(k,k-1)\hat{x}_{k-1\|k-1}$	$\hat{x}_{k\|k} = \hat{x}_{k\|k-1} + K_k(z_k - C_k \hat{x}_{k\|k-1})$



(a) Channel 1 Auto-spectrum



(b) Channel 2 Auto-spectrum



(c) Cross-spectrum of Channels 1 and 2

Figure 4: True Spectrum for the Two Channel Complex Test Case 4

output (MIMO) complex random process output with specified cross-correlation between output channels. This was needed in [8] in verifying the performance of alternative Maximum Entropy spectral estimators down-stream of the simulator (operating on its outputs), which, like a KF, deal only with first and second order statistics. The correct answer for 2-channel spectral estimation should appear as in Fig. 3. Certain modern tracking radars use coherent phase processing, also known as coherent integration (where both magnitude and phase are accounted for in the summation of signal returns but where the distinction arises of having to keep track of real and imaginary components, instead of merely needing to keep track of magnitude alone, as conventional radars do), which jointly treats Primary Polarization (PP) returns in conjunction with Orthogonal Polarization (OP) returns and utilizes the additional target information provided from the cross-correlation of these two separate channels across adjacent range gates for particular target signatures.

An overview of the complete software test coverage offered here through selective use of analytic closed-form "Test Cases of known solution" is provided in Table 2. The utility of this coverage is discussed above. All items indicated in Table 2 should be successfully confirmed in descending order.

From the right column and last line of Table 3, observe that immediately following the Kalman gain, K_k, there is a driving term in the update form of the estimator portion of a Kalman filter known as the Kalman filter residual, $(z_k - C_k\hat{x}_{k|k-1})$. In verifying and debugging an actual Kalman filter software implementation, these residuals are monitored and used as a gauge-of-goodness and indicate good tracking performance when they become "small." The idea being that the measurements z_k match the model representation $C_k\hat{x}_{k|k-1}$ fairly closely when the residuals are "small." However, since residuals are never identically zero, the question is "how small is small enough" [4]? Residuals will almost always initially decrease as the initial transient settles out. Additionally, "small residuals" are necessary but not sufficient indicators of good Kalman filter performance and similar statements can be made for having statistically white residuals (for instance, see [20], [21] [5] which offer an example of a Kalman filter exhibiting white residuals despite known use of an incorrect system model but which also incurs an anomalous bias as the clue that something is wrong). When possible, as with all simulations, one should juxtapose the time evolution of any critical system states along side their Kalman filter estimates to see how closely the estimates are following the actual quantities of interest as a more encompassing gauge of proper Kalman filter performance. The only problem sometimes encountered in certain sensitive applications is that actual estimates may be classified while residuals may be unclassified, in which case attention centers on the residuals in unclassified presentations as a default in justifying good filter performance. For actual real system data, the true system state uncontaminated by measurement noise is seldom available

[4] See [22] for an appealing explicit statistical test on the residuals using Chi-square statistics with appropriately specified degrees-of-freedom for an assortment of likely test conditions that can occur.

[5] Residuals are also sometimes called *innovations*.

so use of residuals must suffice in this situation also.

Restating for emphasis, the modern simulator design, discussed in this section, was pursued so that only a fairly exact mechanization would be used so that the input to the KF is precisely known. This was sought as a reliable testbed that avoids use of uncalibrated approximations in order to avoid confusing artifacts of simulator approximations with possible cross-channel feedthrough (that multichannel spectral estimation implementations are also known to frequently exhibit as a weakness or vulnerability) and which can adversely affect KF testing as well for the same reasons of uncalibrated cross-correlations being present.

By using these or similar examples, certain qualitative and quantitative aspects of the software implementation can be checked for conformance to anticipated behavior as an intermediate benchmark, prior to modular replacement of the various higher-order matrices appropriate to the particular application. This procedure is less expensive in CPU time expenditure during the software debug and checkout phase than using the generally higher n-dimensional matrices of the intended application since the computational burden is generally at least a cubic polynomial in n during the required solution of a Matrix Riccati equation for the associated covariances.

6 Lyapunov Equation Solution

Solve both of the following algebraic matrix equations

$$P_1 = \left[e^{A_4\Delta}\right] P_1 \left[e^{A_4\Delta}\right]^T + Q_d , \tag{48}$$

and

$$0 = A_4 P_2 + P_2 A_4^T + Q , \tag{49}$$

for P_1 and P_2, respectively, where

$$Q_d = \begin{bmatrix} 0.094 & 0.016 \\ 0.016 & 0.634 \end{bmatrix} , \tag{50}$$

and

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} , \tag{51}$$

where, in the above,

$$A_4 = \begin{bmatrix} -5 & -1 \\ 6 & 0 \end{bmatrix} , \tag{52}$$

and $\Delta = 0.5$ (units).

6.1 ANSWERS:

Expanding out the symmetric matrix equations of concern to result in merely linear algebraic equations that are easily solved to yield:

$$P_1 = \begin{bmatrix} 0.116 & -0.0819 \\ -0.0819 & 1.08 \end{bmatrix} , \tag{53}$$

where

$$e^{A_4 t} = \begin{bmatrix} (3e^{-3t} - 3e^{-2t}) & (e^{-3t} - e^{-2t}) \\ (-6e^{-3t} + 6e^{-2t}) & (-2e^{-3t} + 3e^{-2t}) \end{bmatrix} , \tag{54}$$

and therefore

$$e^{A_4 0.5} = \begin{bmatrix} (3e^{-\frac{3}{2}} - 3e^{-1}) & (e^{-\frac{3}{2}} - e^{-1}) \\ (-6e^{-\frac{3}{2}} + 6e^{-1}) & (-2e^{-\frac{3}{2}} + 3e^{-1}) \end{bmatrix} = \begin{bmatrix} -0.0664 & -0.1447 \\ 0.8685 & 0.6574 \end{bmatrix} , \tag{55}$$

and

$$P_2 = \begin{bmatrix} 0.116 & -0.083 \\ -0.083 & 1.11 \end{bmatrix} . \tag{56}$$

7 Riccati Equation Solution: Steady-State

Solve for the steady-state solution of the Riccati equation before and after a measurement update as, respectively, the matrices $\tilde{P}$ and $\hat{P}$ occurring within the covariance propagate step and update step as

$$\tilde{P} = \Phi_5(I - K_5 C_5)\tilde{P}\Phi_5^T + B_5 Q_5 B_5^T , \tag{57}$$

and

$$\hat{P} = (I - K_5 C_5)\tilde{P} , \tag{58}$$

where the corresponding Kalman gain is

$$K_5 = \tilde{P}C_5^T \left[C_5 \tilde{P} C_5^T + R_5\right]^{-1} . \tag{59}$$

The particular parameter values to be used for the system matrix A_5 is

$$\Phi_5 = \begin{bmatrix} 1 & T & \frac{T^2}{2} \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix} , \tag{60}$$

(corresponding to $A_5 = I_3$, where $\Phi_5 = e^{A_5 T}$), the control gain matrix B_5 is

$$B_5 = \begin{bmatrix} \frac{T^3}{6} \\ \frac{T^2}{2} \\ T \end{bmatrix} , \tag{61}$$

and the observation matrix C_5 is

$$C_5 = [\,1 \;\; 0 \;\; 0\,] , \tag{62}$$

with process noise covariance intensity

$$Q_5 = (2.764620)^2 = 7.64620 , \tag{63}$$

and measurement noise covariance intensity

$$R_5 = 16 , \tag{64}$$

where, in the above, $T = 2$ (units) is the fixed time-step (between consecutive discrete measurements).

7.1 ANSWER:

As laid out in [10] after laboring through a lot of algebra and parameter scaling, we have to solve an associated biquartic equation [10, Eq. A9] as an intermediate calculation. In order to make this challenge somewhat easier, instead of first specifying Q_5 and R_5 in Eqs. 63 and 64 arbitrarily beforehand, the new contribution provided in [9] and recounted in abbreviated form here, is to use a trick of convenience by finding the value that makes the following associated biquartic [6] easy to solve for S:

$$S^4 - 6S^3 + 10S^2 - 6(1 + 2r^2)S + (1 + 3r^2) = 0 . \tag{65}$$

The trick is to force a convenient answer, as say,

$$S = 6 \tag{66}$$

to be a solution of Eq. 65 [10, Eq. A9] by choosing the value of r (appearing in Eq. 65) for convenience. This proper value of r can be selected by first performing the

[6] Recall that while quadratic equations are easy to solve, general cubics and quartics/biquartics are extremely challenging and messy in general.

following division exercise:

$$\begin{array}{r l l l l l l}
 & S^3 & +10S & +(54-12r^2) & & \text{remainder}: & +325-69r^2 \\
\hline
S-6 \,\big) & S^4 & -6S^3 & +10S^2 & -6(1+2r^2)S & +(1+3r^2) & \\
 & S^4 & -6S^3 & & & & \\
\cline{2-6}
 & & & 10S^2 & -6(1+2r^2)S & & \\
 & & & 10S^2 & -60S & & \\
\cline{4-6}
 & & & & (54-12r^2)S & +(1+3r^2) & \\
 & & & & (54-12r^2)S & -6(54-12r^2) & \\
\cline{5-6}
 & & & & & 1+3r^2+324-72r^2 &
\end{array} \tag{67}$$

So $S = 6$ is a root of Eq. 65 [7] if the remainder in the above is zero as

$$325 - 69r^2 = 0 \tag{68}$$

or

$$r^2 = \frac{325}{69} \Rightarrow r = \sqrt{\frac{325}{69}} = 2.17028\ . \tag{69}$$

Now from the equation following Eq. 15 in [10], we have that

$$r \triangleq \frac{12\sqrt{R_5}}{\sqrt{Q_5 T^3}}\ . \tag{70}$$

From Eq. 70 above,

$$2.17028 = r = \frac{12\sqrt{R_5}}{\sqrt{Q_5 T^3}}, \tag{71}$$

we can now take

$$T = 2 \tag{72}$$

and

$$\sqrt{R_5} = 4\ , \tag{73}$$

so that rearranging Eq. 71 with these two assignments of Eqs. 72 and 73 yields

$$\sqrt{Q_5} = \frac{12(4)}{2.17028(8)} = \frac{6}{2.17028} = 2.764620 \tag{74}$$

or, referring back, yields the following two specifications of Eqs. 63 and 64 that are necessary to be pinned-down for a well-posed KF. According to [10, prior to Fig.

[7] Arbitrary solutions of Eq. 65 can't be forced (as in seeking to make $S = 2$ be a solution) because the remainder term will correspond to an "imaginary" value for r, which needs to be a real variable to be viable in this application.

1], the dimensionless quantity r defined in Eq. 70 can be interpreted as a type of noise-to-signal ratio.

Now, following the procedure of [10] as explicitly laid out in [9] for this particular example, the 3×3 steady-state covariance prior to one of the periodic measurement updates is

$$\tilde{P}_1 = \begin{bmatrix} 504.6652 & 349.3374 & 126.1666 \\ \cdot & 263.79805 & 108.5650 \\ \cdot & \cdot & 57.6117 \end{bmatrix}, \tag{75}$$

and the 3×3 steady-state covariance immediately after one of the periodic measurement updates is

$$\hat{P}_1 = \begin{bmatrix} 15.5083 & 10.7345 & 3.8771 \\ \cdot & 29.4118 & 23.9143 \\ \cdot & \cdot & 27.0392 \end{bmatrix}, \tag{76}$$

8 Riccati Equation Solution: Time-Varying

Solve the associated continuous-time Riccati equation

$$\dot{P}(t) = A_6 P(t) + P(t) A_6^T - P(t) C_6^T R_6^{-1} C_6 P(t) + B_6 Q_6 B_6^T \tag{77}$$

corresponding to the following state variable model of the system (i.e., specify what the solution for P(t) is at t=8 units):

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_1(t)\,, \tag{78}$$

with statistics or associated expectations being

$$E[x(0)] = \begin{bmatrix} 0 \\ 0 \end{bmatrix};\ P_0 = \mathrm{Cov}[x(0)] = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix};\ E[u_1(t)u_1(\tau)] = \delta(t-\tau).$$

The continuous-time measurement structure or outputs of the measurement sensors are correspondingly defined as:

$$z(t) = \begin{bmatrix} 0 & 2 \end{bmatrix} \left(\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \right), \tag{79}$$

with associated statistics

$$E[v(t)v^T(\tau)] = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \delta(t-\tau),$$

where in the above, u_1, v_1, and v_2 are zero-mean, independent, white Gaussian noises that are all uncorrelated with the Gaussian random vector initial condition $x(0)$. In terms of fairly familiar standard notation for linear systems described by state variables, the following matrices suffice to summarize the parameter values to be encountered in the system of interest in this section:

$$A_6 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \; ; \; B_6 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \; ; \; C_6 = \begin{bmatrix} 0 & 2 \end{bmatrix} \; ; \; P_0 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \; ; \tag{80}$$

$$Q_6 = I_{1 \times 1} \; ; \; R_6 = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \; ; \; \overline{x}(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{81}$$

8.1 ANSWER:

The covariance of estimation error of the continuous-time Kalman filter is the solution of the following continuous-time Riccati equation:

$$\begin{aligned} \tfrac{d}{dt}P(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} P(t) + P(t) \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} &- P(t) \begin{bmatrix} 0 \\ 2 \end{bmatrix} \left[\tfrac{2}{\mathcal{N}_0}\right] [0 \; 2] P(t) + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ \text{with initial condition } P(0) &= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \end{aligned} \tag{82}$$

for $\mathcal{N}_0 = 8$. This is a nonlinear matrix Riccati equation of dimension $n = 2$, whose solution can be obtained by the standard device of solving a related linear problem of twice the dimension, $2n$, formed as

$$\frac{d}{dt}T(t) = \begin{bmatrix} A(t) & \vdots & B(t)QB^T(t) \\ \cdots & \cdots & \cdots \\ C^T(t)R^{-1}C(t) & \vdots & -A^T(t) \end{bmatrix} T(t), \tag{83}$$

with initial condition $T(0) = I_{2n \times 2n}$. In order to relate the differing solutions of Eqs. 82 and 83, $T(t)$ is partitioned as

$$T(t) = \begin{bmatrix} T_{11}(t) & \vdots & T_{12}(t) \\ \cdots & \cdots & \cdots \\ T_{21}(t) & \vdots & T_{22}(t) \end{bmatrix}; \tag{84}$$

from which a solution of the original covariance equation can be obtained [14, Eq. 184, p. 43] as:

$$P(t) = (T_{11}(t)P_0 + T_{12}(t))(T_{21}(t)P_0 + T_{22}(t))^{-1} \triangleq \Gamma_1(t)\Gamma_2^{-1}(t). \tag{85}$$

For the parameters of the present example, the differential equations of Eq. 83 for the time evolution of the matrix $T(t)$ becomes:

$$\frac{d}{dt}T(t) = \begin{bmatrix} 0 & 1 & \vdots & 0 & 0 \\ 0 & 0 & \vdots & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & 0 & 0 \\ 0 & \alpha^2 & \vdots & -1 & 0 \end{bmatrix} T(t) \tag{86}$$

with $T(0) = I_{4\times4}$, where for convenience in notation within the above and in what is to follow take

$$\alpha^2 \triangleq \frac{8}{\mathcal{N}_0} \tag{87}$$

and where the matrix on the right hand side of the differential equation is defined to be the matrix B in what follows. The solution to Eq. 86 can be obtained by first finding $\mathcal{L}^{-1}\{(sI-B)^{-1}\}$, where s is the Laplace transform variable.

It is easily demonstrated by hand calculations for the relatively sparse matrix B, that:

$$(sI-B)^{-1} = \frac{\operatorname{adj}(sI-B)}{\det(sI-B)} = \begin{bmatrix} \frac{1}{s} & \frac{1}{s^2-\alpha^2} & \frac{-1}{s^2(s^2-\alpha^2)} & \frac{1}{s(s^2-\alpha^2)} \\ 0 & \frac{s}{s^2-\alpha^2} & \frac{-1}{s(s^2-\alpha^2)} & \frac{1}{s^2-\alpha^2} \\ 0 & 0 & \frac{1}{s} & 0 \\ 0 & \frac{\alpha^2}{s^2-\alpha^2} & \frac{-1}{s^2-\alpha^2} & \frac{s}{s^2-\alpha^2} \end{bmatrix}. \tag{88}$$

Using partial fraction expansions and appropriately inverse Laplace transforming Eq. 88 yields:

$$\begin{aligned} T(t) &= e^{Bt}T(0) = \mathcal{L}^{-1}\{(sI-B)^{-1}\} \\ &= \begin{bmatrix} 1 & \alpha^{-1}\sinh\alpha t & \vdots & \alpha^{-2}t - \alpha^{-3}\sinh\alpha t & -\alpha^{-2} + \alpha^{-2}\cosh\alpha t \\ 0 & \cosh\alpha t & \vdots & \alpha^{-2} - \alpha^{-2}\cosh\alpha t & \alpha^{-1}\sinh\alpha t \\ \cdots\ \cdots & \cdots & \cdots & \cdots\qquad\cdots & \cdots \\ 0 & 0 & \vdots & 1 & 0 \\ 0 & \alpha\sinh\alpha t & \vdots & -\alpha^{-1}\sinh\alpha t & \cosh\alpha t \end{bmatrix}, \end{aligned} \tag{89}$$

which also satisfies the initial condition $T(0) = I_{4\times4}$ as a check. Now according to Eq. 85 (and Eq. 184 on p. 43 of [14]),

$$\Gamma_1(t) = T_{11}(t)P_0 + T_{12}(t); \quad \Gamma_2(t) = T_{21}(t)P_0 + T_{22}(t) \tag{90}$$

and, therefore,

$$\Gamma_1(t) = \begin{bmatrix} 2 + \alpha^{-2}t - \alpha^{-3}\sinh\alpha t & \vdots & 2\alpha^{-1}\sinh\alpha t + \alpha^{-2}\cosh\alpha t - \alpha^{-2} \\ \cdots & \cdots & \cdots \\ \alpha^{-2} - \alpha^{-2}\cosh\alpha t & \vdots & 2\cosh\alpha t + \alpha^{-1}\sinh\alpha t \end{bmatrix}, \tag{91}$$

$$\Gamma_2(t) = \begin{bmatrix} 1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ -\alpha^{-1}\sinh\alpha t & \vdots & 2\alpha\sinh\alpha t + \cosh\alpha t \end{bmatrix}, \tag{92}$$

from which we can finally reconstruct the sought for covariance of estimation error as:

$$P(t) = \Gamma_1(t)\Gamma_2^{-1}(t)\,, \tag{93}$$

which for $\mathcal{N}_0 = 8$ yields

$$\alpha^2 = \frac{8}{\mathcal{N}_0} = \frac{8}{8} = 1 \tag{94}$$

in the above Eqs. 91 and 92 in 93 to result in

$$\begin{aligned} P(t) &= \Gamma_1(t)\Gamma_2^{-1}(t) \\ &= \frac{1}{2\sinh t + \cosh t}\begin{bmatrix} (3+2t)\sinh t + (2+t)\cosh t & 2\sinh t + \cosh t - 1 \\ 2\sinh t + \cosh t - 1 & 2\cosh t + \sinh t \end{bmatrix}, \end{aligned} \tag{95}$$

which at $t = 8$ is

$$P(8) = \begin{bmatrix} 7.0000 & 0.9998 \\ 0.9998 & 1.0000 \end{bmatrix}. \tag{96}$$

Notice that $p_{11}(t)$ in Eq. 95 is increasing with time t because the state $x_1(t)$ is not observable from the measurements (since the observability grammian is only 1 rather than 2).

9 Pseudo-Inverse

Find the Moore-Penrose pseudo-inverse (of minimum norm) corresponding to the following matrices:

1. **Case 1:**

$$P_1 = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \tag{97}$$

2. **Case 2:**

$$P_2 = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} . \tag{98}$$

9.1 ANSWERS:

As calculated explicitly in [7, App. C], the Moore-Penrose pseudo-inverse (of minimum norm) corresponding to the matrices of Eqs. 97 and 98 are:

1. **Case 1:**

$$P_1^\dagger = \begin{bmatrix} -\frac{1}{3} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} , \tag{99}$$

2. **Case 2:**

$$P_2^\dagger = \frac{1}{25}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} . \tag{100}$$

10 Optimization: Finite Time Horizon

1. **Case 1:** For the scalar linear system

$$\dot{x}(t) = u(t), \text{ for } 0 \le t \le 1 , \tag{101}$$

with initial condition

$$x(0) = 3, \tag{102}$$

where $u(t)$ is the deterministic control to be specified.

MINIMIZE the scalar quadratic cost function : $J[u] = [x(1)]^2 + \alpha \int_0^1 [u(t)]^2\, dt,$ (103)

for fixed scalar

$$\alpha = 10 . \tag{104}$$

Synthesize the optimum controller as a feedback control. What is the minimum cost for the optimal control? Remember that we are only interested in a finite horizon solution up to time=1 (unit).

2. **Case 2:** For the linear system of the form

$$\dot{x}(t) = \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t), \text{ for } 0 \le t \le T , \tag{105}$$

with initial condition $x(0) = [0,0]^T$, find the deterministic control $u(t)$ that transfers the system of Eq. 105 to the specified final state

$$x(T) = x_1 \text{ (known)} \tag{106}$$

and

$$\text{Minimizes the scalar quadratic cost function : } C[u] = \int_0^T [u(t)]^2\, dt. \tag{107}$$

Let T=5 (units) and $x_1 = [10,2]^T$ in the above. A feedback solution is not required here. What is the magnitude of the optimum control at time t=2 (units)?

10.1 ANSWERS:

1. To handle a general linear system with quadratic cost function to be minimized over a finite-time horizon and with a convex function of the final time state as a term in the cost function, application of the *Pontryagin maximum principle* to the original problem reduces to the solution of the following two point boundary value problem (entirely described by the following four equations):

$$\dot{x}^* = 0 \cdot x^* + 1 \cdot \frac{1}{\alpha} \cdot 1 \cdot \eta^* = \frac{1}{\alpha} \cdot \eta^* \,, \tag{108}$$

$$\dot{\eta}^* = x \cdot 0 - \eta^* \cdot 0 = 0 \,, \tag{109}$$

$$\eta^*(1) = -1 \cdot x^*(1) \,, \tag{110}$$

$$x^*(0) = 3 \,. \tag{111}$$

Since a feedback control is desired, assume that the desired control is a time-varying matrix times the fedback state $x(t)$, as:

$$u^*(t) = E^*(t) x^*(t), \tag{112}$$

where

$$E^*(t) = U^{-1}(t) B^T(t) E(t) \,. \tag{113}$$

So for the present scalar problem, the matrices in the above matrix equation reduce to

$$e^*(t) = \frac{1}{\alpha} \cdot 1 \cdot e(t), \tag{114}$$

where $e(t)$ is the solution of the following (now scalar) Riccati equation:

$$\dot{e}(t) = 0 - 0 \cdot e - e \cdot 0 - e \cdot 1 \cdot \frac{1}{\alpha} \cdot 1 \cdot e = -\frac{1}{\alpha} e^2 \,, \tag{115}$$

with final condition

$$e(1) = -1\,. \tag{116}$$

The solution to this Riccati equation can be simply found by *separation of variables* here as:

$$\begin{aligned}
\int_e^{-1} \frac{de}{e^2} &= -\frac{1}{\alpha}\int_t^1 dt & (117)\\
-\frac{1}{e}\Big|_e^{-1} &= -\frac{1}{\alpha}t\Big|_t^1 & (118)\\
\frac{1}{e}+1 &= -\frac{1}{\alpha}[1-t] & (119)\\
\frac{1}{e} &= -1-\frac{1}{\alpha}+\frac{1}{\alpha}t = -(\frac{\alpha+1}{\alpha})+\frac{1}{\alpha}t & (120)\\
e(t) &= \frac{1}{-(\frac{\alpha+1}{\alpha})+\frac{1}{\alpha}t} = \frac{1}{\frac{1}{\alpha}t-(\frac{\alpha+1}{\alpha})} & (121)
\end{aligned}$$

Since $\alpha+1>1$ because $\alpha>0$ (Eq. 104), the denominator in the above Riccati equation solution of Eq. 121 could only be zero if $t=\alpha+1$; however, as already mentioned $\alpha+1>1$ and in the entire problem we are only considering time over the finite time horizon of $0 \le t \le 1$. Therefore, now unraveling, the desired solution for the optimum time-varying feedback gain $e^*(t)$ is:

$$e^*(t) = \frac{1}{\alpha}\cdot 1\cdot e(t) = \frac{1}{t-(\alpha+1)}\,, \tag{122}$$

and the optimum feedback control is

$$u^*(t) = e^*(t)x^*(t) = \frac{x^*}{t-(\alpha+1)}\,, \tag{123}$$

and the resulting state-variable trajectory is described by the solution of the following equation:

$$\dot{x}^* = 0\cdot x^* + 1\cdot u^*(t) = \frac{x^*}{t-(\alpha+1)}\,. \tag{124}$$

The differential equation of Eq. 124 can be solved via *separation of variables* also as:

$$\begin{aligned}
\int_3^{x^*}\frac{dx^*}{x^*} &= \int_0^t \frac{dt}{t-(\alpha+1)}\,, & (125)\\
\ln\frac{x^*}{3} &= \ln\frac{|t-(\alpha+1)|}{|-(\alpha+1)|} = \ln\left[\frac{(\alpha+1)-t}{(\alpha+1)}\right], & (126)\\
\exp\left[\ln\frac{x^*}{3}\right] &= \exp\{\ln\left[1-\frac{1}{(\alpha+1)}t\right]\}\,, & (127)\\
x^* &= 3\left[1-\frac{1}{(\alpha+1)}t\right]\,. & (128)
\end{aligned}$$

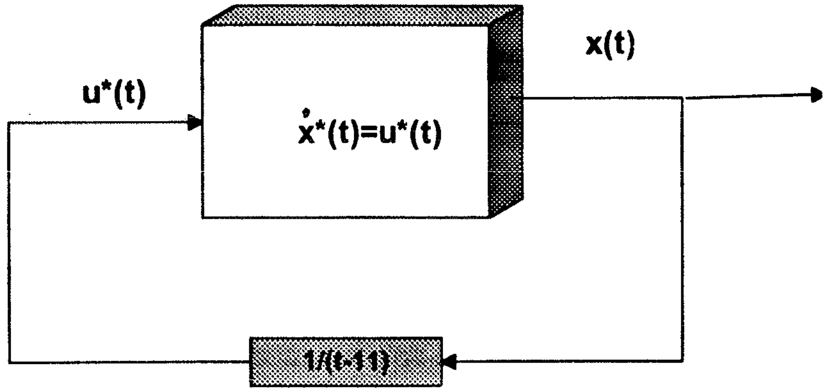


Figure 5: Feedback implementation of the optimal control solution for **Case 1**

From the above Eqs. 123 and 128, it may be concluded that the optimal control is:

$$u^*(t) = \frac{x^*(t)}{t-(\alpha+1)} = \frac{3\left[(\alpha+1)-t\right]}{(\alpha+1)\left[t-(\alpha+1)\right]} = \frac{-3}{(\alpha+1)} \tag{129}$$

and the square of the above optimal control is:

$$\left[u^*(t)\right]^2 = \frac{9}{(\alpha+1)^2} = \frac{9}{(11)^2} = 0.07438\ , \tag{130}$$

which, when substituted back into the original cost function to be minimized, yields the minimum cost to be:

$$\begin{aligned}
J[u^*] &= \left[x^*(1)\right]^2 + \int_0^1 \alpha\left[u^*(t)\right]^2 dt && (131)\\
&= 9\left[1-\frac{1}{(\alpha+1)}\right]^2 + \int_0^1 \alpha\frac{9}{(\alpha+1)^2}\,dt && (132)\\
&= 9\left[\frac{\alpha}{\alpha+1}\right]^2 + \frac{9\alpha}{(\alpha+1)^2} = \frac{9}{(\alpha+1)^2}\alpha(\alpha+1) && (133)\\
&= 9\left[\frac{\alpha}{\alpha+1}\right] = \frac{90}{11} = 8.1818 \text{ where } \alpha = 10 > 0\ . && (134)
\end{aligned}$$

This completes the problem of **Case 1.**

2. The system of Eq. 105 is a linear system of the form of

$$\underline{\dot{x}}(t) = F\underline{x}(t) + \underline{b}u(t)\ , \tag{135}$$

with

$$F = \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix}, \tag{136}$$

and

$$\underline{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{137}$$

and further the system matrix F of Eq. 136 is observed to be idempotent in that

$$F\,F = F. \tag{138}$$

One benefit of dealing with an idempotent system matrix, F, is that the Kalman "rank test" for "controllability" (as the standard regularity condition that must be satisfied before an optimal control can be sought) usually degenerates into a much more tractable expression for the *Controllability Grammian* that one must check the rank of but, in the case of the present 2-dimensional example, the simplified expression and the general expression are identical here (there would have been a simplification if the 3-dimensional idempotent matrix of Eq. 22 had been selected for use as the system matrix in optimization problem **Case 2** instead of the 2-dimensional idempotent matrix of Eq. 21), both being simply:

$$\left[\underline{b} \vdots A\underline{b}\right] = \begin{bmatrix} 0 & \vdots & -\frac{2}{5} \\ & \cdot & \\ 1 & \vdots & \frac{1}{5} \end{bmatrix}, \tag{139}$$

which by examination is of rank=2 (=n in this problem), so the system of Eq. 135 is in fact controllable.

Recall from [2, pp. 97-100], that for a controllable system (as established above) of the form of Eq. 135, with zero initial condition, and final condition x_1, where the minimization of a simplified or abbreviated finite-horizon quadratic cost function of the form of Eq. 107 is sought, then the final form of the optimal control that solves this problem (as worked out in [2, App. C] from a referenced theorem in Luenberger's 1969 textbook *Optimization by Vector Space Methods*) is:

$$u(t) = \underline{b}' e^{F'(T-t)} \left[\int_0^T e^{F(T-s)}\, \underline{b}\underline{b}'\, e^{F'(t-s)}\, ds\right]^{-1} x_1 \,. \tag{140}$$

Please notice the similarity between the expression within brackets to be inverted in the above Eq. 140 and the structure that is routinely encountered in computing the *exact* discrete-time equivalent to continuous-time white process noise covariance intensity matrix as Eqs. 17 and 38. The asserted similarity is even more striking when the two pre- and post-multiplying matrices in Eq. 17 are brought back under the integral sign as the following rerepresentation:

$$Q_d = \left[\int_0^\Delta e^{F(\Delta-\tau)}\, B\, Q\, B^T\, e^{F^T(\Delta-\tau)}\, d\tau\right] \delta_{kj}\,, \tag{141}$$

where Δ is used here in Eq. 38 in the role of T in Eq. 140, B serves in the role of $\underline{b}$, Q serves in the role of $I_{n\times n}$, and the dummy variable of integration τ is used here in the role of s in Eq. 140. From the simplification offered in [1, Eq.

42] as a closed-form evaluation of Eq. 141 when F is an idempotent matrix, we have that Eq. 39 again applies. The form of the answer in Eq. 140 may be further simplified as:

$$u(t) = \underline{b}'e^{F'(T-t)}M^{-1}x_1 \,, \tag{142}$$

where M is explicitly defined further below in Eq. 146 (using the result of Eq. 39), the apostrophe,', is now used to indicate matrix transpose to avoid confusion with T that is now used to represent the fixed known (but arbitrary) final time, and calculation of the exponential evaluation as a component of the first term on the right hand side of Eq. 142 simplifies for idempotent matrices to be:

$$\begin{aligned}
e^{F(T-t)} &= \sum_{k=0}^{\infty} \tfrac{F^k}{k!}(T-t)^k \\
&= I + \tfrac{F}{1!}(T-t) + \tfrac{F^2}{2!}(T-t)^2 + \tfrac{F^3}{3!}(T-t)^3 + \cdots \\
&= I + F(\tfrac{(T-t)}{1!} + \tfrac{(T-t)^2}{2!} + \tfrac{(T-t)^3}{3!} + \cdots) \\
&= I + F(e^{(T-t)} - 1) \\
&= I + \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix} (e^{(T-t)} - 1) = \begin{bmatrix} 1 + \frac{4}{5}(e^{(T-t)} - 1) & -\frac{2}{5}(e^{(T-t)} - 1) \\ -\frac{2}{5}(e^{(T-t)} - 1) & 1 + \frac{1}{5}(e^{(T-t)} - 1) \end{bmatrix},
\end{aligned} \tag{143}$$

by arguments identical to those offered for deriving Eq. 25. Therefore, the lead component in Eq. 142 is

$$\begin{aligned}
\underline{b}'e^{F'(T-t)} &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 + \frac{4}{5}(e^{(T-t)} - 1) & -\frac{2}{5}(e^{(T-t)} - 1) \\ -\frac{2}{5}(e^{(T-t)} - 1) & 1 + \frac{1}{5}(e^{(T-t)} - 1) \end{bmatrix} \\
&= \begin{bmatrix} -\frac{2}{5}(e^{(T-t)} - 1) & 1 + \frac{1}{5}(e^{(T-t)} - 1) \end{bmatrix} .
\end{aligned} \tag{144}$$

The remaining challenging term on the right hand side of Eq. 142, consisting of the inverse of the expression of Eq. 141 to be evaluated for the variables assignment of the problem of **Case 2** Eq. 39, becomes:

$$M \triangleq \left[\int_0^T e^{F\,(T-s)}\, \underline{bb}'\, e^{F'\,(T-s)}\, ds\right] \tag{145}$$

$$\begin{aligned}
&= \left[\mathrm{I} + \mathrm{F}(\mathrm{e}^{\mathrm{T}} - 1)\right] \{\int_0^{\mathrm{T}} \left[\mathrm{I} + \mathrm{F}(\mathrm{e}^{-s} - 1)\right] \underline{\mathrm{bb}}' \left[\mathrm{I} + \mathrm{F}'(\mathrm{e}^{-s} - 1)\right] \mathrm{ds}\} \left[\mathrm{I} + \mathrm{F}'(\mathrm{e}^{\mathrm{T}} - 1)\right] \\
&= \left[I + F(e^T - 1)\right] \\
&\quad \int_0^T \left[\underline{bb}' + (F\underline{bb}' + \underline{bb}'F')(e^{-s} - 1) + F\underline{bb}'F'(e^{-2s} - 2e^{-s} + 1)\right] ds
\end{aligned}$$

$$
\begin{aligned}
&\Big[I + F'(e^T - 1)\Big] \\
&= \Big[I + F(e^T - 1)\Big] \\
&\quad \Big[\underline{bb}'\mathrm{T} + (\mathrm{F}\underline{bb}' + \underline{bb}'\mathrm{F}')(1 - \mathrm{e}^{-\mathrm{T}} - \mathrm{T}) + \mathrm{F}\underline{bb}'\mathrm{F}'(-\frac{3}{2} - \frac{1}{2}\mathrm{e}^{-2\mathrm{T}} + 2\mathrm{e}^{-\mathrm{T}} + \mathrm{T})\Big] \\
&\quad \Big[I + F'(e^T - 1)\Big] \\
&= \begin{bmatrix} 1 + \frac{4}{5}(e^T - 1) & -\frac{2}{5}(e^T - 1) \\ -\frac{2}{5}(e^T - 1) & 1 + \frac{1}{5}(e^T - 1) \end{bmatrix} \\
&\Big[\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\mathrm{T} + (\begin{bmatrix} 0 & -\frac{2}{5} \\ -\frac{2}{5} & \frac{2}{5} \end{bmatrix})(1 - \mathrm{e}^{-\mathrm{T}} - \mathrm{T}) + \begin{bmatrix} \frac{4}{25} & -\frac{2}{25} \\ -\frac{2}{25} & \frac{1}{25} \end{bmatrix}(-\frac{3}{2} - \frac{1}{2}\mathrm{e}^{-2\mathrm{T}} + 2\mathrm{e}^{-\mathrm{T}} + \mathrm{T})\Big] \\
&\quad \begin{bmatrix} 1 + \frac{4}{5}(e^T - 1) & -\frac{2}{5}(e^T - 1) \\ -\frac{2}{5}(e^T - 1) & 1 + \frac{1}{5}(e^T - 1) \end{bmatrix} \\
&= \begin{bmatrix} 1 + \frac{4}{5}(e^T - 1) & -\frac{2}{5}(e^T - 1) \\ -\frac{2}{5}(e^T - 1) & 1 + \frac{1}{5}(e^T - 1) \end{bmatrix} \\
&\quad \begin{bmatrix} \frac{4}{25}(-\frac{3}{2} - \frac{1}{2}e^{-2T} + 2e^{-T} + T) & -\frac{2}{5}(1 - e^{-T} - T) - \frac{2}{25}(-\frac{3}{2} - \frac{1}{2}e^{-2T} + 2e^{-T} + T) \\ -\frac{2}{5}(1 - e^{-T} - T) - \frac{2}{25}(-\frac{3}{2} - \frac{1}{2}e^{-2T} + 2e^{-T} + T) & T + \frac{2}{5}(1 - e^{-T} - T) + \frac{1}{25}(-\frac{3}{2} - \frac{1}{2}e^{-2T} + 2e^{-T} + T) \end{bmatrix} \\
&\quad \begin{bmatrix} 1 + \frac{4}{5}(e^T - 1) & -\frac{2}{5}(e^T - 1) \\ -\frac{2}{5}(e^T - 1) & 1 + \frac{1}{5}(e^T - 1) \end{bmatrix},
\end{aligned} \tag{1
$$

and, since

$$
M = M_1 M_2 M_3 \Longrightarrow M^{-1} = M_3^{-1} M_2^{-1} M_1^{-1}\ , \tag{147}
$$

we have that the explicit components of the requisite inverse are:

$$
M^{-1} = \frac{1}{25}\begin{bmatrix} 4e^{-T} + 1 & 2 - 2e^{-T} \\ 2 - 2e^{-T} & e^{-T} + 4 \end{bmatrix} M_2^{-1} \begin{bmatrix} 4e^{-T} + 1 & 2 - 2e^{-T} \\ 2 - 2e^{-T} & e^{-T} + 4 \end{bmatrix}, \tag{148}
$$

where

$$
M_2^{-1} = \frac{1}{\ell}\begin{bmatrix} T + \frac{2}{5}(1 - e^{-T} - T) + \frac{1}{25}(-\frac{3}{2} - \frac{1}{2}e^{-2T} + 2e^{-T} + T) & \frac{2}{5}(1 - e^{-T} - T) + \frac{2}{25}(-\frac{3}{2} - \frac{1}{2}e^{-2T} + 2e^{-T} + T) \\ \frac{2}{5}(1 - e^{-T} - T) + \frac{2}{25}(-\frac{3}{2} - \frac{1}{2}e^{-2T} + 2e^{-T} + T) & \frac{4}{25}(-\frac{3}{2} - \frac{1}{2}e^{-2T} + 2e^{-T} + T) \end{bmatrix} \tag{149}
$$

and where $\ell = \frac{4}{25}\Big[T(-\frac{3}{2} - \frac{1}{2}e^{-2T} + 2e^{-T} + T) - (1 - e^{-T} - T)^2\Big]$. Now to complete this problem (by evaluating Eq. 142) all we have to do is multiply out the result of Eq. 144 times that of Eq. 148 with the result of Eq. 149 inserted for M_2^{-1} in Eq. 148, then post-multiplied by the final condition $x_1 = [10, 2]^T$ to yield (at $t = 2$ and $T = 5$):

$$u^*(2) = -0.0073\,. \tag{150}$$

10.2 Distinct optimization problems that appear to be similar to CASE 1:

1. **Case 3:** For the scalar linear system

$$\dot{x} = u(t), \text{ for } 0 \leq t \leq 1\,, \tag{151}$$

with initial condition

$$x(0) = 3, \tag{152}$$

where $u(t)$ is the deterministic control to be specified.

$$\text{MINIMIZE the scalar quadratic cost function : } J[u] = \underline{x(1)} + \alpha \int_0^1 [u(t)]^2\, dt, \tag{153}$$

for general, fixed scalar

$$\alpha > 0\,. \tag{154}$$

Unlike the feedback formulation obtained for the solution to **Case 1** at the beginning of Section 10, it is suggested that the optimum controller not be synthesized here as a feedback control for this problem but merely as an explicitly stated control solution as a function of time. What is the minimum cost for the optimal control? Remember that we are only interested in a finite horizon solution up to time=1.

2. **Case 4:** For the scalar linear system

$$\dot{x} = u(t), \text{ for } 0 \leq t \leq 1\,, \tag{155}$$

with initial condition

$$x(0) = 3, \tag{156}$$

and final condition

$$x(1) = 0, \tag{157}$$

where $u(t)$ is the deterministic control to be specified.

$$\text{MINIMIZE the scalar quadratic cost function : } J[u] = +\alpha \int_0^1 [u(t)]^2\, dt, \tag{158}$$

for general, fixed scalar

$$\alpha > 0\,. \tag{159}$$

Unlike the feedback formulation obtained for the solution to **Case 1** at the beginning of Section 10, it is suggested that the optimum controller not be synthesized here as a feedback control for this problem but merely as an explicitly stated control solution as a function of time. What is the minimum cost for the optimal control? Remember that we are only interested in a finite horizon solution up to time=1.

Notice that these two problems differ between themselves and from the first problem, **Case 1** in Section 10, only in a small aspect in the form of the cost function to be minimized. However, the form and nature of the subsequent solutions will be significantly different between these three problems. It may be helpful to utilize the theorems within the textbook *Foundations of Optimal Control*, John Wiley, 1967, written by E. B. Lee and L. Markus; and the textbook *Linear Optimal Control Systems*, John Wiley, 1972, written by H. Kwakernaak and R. Sivan.

10.3 Radically different answers to similar looking problems:

1. **Answer to Case 3:** For a general linear system with quadratic cost function to be minimized over a finite-time horizon and with a slightly degenerate form of a convex function (linear) of the final time state as a term in the cost function, as in Section 3.3 of the textbook *Linear Optimal Control Systems* by Kwakernaak and Sivan <u>or</u> by Theorem 5 in the textbook by E. B. Lee and L. Markus entitled *Foundations of Optimal Control*), the original problem reduces to the solution of the following two-point-boundary-value-problem TPBVP (entirely described by the following four equations similar to what were obtained in the answer to **Case 1** above):

$$\dot{x}^* = 0 \cdot x^* + 1 \cdot \frac{1}{\alpha} \cdot 1 \cdot \eta^* = \frac{1}{\alpha} \cdot \eta^* \,, \tag{160}$$

$$\dot{\eta}^* = x^* \cdot 0 - \eta^* \cdot 0 = 0 \,, \tag{161}$$

$$\eta^*(1) = -\frac{1}{2} \cdot (\text{grad}_x\,[x(1)]) = -\frac{1}{2} \cdot (1) = -\frac{1}{2} \,, \tag{162}$$

$$x^*(0) = 3 \,. \tag{163}$$

Since a feedback control is <u>not</u> sought, we can avoid assuming that the desired control is a time-varying matrix times the fedback state $x(t)$. Since we are <u>not</u> interested in feedback solutions, we can avoid entirely having to deal with any Riccati equation.

Instead, we now unravel the optimum solution directly by just using the *maximum principle* found in Lee and Markus' textbook (or as a direct application of the *Pontryagin Maximum Principle*), the desired solution for the optimum control can be found from the following solution procedure: From Eq. 161,

$$\dot{\eta}^* = 0 \Rightarrow \text{ solution} : \eta^*(t) = c_1 \text{ (a constant)} . \tag{164}$$

Therefore, by using the known final value for $\eta^*(1)$ from the above Eq. 162 yields

$$-\frac{1}{2} = \eta^*(1) = c_1 \Rightarrow c_1 = -\frac{1}{2} \Rightarrow \eta^*(t) = -\frac{1}{2} \text{ for } 0 \leq t \leq 1 . \tag{165}$$

Now from Eqs. 160 and 164, the optimal state-variable trajectory is described by the solution of the following equation:

$$\dot{x}^* = \frac{1}{\alpha} \cdot \eta^* = \frac{1}{\alpha} \cdot -\frac{1}{2} = -\frac{1}{2\alpha} , \tag{166}$$

from which a solution may be obtained as

$$\int_3^{x^*} dx^* = -\frac{1}{2\alpha} \int_0^t dt \text{ for } 0 \leq t \leq 1 , \tag{167}$$

hence

$$x^*(t) = 3 - \frac{1}{2\alpha} t \text{ for } 0 \leq t \leq 1 , \tag{168}$$

From the maximum principle applied to linear systems and for a quadratic cost function (Theorem 5 of Lee and Markus), it may be concluded that the optimal control is:

$$u^*(t) = \frac{1}{\alpha} \cdot 1 \cdot (-\frac{1}{2}) = -\frac{1}{2\alpha} \tag{169}$$

and the square of the above optimal control is:

$$[u^*(t)]^2 = \frac{1}{4(\alpha)^2} , \tag{170}$$

which, when substituted back into the original cost function to be minimized, yields the minimum cost to be:

$$\begin{aligned} J[u^*] &= [x^*(1)] + \int_0^1 \alpha \, [u^*(t)]^2 \, dt & (171) \\ &= \left[3 - \frac{1}{2\alpha}(1)\right] + \int_0^1 \frac{\alpha}{4\alpha^2} \, dt & (172) \\ &= \left[3 - \frac{1}{2\alpha}\right] + \frac{1}{4\alpha} = 3 - \frac{1}{4\alpha} . & (173) \end{aligned}$$

This completes solution to the problem of **Case 3**. Notice that $x^2(0) = 9$.

2. **Answer to Case 4:** Since a feedback control is <u>not</u> sought, we can avoid assuming that the desired control is a time-varying matrix times the fedback state $x(t)$. Since we are <u>not</u> interested in feedback solutions, we can avoid entirely having to deal with any Riccati equation. To handle a general linear system with quadratic cost function to be minimized over a finite-time horizon and with a degenerate form of a convex function of the final time state as a term in the cost function, as in Section 3.3 of the textbook *Linear Optimal Control Systems* by Kwakernaak and Sivan <u>or</u> by Theorem 3 in the textbook by E. B. Lee and L. Markus entitled *Foundations of Optimal Control*, the original problem reduces to the solution of the following two-point-boundary-value-problem TPBVP (entirely described by the following four equations similar to what were obtained in the answer to **Case 1** as Eqs. 108 to 111):

$$\dot{x}^* = 0 \cdot x^* + 1 \cdot u^*(t) \; , \tag{174}$$

$$u^*(t) = -\frac{1}{2\eta_o} \cdot \frac{1}{\alpha} \cdot 1 \cdot \eta^*(t) \; , \tag{175}$$

$$\dot{\eta}^* = -2\eta_o \cdot x^*(t) \cdot 0 - \eta^* \cdot 0 = 0 \; , \tag{176}$$

where

$$\eta_o < 0 \; , \tag{177}$$

and with prescribed initial and final conditions, respectively,

$$x^*(1) = 0 \; , \tag{178}$$

$$x^*(0) = 3 \; . \tag{179}$$

For specificity, select

$$\eta_(o) = -\frac{1}{2} \; , \tag{180}$$

then Eq. 175 becomes

$$u^*(t) = \frac{1}{\alpha}\eta^*(t) \tag{181}$$

and Eq. 175 has the solution

$$\dot{\eta}^*(t) = 0 \Rightarrow \eta^*(t) = c_1 \text{ (yet to be specified)} \; , \tag{182}$$

Then Eq. 174 has the solution

$$\int_3^{x^*} dx^* = \frac{c_1}{\alpha} \int_0^t dt \text{ for } 0 \le t \le 1 \; . \tag{183}$$

Therefore, the form of the solution is

$$x^*(t) = 3 + \frac{c_1}{\alpha} t \; , \tag{184}$$

and by Eqs. 178 and 184,

$$0 = x^*(1) = 3 + \frac{c_1}{\alpha}(1) , \tag{185}$$

so, as a consequence, we can now solve the above for c_1 as

$$c_1 = -3\alpha , \tag{186}$$

hence Eq. 184 becomes

$$x^*(t) = 3\left[1 - t\right] , \tag{187}$$

and from Eq. 181, the optimal control is

$$u^*(t) = \frac{1}{\alpha} \cdot 1 \cdot (-3\alpha) = -3 , \tag{188}$$

and the associated minimum cost is

$$J\left[u^*(t)\right] = \int_0^1 \alpha \left[u^*(t)\right]^2 dt = \int_0^1 \alpha \left[-3\right]^2 dt = 9\alpha . \tag{189}$$

This completes the answer to **Case 4**. Notice that $x^2(0) = 9$. The solution could have been worked out entirely for arbitrary nonnegative initial condition $x(0)$ but this would have made the algebra a little more challenging than it was here.

11 Handling Serially Time-Correlated Process Noise

For the linear system and measurement model of Test Case 2 in Section 5.2 (Table 1), assume that the process noise depicted is serially time correlated Gaussian noise having the following correlation matrix:

$$R_{uu}(\tau) = \begin{bmatrix} \frac{1}{6}e^{-2|\tau|} + \frac{1}{6}e^{-|\tau|} & \vdots & \frac{1}{4}e^{-2|\tau|} \\ \cdots & \cdots & \cdots \\ \frac{1}{4}e^{-2|\tau|} & \vdots & \frac{1}{2}e^{-|\tau|} \end{bmatrix} , \tag{190}$$

rather than uncorrelated white unit variance Gaussian noise, as depicted in Table 1 for Test Case 2. Show how this system should be configured (or augmented) for Kalman filtering (which theoretically expects only uncorrelated white noises) to still yield an optimal unbiased estimate of the state of this linear system.

11.1 ANSWER:

For the continuous-time linear system of Test Case 2 of the form

$$\dot{x}(t) = A\, x(t) + B\, u(t) \tag{191}$$

with sensor measurements

$$z(t) = C\, x(t) + v(t)\ , \text{where } v(t) \sim \mathcal{N}(0, R)\ , \tag{192}$$

and $u(t)$ has the correlation matrix of Eq. 190.

The associated *rational* power spectral density matrix (obtained as the Bilateral Laplace [8] transform of Eq. 190) is

$$\Phi_{uu}(s) = \begin{bmatrix} \frac{2-s^2}{(4-s^2)(1-s^2)} & \frac{1}{4-s^2} \\ \frac{1}{4-s^2} & \frac{1}{1-s^2} \end{bmatrix}, \tag{193}$$

where the "abscissa of convergence" in Eq. 193 is for $|Re(s)| < 1$. It is desired to model the above continuous-time power spectral matrix using a continuous-time multi-input/multi-output linear shaping filter driven by white noise. The proper shaping filter to be used to accomplish this can be obtained by applying standard Matrix Spectral Factorization algorithms [18], [19], to yield a solution of the form

$$\Phi_{uu}(s) = W^T(-s)\, W(s). \tag{194}$$

Since the relationship between the input spectral matrix, $\Phi_{ww}(s)$, and the output spectral matrix, $\Phi_{uu}(s)$, for a strictly stable linear time-invariant system is

$$\Phi_{uu}(s) = \mathcal{H}(-s)\, \Phi_{ww}(s)\, \mathcal{H}^T(s) \tag{195}$$

where

$\mathcal{H}(s) =$ a not necessarily square transfer function matrix of the linear system;

it is convenient to treat the modeling as involving a zero mean, uncorrelated Gaussian white noise input (having spectral matrix being the identity matrix)

$$\Phi_{ww}(s) = I_n \tag{196}$$

as driving the linear system to yield the following degenerate simplification of Eq. 195 as representing the output spectral matrix by

$$\Phi_{uu}(s) = \mathcal{H}(-s)\, \mathcal{H}^T(s)\ . \tag{197}$$

[8] Throughout this discussion, s is the complex bilateral Laplace transform variable as in the Laplace transform kernel $\exp(-st)$. The Bilateral Laplace transform is closely related to the Fourier transform (i.e., $s = \jmath\omega$) but is more convenient for the manipulations to be performed here of spectral factorization.

By an obvious association between the factors of Eq. 194 and the representation of Eq. 197, it is seen that an appropriate transfer function to model the power spectral matrix $\Phi_{uu}(s)$ is

$$\mathcal{H}(s) \equiv W^T(s) \; . \tag{198}$$

That the final result of computations does in fact yield a valid solution can easily be verified or confirmed merely by multiplying the asserted solution matrix by its conjugate transpose to again obtain the original power spectral density matrix of Eq. 193 as a check. Additionally, one should then proceed by checking the solution matrix to be certain that no poles occur in the right half s-plane (LHP) or on the $j\omega$-axis, otherwise the result does not correspond to a *strictly stable* or to a *stationary* random process.

For the specific example of Eq. 193, Matrix Spectral Factorization using Youla's first approach (detailed in [4, App. B]) yields the following factor as an answer

$$W(s) \;=\; \begin{bmatrix} \frac{-s-(\sqrt{7}/2)}{(2+s)(1+s)} & \frac{-s-(\sqrt{7}/2)}{(2+s)(1+s)} \\ & \\ \frac{-1/2}{(2+s)(1+s)} & \frac{3/2}{(2+s)(1+s)} \end{bmatrix} , \tag{199}$$

which checks since for the factor in Eq. 199, the product $W^T(-s)\,W(s)$ is again the specific power spectral matrix of Eq. 193.

Now we need to determine a linear system realization of the transfer function consisting of the transpose of Eq. 199 which is of the form:

$$\dot{x} = F_{TOTAL}\, x + G_{TOTAL}\, w \; ; y = H_{TOTAL}\, x \; , \tag{200}$$

where

$$F_{TOTAL} = \mathrm{diag}\left[F_1, F_2, \ldots, F_m\right] \; , \tag{201}$$

$$G_{TOTAL} = \left[G_1, \ldots, G_m\right]^T \; , \tag{202}$$

$$H_{TOTAL} = \mathrm{diag}\left[H_1, \ldots, H_m\right] \; , \tag{203}$$

and x is the augmented state vector of dimension $n = \sum_{i=1}^{m} n_i$. The obvious check is that the original transfer function matrix should satisfy the following as an identity:

$$\mathcal{H}(s) = H_{TOTAL}\,(sI - F_{TOTAL})^{-1}\, G_{TOTAL} \; . \tag{204}$$

For concreteness, the above mentioned technique is demonstrated next as applied to the realization having Eq. 193 as the output power spectral density matrix.

Referring to the Matrix Spectral Factorization result of Eq. 199, the associated transfer function matrix is its transpose. The corresponding matrix transfer function block diagram is depicted in Fig. 6, as can be verified by inspection. From the transfer function block diagram, the system differential equations are obtained which, unfortunately, also happen to contain differentiation of the inputs as the more general

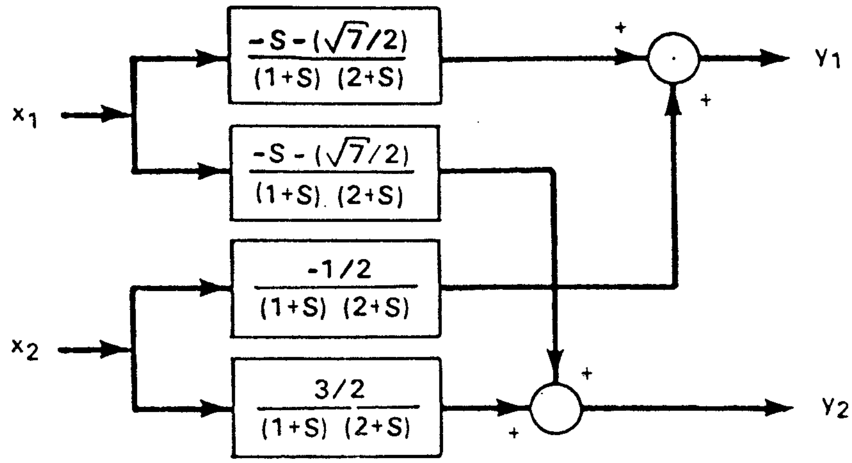


Figure 6: Two-Input/Two-Output Transfer Function Implementation of the Spectral Factor of Eq. 199.

situation that is most likely to be encountered in practice. For this example, these differential equations are:

$$\ddot{y}_1 + 3\,\dot{y}_1 + 2\,y_1 = -\dot{x}_1 - \sqrt{7}/2\,x_1 - (1/2)\,x_2, \tag{205}$$

$$\ddot{y}_2 + 3\,\dot{y}_2 + 2\,y_2 = -\dot{x}_1 - \sqrt{7}/2\,x_1 + (3/2)\,x_2\,, \tag{206}$$

where in the above the use of $\cdot$ above a variable denotes differentiation with respect to time as $\frac{d}{dt}$.

In the next few steps, a procedure outlined in [3, pp. 334-335] (and generalized for large dimensional systems as an algorithm in [5]) for removing the differentiated input occurring in the differential equations is now applied here. Let

$$w_1 = y_1 \text{ and } w_2 = \dot{y}_1 - k_1\,x_1; \tag{207}$$

then

$$\dot{w}_1 = w_2 + k_1\,x_1 \tag{208}$$

and substituting the above in Eq. 205 and rearranging, yields

$$\dot{w}_2 = (-1 - k_1)\,\dot{x}_1 - (\sqrt{7}/2 + 3\,k_1)\,x_1 - (1/2)\,x_2 - 2\,w_1 - 3\,w_2\,; \tag{209}$$

which upon taking

$$k_1 = -1, \tag{210}$$

then zeroes out the coefficient of the $\dot{x}_1$ term to yield

$$\dot{w}_2 = -2\,w_1 - 3\,w_2 + (6 - \frac{\sqrt{7}}{2})\,x_1 - (1/2)\,x_2. \tag{211}$$

Similarly, let

$$r_1 = y_2 \text{ and } r_2 = \dot{y}_2 - k_2\,x_1; \tag{212}$$

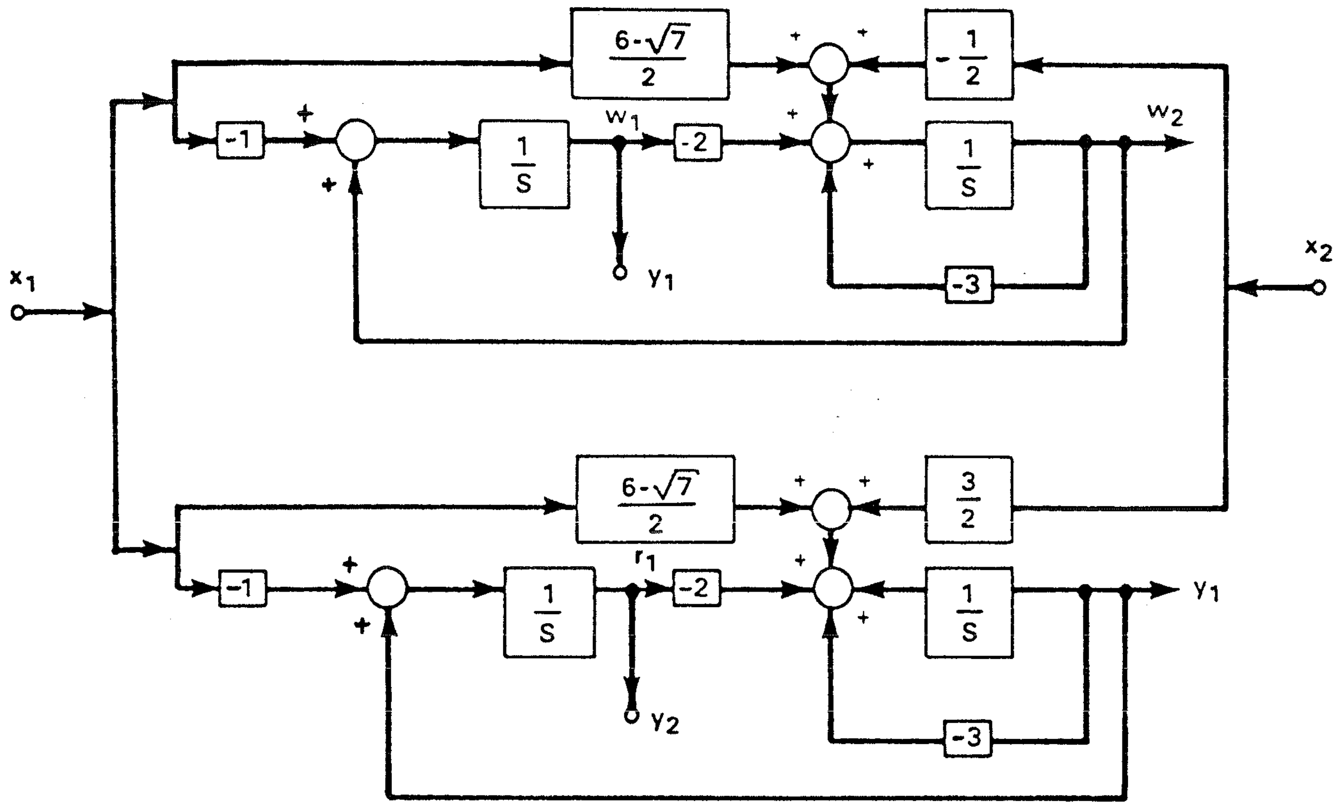


Figure 7: Implementation Detail of Transfer Function Realization of Matrix Factor Solution of Eq. 199.

then

$$\dot{r}_1 = r_2 + k_2\, x_1\ , \tag{213}$$

and substituting the above in Eq. 206 and rearranging with

$$k_2 = -1, \tag{214}$$

then zeroes out the coefficient of the differentiated input term to yield

$$\dot{r}_2 = -2\, r_1 - 3\, r_2 + (6 - \frac{\sqrt{7}}{2})\, x_1 + (3/2)\, x_2. \tag{215}$$

The analog computer block diagram which summarizes this procedure is offered in Fig. 7, where x_1 and x_2 are independent Gaussian white noises with unit variance. Notice that only four integrators are required for this simulation!

The resulting augmented system, using the technique offered in [5], can be reexpressed in vector/matrix form as

$$\frac{d}{dt}\begin{bmatrix} w_1 \\ w_2 \\ r_1 \\ r_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & -3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & -3 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \\ r_1 \\ r_2 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ (6-\sqrt{7})/2 & -1/2 \\ -1 & 0 \\ (6-\sqrt{7})/2 & 3/2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{216}$$

with

$$u(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \\ r_1(t) \\ r_2(t) \end{bmatrix}. \tag{217}$$

The stability of this augmented system can easily be checked by applying the Routh-Hurwitz criterion to the characteristic equation for the system matrix F in Eq. 216. The characteristic equation is

$$\begin{aligned} 0 &= \det[\lambda I_4 - F] = \det \begin{bmatrix} \lambda & -1 & 0 & 0 \\ 2 & (\lambda+3) & 0 & 0 \\ 0 & 0 & \lambda & -1 \\ 0 & 0 & 2 & (\lambda+3) \end{bmatrix} \\ &= \lambda^4 + 6\lambda^3 + 13\lambda^2 + 12\lambda + 4. \end{aligned} \tag{218}$$

The corresponding Routh-Hurowitz array formed from the coefficients of Eq. 218, from which stability may be inferred, is

$$\begin{array}{c|ccc} \lambda^4 & 1 & 13 & 4 \\ \lambda^3 & 6 & 12 & \\ \lambda^2 & 11 & 4 & \\ \lambda^1 & \frac{108}{11} & & \\ \lambda^0 & 4. & & \end{array} \tag{219}$$

Since there are no changes in sign along the first column to the right of the vertical line in the above Routh-Hurowitz array, the interpretation from its use is that there are no zeroes of the characteristic equation that have real parts greater than or equal to zero. Therefore, this particular augmented system of Eqs. 216 and 217 is stable (and, as a consequence, is a stationary random process).

By the Kalman rank test on the controllability Grammian, for the augmented system (F, G) to be a controllable pair, it must be that

$$\text{rank}[G \vdots FG \vdots F^2G \vdots F^3G] = 4. \tag{220}$$

It is sufficient to just test the 4×4 matrix that results from adjoining the first two indicated components as:

$$\det\left[G \vdots FG\right] = \det \begin{bmatrix} -1 & 0 & (6-\sqrt{7})/2 & -1/2 \\ (6-\sqrt{7})/2 & -1/2 & (3\sqrt{7}/2)-7 & 3/2 \\ -1 & 0 & (6-\sqrt{7})/2 & 3/2 \\ (6-\sqrt{7})/2 & 3/2 & (3\sqrt{7}/2)-7 & -9/2 \end{bmatrix} = 5\sqrt{7} - 15 \neq 0. \tag{221}$$

Since this determinant is nonsingular, we have that Eq. 220 holds. This in turn assures that (F_1, G_1) is a controllable pair so the system of Eq. 216 is controllable.

By a similar Kalman rank test on the observability Grammian, for (H, F) to be an observable pair, it must be that

$$\text{rank}\left[H^T \vdots F^T H^T \vdots (F^T)^2 H^T \vdots (F^T)^3 H^T\right] = 4. \tag{222}$$

and since

$$\left[H^T \vdots F^T H^T\right] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{223}$$

has four independent columns, it follows that Eq. 222 does in fact hold. Therefore, the augmented system (F, H) is an observable pair so the system of Eqs. 216 and 217 is observable.

Further,

$$\begin{aligned} (sI - F_1)^{-1} &= \frac{\text{adj}[sI - F_1]}{\det[sI - F_1]} = \frac{\text{adj}[sI - F_1]}{(s^2+3s+2)^2} \\ &= \frac{1}{(s^2+3s+2)^2} \cdot \\ & \quad \begin{bmatrix} (s+3)(s^2+3s+2) & -2(s^2+3s+2) & 0 & 0 \\ (s^2+3s+2) & s(s^2+3s+2) & 0 & 0 \\ 0 & 0 & (s+3)(s^2+3s+2) & -2(s^2+3s+2) \\ 0 & 0 & (s^2+3s+2) & s(s^2+3s+2) \end{bmatrix} \\ &= \frac{1}{s^2+3s+2} \begin{bmatrix} s+3 & -2 & 0 & 0 \\ 1 & s & 0 & 0 \\ 0 & 0 & s+3 & -2 \\ 0 & 0 & 1 & s \end{bmatrix}. \end{aligned} \tag{224}$$

and pre- and post-multiplying the result of Eq. 224 by the observation matrix of Eq. 217 and the noise gain matrix of Eq. 216, respectively, again yields the transfer function of Eq. 199, transposed, as a type of check on the correctness of these calculations.

Now the correlated process noise $u(t)$ of Eq. 191 can be represented dynamically in terms of the above linear system parameters (corresponding explicitly to Eqs. 216 and 217):

$$\dot{u}'(t) = F\, u'(t) + G\, w(t)\ \text{, where } w(t) \sim \mathcal{N}(0, I_2)\ , \tag{225}$$

$$u(t) = H\, u'(t)\ , \tag{226}$$

so the resulting system corresponding to augmenting the linear system of Test Case 2 with the current represention of the time-correlated $u(t)$, as driven by Gaussian white noise, $w(t)$, being

$$\underline{x}(t) = \begin{bmatrix} x(t) \\ \cdots \\ u'(t) \end{bmatrix}, \tag{227}$$

yields the following representation:

$$\begin{aligned}
\frac{d}{dt}\begin{bmatrix} x(t) \\ \cdots \\ u'(t) \end{bmatrix} &= \begin{bmatrix} A & \vdots & BH \\ \cdots & \cdots & \cdots \\ 0 & \vdots & F \end{bmatrix}\begin{bmatrix} x(t) \\ \cdots \\ u'(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \cdots \\ G \end{bmatrix} w(t)\,, \\
&= \begin{bmatrix} -5 & -1 & \vdots & 1 & 0 & 0 & 0 \\ 6 & 0 & \vdots & 0 & 0 & 1 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & 0 & 1 & 0 & 0 \\ 0 & 0 & \vdots & -1 & -3 & 0 & 0 \\ 0 & 0 & \vdots & 0 & 0 & 0 & 1 \\ 0 & 0 & \vdots & 0 & 0 & -2 & -3 \end{bmatrix}\begin{bmatrix} x(t) \\ \cdots \\ u'(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \cdots & \cdots \\ -1 & 0 \\ (6-\sqrt{7})/2 & -1/2 \\ -1 & 0 \\ (6-\sqrt{7})/2 & 3/2 \end{bmatrix}\begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}
\end{aligned} \tag{228}$$

with measurements now represented as

$$\begin{aligned}
z(t) &= [C\vdots 0]\underline{x}(t) + v(t)\,, \text{where } v(t) \sim \mathcal{N}(0, R)\,, \\
&= \begin{bmatrix} 1 & 0 & \vdots & 0 & 0 & 0 & 0 \\ 0 & 1 & \vdots & 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \\ \cdots \\ u_1'(t) \\ u_2'(t) \\ u_3'(t) \\ u_4'(t) \end{bmatrix} + v(t)\,.
\end{aligned} \tag{229}$$

This concludes the handling of serially time correlated process noise within an augmented linear state variable representation that is completely compatible with optimal Kalman filtering by only having effective uncorrelated Gaussian white process and measurement noises present. The drawback or down side is that the computational burden being the cube of the state size now goes as $(6)^3$ rather than just $(2)^3$ for Test Case 2 of Table 1 devoid of time-correlated process noise.

12 Philosophy of Software Structural Validity Confirmation Using Augmented Test Cases of Known Closed-Form Solution Such as These

A difficulty, as discussed in [6, Sec. I], is that most closed-form KF covariance solutions are of either dimension 1 or 2 (as in [15, pp. 138-142, pp. 243-244, p. 246, pp. 255-257, pp. 318-320]) or 3 (as in [10]). To circumvent this dimensional mismatch to higher dimensional real applications that may be hard-wired, we can achieve the dimension n goal by augmenting matrices and vectors with a concatenation of several existing test problems. Use of only totally diagonal decoupled test problems is notorious for being too benign or lenient and not taxing enough to uncover software implementation defects (when the problems exist in the portion of the code that handles cross-term effects). Augmenting either several low-dimensional 2-state problems or fewer 3-state problems is the way to proceed in order to easily obtain a general n-state non-trivial non-diagonal test problem. A confirmation that this proposed augmentation is valid in general is provided next for a closed-form steady-state radar target tracking solution that was successfully used as a check on the software implementation of [16].

An initial worry in adjoining the same 3-state problem with itself relates to whether "controllability and observability" are destroyed, while the 3-state problem by itself does possess the requisite "controllability and observability". "Controllability and observability" conditions, or at least more relaxed but similar "stabilizability and detectability" conditions [17, pp. 62-64, pp. 76-78, pp. 462-465], need to be satisfied in order that the covariance of a KF be well-behaved [15, p. 70, p. 142], [17]. The following mathematical manipulations establish that such an adjoining of two 3-state test problems does not destroy the "controllability and observability" of the resulting 6-state test problem even though it already exists for the 3-state test problem by itself.

First consider the 3-state test problem of [10] of the following form:

$$\overset{(3\times1)}{x} = \begin{bmatrix} \text{position} \\ \text{velocity} \\ \text{acceleration} \end{bmatrix}, \tag{230}$$

with

$$\dot{x} = A_1 x + B_1 \xi, \quad \xi \sim \mathcal{N}(0, Q_1),$$

$$v = C_1 x + \eta, \quad \eta \sim \mathcal{N}(0, R_1),$$

and assumed to be already satisfying Kalman's "controllability and observability" rank test criteria ([15, p. 70]), respectively, as

$$rank[B_1 \vdots A_1 B_1 \vdots A_1^2 B_1] = n_1 = 3, \tag{231}$$

$$rank[C_1^T \vdots A_1^T C_1^T \vdots (A_1^T)^2 C_1^T] \;=\; n_1 = 3. \tag{232}$$

Now the augmented system of the form

$$x = \begin{bmatrix} \text{position} \\ \text{velocity} \\ \text{acceleration} \\ \cdots \\ \text{position} \\ \text{velocity} \\ \text{acceleration} \end{bmatrix}, \tag{233}$$

with

$$\dot{x} = \begin{bmatrix} A_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & A_1 \end{bmatrix} x + \begin{bmatrix} B_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & B_1 \end{bmatrix} \begin{bmatrix} \xi \\ \cdots \\ \xi \end{bmatrix} \tag{234}$$

$$v = \begin{bmatrix} C_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & C_1 \end{bmatrix} x + \begin{bmatrix} 1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 1 \end{bmatrix} \begin{bmatrix} \eta \\ \cdots \\ \eta \end{bmatrix} \tag{235}$$

has system, process noise gain, and observation matrices, respectively, of the form

$$A_2 = \begin{bmatrix} A_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & A_1 \end{bmatrix}, \tag{236}$$

$$B_2 = \begin{bmatrix} B_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & B_1 \end{bmatrix}, \tag{237}$$

$$C_2 = \begin{bmatrix} C_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & C_1 \end{bmatrix}. \tag{238}$$

In testing for controllability of this augmented system, form

$$rank[B_2 \vdots A_2B_2 \vdots A_2^2B_2 \vdots A_2^3B_2 \vdots A_2^4B_2 \vdots A_2^5B_2] =$$

$$\text{rank}\begin{bmatrix} B_1 & \vdots & 0 & \vdots & A_1B_1 & \vdots & 0 & \vdots & A_1^2B_1 & \vdots & 0 & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\text{other stuff} \\ 0 & \vdots & B_1 & \vdots & 0 & \vdots & A_1B_1 & \vdots & 0 & \vdots & A_1^2B_1 & \vdots & \end{bmatrix} =$$

$$\text{rank}\begin{bmatrix} B_1 & \vdots & A_1B_1 & \vdots & A_1^2B_1 & \vdots & 0 & \vdots & 0 & \vdots & 0 & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\text{other stuff} \\ 0 & \vdots & 0 & \vdots & 0 & \vdots & B_1 & \vdots & A_1B_1 & \vdots & A_1^2B_1 & \vdots & \end{bmatrix}$$

$$= 3 + 3 = 6\,. \tag{239}$$

In the next to the last line of Eq. 239, the columns of the Controllability Grammian are rearranged for convenience to provide the necessary insight. Permuting columns of a matrix doesn't alter its rank but can alter at-a-glance conclusions. Since we are able to show that the augmented system rank is 6, this system is confirmed to be controllable. A similar conclusion (on the requisite observability being satisfied) can be obtained by identical steps using the duality that exists between controllability and observability results and the associated forms of arguments or proofs when similar matrix structures, such as are present here, are involved. The above described augmented system of Eqs. 234 and 235 can be used with

$$R_2 = \begin{bmatrix} R_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & R_1 \end{bmatrix}, \tag{240}$$

$$Q_2 = \begin{bmatrix} Q_1 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & Q_1 \end{bmatrix}, \tag{241}$$

$$P_2(0) = \begin{bmatrix} P_1(0) & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & P_1(0) \end{bmatrix}, \tag{242}$$

since now the augmented system has been demonstrated above to be both "observable and controllable" and the measurement noise covariance R_2 of Eq. 240 to be utilized is positive definite. This final observation allows us to use this 6-state augmented test problem with confidence to check out the software implementation as it is currently configured without making any further changes to the software.

The importance of this section is that subsequent software verifiers, when faced with validating newly coded or newly procured Monte-Carlo simulator subroutine software modules and Kalman filters of their own, can treat the entire exercise as one of confirming the proper performance behavior of the new modules merely as an exercise with black boxes. Time can then be saved by just confirming the outputs corresponding to the designated low-dimensional test cases of known closed-form solution provided herein and matching critical intermediate computational benchmarks (without having to necessarily further probe the internal theoretical intricacies that are already justified in [9], in [1] and in [8], where the veracity and utility of these test cases is established and explained in more detail) but can instead check the code, with helpful clues as to the real software culprits and bugs being revealed by these recommended tests when output results don't jibe. Thus, the software verification/validation job is simplified by using the results presented here and used to pinpoint or isolate any problems that exist in the code. This entire exercise of using simple transparent test problems may be interpreted as an initial calibration of the available software before proceeding to use the parameters of the actual application.

By using these or similar examples, certain qualitative and quantitative aspects of the software implementation can be checked for conformance to anticipated behavior as an intermediate benchmark, prior to modular replacement of the various higher-order matrices appropriate to the particular application. This procedure is less expensive in CPU time expenditure during the software debug and checkout phase than using the generally higher n-dimensional matrices of the intended application since the computational burden is generally at least a cubic polynomial in n during the required solution of a Matrix Riccati equation for the associated covariances (also needed to specify the Kalman gain at each time-step). The main contribution of these Test Cases is that one now knows what the answers should be beforehand and is alarmed if resolution is not immediately forthcoming from the software under test. Correct answers could be "hardwired" within candidate software under test, but appropriate scaling [9] of the original test problems to be used as inputs can foil this possible stratagem of such an unscrupulous software supplier/developer.

Another scaling trick that increases the realm of test case possibilities of known closed-form analytic solutions is to use

$$A' = K\,A\,, \tag{243}$$

as the system matrix in Eqs. 1 or 20, where K is some scalar constant selected for convenience and A above in Eq. 243 satisfies the idempotent property of Eq. 24 (e.g., Eqs. 30, 31). Then even though A' doesn't satisfy the idempotent property of Eq. 24 itself, we do have that A' times itself is

$$A'A' = (K\,A)(K\,A) = K^2\,AA = K^2\,A\,, \tag{244}$$

[9]By virtue of the superposition property of linear systems, input $u_A(t)$ that results in known output $y_A(t)$ and input $u_B(t)$ that results in known output $y_B(t)$, should yield known output $[\alpha_1 y_A(t) + \alpha_2 y_B(t)]$ for input $[\alpha_1 u_A(t) + \alpha_2 u_B(t)]$ for any fixed scalars α_1, α_2 (for an infinite set of possibilities that defy hardwiring answers in the software).

and, moreover,

$$(A')^3 = (K\,A)^3 = K^3\,A\,, \tag{245}$$

and the following manipulation corresponding to Eq. 25 is again of closed-form as the known solution:

$$\begin{aligned} e^{A'\Delta} &= \textstyle\sum_{k=0}^{\infty} \frac{A'^k}{k!}\Delta^k \\ &= I + \frac{A'}{1!}\Delta + \frac{A'^2}{2!}\Delta^2 + \frac{A'^3}{3!}\Delta^3 + \cdots \\ &= I + A(\frac{\Delta\,K}{1!} + \frac{\Delta^2\,K^2}{2!} + \frac{\Delta^3\,K^3}{3!} + \cdots) \\ &= I + A(1 + \frac{\Delta\,K}{1!} + \frac{\Delta^2\,K^2}{2!} + \frac{\Delta^3\,K^3}{3!} + \cdots - 1) \\ &= I + A(e^{\Delta\,K} - 1)\,, \end{aligned} \tag{246}$$

where again the last line of Eq. 246 can be explicitly evaluated exactly with simply one scalar multiplication and a matrix addition.

The benefits of using these recommended or similarly justified test cases are the reduced computational expense incurred during software debug by using such low-dimensional test cases and the insight gained into software performance as gauged against test problems of known closed-form solution behavior. However, a modular software design has to be adopted in order to accommodate this approach, so that upon completion of successful verification of the objective computer program implementation with these low-dimensional test problems, the matrices corresponding to the actual application can be conveniently inserted as replacements without perturbing the basic software structure and interactions between subroutines. Even time-critical, real-time applications can be validated in this manner even when using matrix dimensions that are "hardwired" to the particular application by tailoring to the specified dimension using the technique espoused here.

Having validated software that one is confident of is a necessary prerequisite before venturing into mission critical Navigation applications (e.g., [29]-[33]), sonobuoy target tracking [43], radar target tracking [44], and novel research areas such as applications of decentralized Kalman filters [33], [34] or sophisticated algorithms for further post-processing and massaging the outputs of a Kalman filter, as occurs in certain approaches to signal detection in nonstationary systems [6] and in NAV system failure detection [33, Secs. II, III], [35]-[41], and maneuver detection [42]. Explicit analytic closed-form examples or counterexamples are also useful for exposing existing problems or weaknesses in other areas of control and estimation theory (as provided in [2, Secs. I, II, VB, VI], [23]-[27], [28], [33, Sec. II], [39]) so that these unfortunate holes may be shored up in a timely fashion before damage is done in actual applications.

References

[1] Kerr, T. H., "A Constructive Use of Idempotent Matrices to Validate Linear Systems Analysis Software," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-26, No. 6, pp. 935-952, Nov. 1990 (minor corrections in same journal, Vol. AES-27, No. 6, pp. 951-952, Nov. 1991).

[2] Kerr, T. H., "Computational Techniques for the Matrix Pseudoinverse in Minimum Variance Reduced-Order Filtering and Control," in *Control and Dynamic Systems-Advances in Theory and Applications,* Vol. XXVIII: Advances in Algorithms and computational Techniques for Dynamic Control Systems, Part 1 of 3, C. T. Leondes (Ed.), Academic Press, N.Y., 1988.

[3] DeRusso, P. M., Roy, R. J., and Close, C. M., *State Variables for Engineers,* John Wiley and Sons, N.Y., 1967.

[4] Kerr, T. H., "Multichannel Shaping Filter Formulations for Vector Random Process Modeling Using Matrix Spectral Factorization," MIT Lincoln Laboratory Project Report No. PA-500, Lexington, MA, 27 March 1989.

[5] Kerr, T. H., "Emulating Random Process Target Statistics (using MSF)," *IEEE Transactions on Aerospace and Electronic Systems,* to appear in Vol. AES-30, April 1994.

[6] Kerr, T. H., "An Analytic Example of a Schweppe Likelihood Ratio Detector," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-25, No. 4, pp. 545-558, July 1989.

[7] Kerr, T. H., "The Proper Computation of the Matrix Pseudo-Inverse and its Impact in MVRO Filtering," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-21, No. 5, September 1985.

[8] Kerr, T. H., "Rationale for Monte-Carlo Simulator Design to Support Multichannel Spectral Estimation and/or Kalman Filter Performance Testing and Software Validation/Verification Using Closed-Form Test Cases," MIT Lincoln Laboratory Report No. PA-512, Lexington, MA, 22 December 1989.

[9] Kerr, T. H., "Numerical Approximations and Other Structural Issues in Practical Implementations of Kalman Filtering," a chapter in *Approximate Kalman Filtering*, edited by Guanrong Chen, scheduled to appear in summer of 1993.

[10] Ramachandra, K. V., "Optimum Steady State Position, Velocity, and Acceleration Estimation Using Noisy Sampled Position Data," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-23, No. 5, pp. 705-708, September 1987 (a correction appears in 1989 issue).

[11] Healy, M. J. R., *Matrices for Statistics*, Clarendon Press, Oxford, England, 1986.

[12] Burlington, R. S., *Handbook of Mathematical Tables and Formulas*, McGraw-Hill, NY, 1962.

[13] Getson, A. J., and Hsuan, F. C., *{2}-Inverses and Their Statistical Applications*, Springer-Verlag, N.Y., 1988.

[14] Van Trees, H. L., *Detection, Estimation and Modulation Theory, Part III: Radar-Sonar Signal Processing and Gaussian Signals in Noise*, John Wiley and Sons, Inc., N.Y., 1971.

[15] Gelb, A. (Ed.), *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.

[16] *Kalman Filtering Software: User's Guide*, Optimization Software, Inc. Publications Division, N.Y., 1984.

[17] Kwakernaak, H., and Sivan, R., *Linear Optimal Control Systems*, Wiley-Interscience, N. Y., 1972.

[18] Youla, D. C., "On the Factorization of Rational Matrices," *IRE Transactions on Information Theory,* Vol. IT-7, No. 4, pp. 172-189, July 1961.

[19] Davis, M. C., "Factoring the Spectral Matrix," *IEEE Transactions on Automatic Control,* Vol. AC-8, No. 5, pp. 296-305, October 1963.

[20] Boozer, D. D., and McDaniel, W. L., "On Innovation Sequence Testing of the Kalman Filter," *IEEE Transactions on Automatic Control*, Vol. AC-17, No. 1, February 1972.

[21] Martin, W. C., and Stubberud, A. R., "An Additional requirement for Innovations Testing in System Identification," *IEEE Transactions on Automatic Control*, Vol. AC-19, No. 5, October 1974.

[22] Bar-Shalom, Y., and Fortmann, T. E., *Tracking and Data Association*, Academic Press, N. Y., 1988.

[23] Kerr, T. H., "An Invalid Norm Appearing in Control and Estimation," *IEEE Transactions on Automatic Control*, Vol. AC-23, No. 1, pp. 73-74, February 1978 (Correction on pp. 1117-1118, December 1978).

[24] Kerr, T. H., "Three Important Matrix Inequalities Currently Impacting Control and Estimation Applications," *IEEE Transactions on Automatic Control*, Vol. AC-23, No. 6, December 1978.

[25] Kerr, T. H., "Testing Matrices for Definiteness and Application Examples that Spawn the Need," *AIAA Journal of Guidance, Control, and Dynamics,* Vol. 10, No. 5, pp. 503-506, Sept.-Oct., 1987 (reply to and rebuttal by author in Vol. 12, No. 5, p. 767, Sept.-Oct. 1989 issue).

[26] Kerr, T. H., "On Misstatements of the Test for Positive Semidefinite Matrices," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 3, pp. 571-572, May-June 1990.

[27] Kerr, T. H., "Fallacies in Computational Testing of Matrix Positive Definiteness/Semidefiniteness," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-26, No. 2, pp. 415-421, March 1990.

[28] Kerr, T. H., "Status of CR-Like Lower bounds for Nonlinear Filtering," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-25, No. 5, pp. 590-601, September 1989 (Author's reply in Vol. AES-26, No. 5, pp. 896-898, September 1990).

[29] Kerr, T. H., "Modeling and Evaluating an Empirical INS Difference Monitoring Procedure Used to Sequence SSBN Navaid Fixes," *Proceedings of the Annual Meeting of the Institute of Navigation*, U.S. Naval Academy, Annapolis, Md., 9-11 June 1981 (Selected for reprinting in *Navigation: Journal of the Institute of Navigation*, Vol. 28, No. 4, pp. 263-285, Winter 1981-82).

[30] Kerr, T. H., and Chin, L., "A Stable Decentralized Filtering Implementation for JTIDS RelNav," *Proceedings of IEEE Position, Location, and Navigation Symposium (PLANS)*, Atlantic City, NJ, 8-11 December 1980.

[31] Kerr, T. H., "Impact of Navigation Accuracy in Optimized Straight-Line Surveillance/Detection of Undersea Buried Pipe Valves," *Proceedings of National Marine Meeting of the Institute of Navigation (ION)*, Cambridge, MA, 27-29 October 1982.

[32] Kerr, T. H., "Phase III GPS Integration; Volume 1: GPS U.E. Characteristics," Intermetrics Report IR-MA-177, Cambridge, MA, January 1983, for Navair.

[33] Kerr, T. H., "Decentralized Filtering and Redundancy Management for Multisensor Navigation," *IEEE Trans. on Aerospace and Electronic Systems,* Vol. AES-23, No. 1, pp. 83-119, Jan. 1987 (minor corrections appear on p. 412 of May and on p. 599 of July 1987 issues of same journal) [receiving the IEEE AES Barry Carlton award for best paper in AES for 1987].

[34] Kerr, T.H., and Chin, L., "The Theory and Techniques of Discrete-Time Decentralized Filters," in *Advances in the Techniques and Technology in the Application of Nonlinear Filters and Kalman Filters*, edited by C.T. Leondes, NATO Advisory Group for Aerospace Research and Development, AGARDograph, Noordhoff International Publishing, Lieden, 1981.

[35] Kerr, T. H., "Failure Detection Aids for Human Operator Decisions in a Precision Inertial Navigation System Complex," *Proceedings of Symposium on Applications of Decision Theory to Problems of Diagnosis and Repair*, sponsored by the local Dayton Chapter of the American Statistical Association, Fairborn, Ohio, June 1976, pp. 98-127, Keith Womer (Ed.), Wright-Patterson AFB, OH: AFIT TR 76-15, AFIT/EN, Oct. 1976.

[36] Kerr, T. H., "Real-Time Failure Detection: A Static Nonlinear Optimization Problem that Yields a Two Ellipsoid Overlap Test," *Journal of Optimization Theory and Applications*, Vol. 22, No. 4, August 1977.

[37] Kerr, T. H., "Statistical Analysis of a Two Ellipsoid Overlap Test for Real-Time Failure Detection," *IEEE Transactions on Automatic Control*, Vol. AC-25, No. 4, August 1980.

[38] Kerr, T. H., "False Alarm and Correct Detection Probabilities Over a Time Interval for Restricted Classes of Failure Detection Algorithms," *IEEE Transactions on Information Theory,* Vol. IT-28, No. 4, pp. 619-631, July 1982.

[39] Kerr, T. H., "Examining the Controversy Over the Acceptability of SPRT and GLR Techniques and Other Loose Ends in Failure Detection," *Proceedings of the American Control Conference*, San Francisco, CA, 22-24 June 1983.

[40] Kerr, T. H., "A Critique of Several Failure Detection Approaches for Navigation Systems," *IEEE Transactions on Automatic Control*, Vol. AC-34, No. 7, pp. 791-792, July 1989.

[41] Kerr, T. H., "Comments on 'A Chi-Square Test for Fault Detection in Kalman Filters'," *IEEE Transactions on Automatic Control*, Vol. AC-35, No. 11, pp. 1277-1278, November 1990.

[42] Kerr, T. H., "On Duality Between Failure Detection and Radar/Optical Maneuver Detection," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-25, No. 4, pp. 581-583, July 1989.

[43] Kerr, T. H., "Functional and Mathematical Structural Analysis of the Passive Tracking Algorithm (PTA)," Intermetrics Report No. IR-MA-208, Cambridge, MA, 25 May 1983, for NADC.

[44] Kerr, T. H., "Streamlining Measurement Iteration for EKF Target Tracking," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-27, No. 2, pp. 408-420, March 1991 (minor correction appears in Nov. 1991 issue).