III. Conclusion

We have shown for symmetric modular sources the existence of a sequence of block codes B^n of fixed rate $R > R(D)$ whose average distortions $\rho_n(B^n)$ converge to D at a doubly exponential rate in block length n. This generalizes the results of Omura and Shohara [5] to context-dependent fidelity criteria of local span.

References

[1] R. G. Gallager, "Tree encoding for symmetric sources with a distortion measure," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 65–76, Jan. 1974.
[2] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec.*, pt. 4, pp. 142–163, 1959.
[3] T. Berger and W. C. Yu, "Rate-distortion theory for context-dependent fidelity criteria," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 378–384, May 1972.
[4] T. J. Goblick, Jr., "Coding for a discrete information source with a distortion measure," Ph.D. dissertation, Dep. Elec. Eng., M.I.T., Cambridge, 1962.
[5] J. K. Omura and A. Shohara, "On convergence of distortion for block and tree encoding of symmetric sources," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 573–577, July 1973.
[6] F. Jelinek, *Probabilistic Information Theory*. New York: McGraw Hill, 1968.
[7] P. Lancaster, *Theory of Matrices*. New York: Academic, 1969.
[8] T. Kato, *Perturbation Theory for Linear Operators*. New York: Springer-Verlag, 1966.

Estimation and Detection of Signals in Multiplicative Noise

ALAN S. WILLSKY, MEMBER, IEEE

***Abstract*—We consider a class of matrix signal processes that are received in the presence of multiplicative observation noise. By examining the differential version of the observation, we are able to derive finite-dimensional optimal detection-estimation equations that involve a linear filter with gain computed on-line using the incoming observations. An example involving the detection of an actuator failure on a rotating rigid body is considered.**

I. Introduction

Kailath and Duncan [1]–[3] and Lo [4] have obtained solutions for rather general classes of continuous-time detection problems. These solutions require the explicit calculation of the conditional expectation of the signal process given the observations, and this calculation cannot, in general, be reduced to the solution of a finite-dimensional set of ordinary stochastic differential equations (which then presumably could be solved on-line, perhaps with the aid of a digital computer). It is thus of interest, from the point of view of practical implementation, to understand the structure of the continuous-time nonlinear estimation-detection problem in greater detail and, in particular, to uncover classes of problems for which finite-dimensional solutions *are* possible.

The linear-Gaussian case is the best known example of an estimation-detection problem for which we have a finite-dimensional solution. In addition, Lo [4] has pointed out another class for which he obtains finite-dimensional solutions (briefly described in the following). In this correspondence we describe yet another class of problems for which finite-dimensional solutions are possible. The solution here is highly nonlinear in nature and possesses a rather interesting structure.

Recently, a great deal of effort has gone into the analysis of bilinear stochastic systems [4]–[8], and finite-dimensional solutions have been obtained in certain cases. Lo [4] has considered a class of estimation and detection problems involving the injection of nonlinear signal and noise processes into Lie groups via right-invariant equations (see the next section). After the construction of an inverse to the injection mapping, Lo applies the result of Kailath and Duncan to obtain the solution to the detection problem in terms of the conditional mean of the signal process. Again this solution cannot be realized, in general, by a finite-dimensional set of equations; however, by considering a special case (right-invariant bilinear signals with Gaussian driving terms), Lo obtains a finite-dimensional solution consisting of a nonlinear preprocessor (to invert the injection mapping), followed by a Kalman–Bucy filter (with precomputed gains) and the likelihood ratio calculation.

In this correspondence we consider another class of estimation-detection problems on Lie groups for which one obtains finite-dimensional optimal solutions. These problems involve observation noise that enters multiplicatively, and the particular formulation is motivated by an attitude estimation-inertial navigation problem described in Section III. By considering the differential forms of the hypotheses, we obtain equivalent hypotheses that involve bilinear equations that are neither right- nor left-invariant. Following the application of the inverse injection procedure described by Lo [4], we obtain equivalent vector space hypotheses. Analysis of these hypotheses leads to a finite-dimensional solution consisting of a nonlinear preprocessor, followed by a nonlinear filter made up of a Kalman–Bucy filter *with gains that must be computed on-line using incoming measurement values*, followed by the likelihood ratio calculation.

II. Main Result

Let G be an n-dimensional matrix Lie group of $N \times N$ matrices with associated Lie algebra L having $A_1,\cdots,A_n$ as a basis (see [4]–[7] for a discussion of the relevant aspects of Lie theory). Consider the processes $x \in R^k$, $y \in R^n$ satisfying

$$dx(t) = F(t)x(t)\,dt + G(t)\,dw(t) \tag{1}$$

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

where w is an m-dimensional Brownian motion independent of $x(0)$ with $E[dw(t)\,dw'(t)] = Q(t)\,dt$. We inject y into G via a right-invariant bilinear equation [4]–[6]

$$dY(t) = \left[\sum_{i=1}^{n} A_i y_i(t)\right] Y(t). \tag{3}$$

The reason for calling this equation right-invariant comes from the fact that the process $Y(t) = Y(t)D$ also satisfies (3) (here D is an arbitrary matrix). Let v be an n-dimensional Brownian motion, independent of x, with $E[dv(t)\,dv'(t)] = R(t)\,dt$. We inject v into G via a *left-invariant* bilinear stochastic equation

$$dV(t) = V(t)\left[\sum_{i=1}^{n} A_i\,dv_i(t) + \sum_{i=1}^{n}\sum_{j=1}^{n} R_{ij}(t)A_iA_j\,dt\right]. \tag{4}$$

We define two hypotheses on G

$$H_{1G}: M(t) = Y(t)V(t) \tag{5}$$

$$H_{0G}: M(t) = V(t). \tag{6}$$

Manuscript received September 30, 1974; revised January 6, 1975. This work was supported in part by NASA under Grant NGL-22-009-124.
The author is with the Electronic Systems Laboratory and the Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

The problem is to determine the likelihood ratio for these two hypotheses and to display the associated filtering equations that arise. Here Y should be interpreted as the signal process and V as observation noise (see Example 1 in Section III for motivation for this formulation).

Computing the differential forms of the hypotheses, we obtain

$$H_{1G}: dM(t) = \left[\sum_{i=1}^{n} A_i y_i(t)\, dt\right] M(t) + M(t)\left[\sum_{i=1}^{n} A_i\, dv_i(t) + \sum_{i=1}^{n}\sum_{j=1}^{n} R_{ij}(t)A_iA_j\, dt\right] \quad (7)$$

$$H_{0G}: dM(t) = M(t)\left[\sum_{i=1}^{n} A_i\, dv_i(t) + \sum_{i=1}^{n}\sum_{j=1}^{n} R_{ij}(t)A_iA_j\, dt\right]. \quad (8)$$

Note that dM in (7) is *neither left- nor right-invariant* unless G is Abelian, in which case our problem reduces to one considered in [4] for which the optimal estimation-detection system consists of a nonlinear preprocessor followed by a linear filter. If G is not Abelian, the solution is somewhat more complicated but still is finite dimensional in nature. We first invert the injection procedure as described in [4]. Premultiply (7) and (8) by $M^{-1}(t)$. Recalling [5] that $X^{-1}AX \in L$, for all $X \in G$, $A \in L$, we have transformed (7), (8) to equivalent hypotheses on L, which when coordinatized using the basis $A_1,\cdots,A_n$, take the form

$$H_{1L}: dz(t) = H(M(t),t)x(t)\, dt + dv(t) \quad (9)$$

$$H_{0L}: dz(t) = dv(t) \quad (10)$$

where $H(M(t),t)$ is an $n \times k$ matrix that depends on $M(t)$. This matrix can be computed as follows: write

$$M^{-1}(t)A_iM(t) = \sum_{j=1}^{n} \gamma_{ij}(M(t))A_j. \quad (11)$$

Then it is easy to show that

$$H(M(t),t) = \Gamma'(M(t))C(t) \quad (12)$$

where the ijth element of Γ is γ_{ij}. Note that H explicitly depends on M unless G is Abelian. Also, H can be computed with relative ease (see the next section).

Since $M(t)$ is known at time t, (9) represents a "conditionally linear" hypothesis. Thus one has that the optimal estimate of x given (9) can be computed by a Kalman–Bucy filter with optimal gains and covariance *computed on-line using incoming values of M and z* (see [8] for details):

$$d\hat{x}(t \mid t) = F(t)\hat{x}(t \mid t)\, dt + K(t \mid t)[dz(t) - H(M(t),t)\hat{x}(t \mid t)\, dt] \quad (13)$$

$$\dot{P}(t \mid t) = F(t)P(t \mid t) + P(t \mid t)F'(t) + G(t)Q(t)G'(t) - K(t \mid t)R(t)K'(t \mid t) \quad (14)$$

$$K(t \mid t) = P(t \mid t)H'(M(t),t)R^{-1}(t) \quad (15)$$

(it should be noted that an example of a scalar discrete-time system for which the optimal filter requires on-line solution of the Riccati equation is reported in [12]). The likelihood ratio $\mathrm{LR}(t \mid t)$ for hypothesis H_1 over H_0, given observations up to time t, is then given by [1]–[3]

$$\mathrm{LR}\,(t \mid t) = \exp\left\{-\frac{1}{2}\int_0^t \hat{x}'(s \mid s)H'(M(s),s) \cdot R^{-1}(s)H(M(s),s)\hat{x}(s \mid s)\, ds + \int_0^t \hat{x}'(s \mid s)H'(M(s),s)R^{-1}(s)\, dz(s)\right\}. \quad (16)$$

III. Two Examples

Example 1: Consider the Lie group $SO(3)$ of all 3×3 orthogonal matrices with positive determinant. The associated Lie algebra has the basis

$$A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (17)$$

Suppose $x \in R^3$ satisfies one of the two hypotheses

$$H_0: dx(t) = f(t)\, dt + dw(t) \quad (18)$$

and

$$H_1: dx(t) = \xi\, dt + f(t)\, dt + dw(t) \quad (19)$$

where f is deterministic and ξ is a Gaussian random vector independent of $x(0)$.

The x process is injected into $SO(3)$ via (3) with $y = x$ and $n = 3$. If we think of Y as a direction cosine matrix representing the orientation of a rigid body with respect to inertial space, then x represents the angular velocity vector (see [5], [8]). Also, f represents known torques applied to the body, and ξ represents a possible actuator failure (e.g., a jammed reaction jet on a spacecraft). For a more detailed description of this dynamical model, we refer the reader to [5] and [8].

Suppose that the rigid body is equipped with an inertial platform that is to be kept fixed in inertial space [9]. Because of errors in the gyroscopes used to sense rotation, the platform drifts. Thus we actually observe

$$M(t) = Y(t)V(t) \quad (20)$$

where M is the orientation of the body with respect to the platform and V represents platform misalignment with respect to inertial space. One can show [5], [10] that a good model for V is a *left-invariant* process driven by the gyro noise. For simplicity, we assume that the noise is white, in which case V is described by (4) with $n = 3$ (there are no difficulties in taking the noise to be colored).

In this case, one can compute

$$H(M) = M' \quad (21)$$

and the likelihood ratio is

$$\mathrm{LR}(t \mid t) = \mathrm{LR}_1(t \mid t)/\mathrm{LR}_0(t \mid t) \quad (22)$$

where LR_i is computed by substituting $\hat{x}_i(t \mid t)$ into (16). Here $\hat{x}_i(t \mid t)$ is the conditional mean of $x(t)$ given z^t, assuming H_i holds. The equations for $\hat{x}_i$ can be easily derived (see [8]; note that for $i = 1$ we must augment the state with the bias ξ).

Example 2: Consider the Lie algebra of all 2 × 2 matrices with basis

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$
$$A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad A_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{23}$$

Let $x \in R^k$ and $y,v \in R^4$ be as in (1)–(4) with $E[dv(t)\,dv'(t)] = I\,dt$. In this case, (4) becomes

$$dV(t) = V(t)\left[\sum_{i=1}^{4} A_i\,dv_i(t) + (A_1 + A_4)\,dt\right]. \tag{24}$$

Defining the two hypotheses as in (5) and (6), we obtain the optimal estimation-detection equations (12)–(16) where γ_{ij}, the ij element of Γ is given by

$$\gamma_{i1}(M) = (M^{-1}A_iM)_{11} \qquad \gamma_{i2}(M) = (M^{-1}A_iM)_{12}$$
$$\gamma_{i3}(M) = (M^{-1}A_iM)_{21} \qquad \gamma_{i4}(M) = (M^{-1}A_iM)_{22}. \tag{25}$$

For instance,

$$\gamma_{11}(M) = \frac{M_{11}M_{22}}{M_{11}M_{22} - M_{12}M_{21}}. \tag{26}$$

IV. Conclusions

In this correspondence we have considered a class of optimal estimation-detection problems involving multiplicative observation noise. By considering the differential form of the observation process, we have obtained optimal estimation and likelihood ratio equations that are quite interesting in that they are identical to those in the linear-Gaussian case except that the estimation error covariance depends on the observations and thus must be computed on-line. We have noted that these results are potentially useful in the detection of failures or changes in system dynamics. This potentiality was illustrated by examining an actuator failure detection problem associated with rigid body rotation and inertial navigation systems.

Finally, we remark that in this correspondence we have considered only the estimation of the vector space processes x and y. The estimation of the injected process Y is more difficult, and we refer the reader to [5]–[7] and [11] for some results for this problem.

References

[1] T. Kailath, "A general likelihood-ratio formula for random signals in Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 350–361, May 1969.

[2] ——, "A further note on a general likelihood formula for random signals in Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 393–396, July 1970.

[3] T. E. Duncan, "Evaluation of likelihood functions," *Inform. Contr.*, vol. 13, 1968.

[4] J. T. Lo, "Signal detection on lie groups," in *Geometric Methods in System Theory*, D. Q. Mayne and R. W. Brockett, Eds. The Netherlands: Reidel, 1973.

[5] A. S. Willsky, "Dynamical systems defined on groups: structural properties and estimation," Ph.D. dissertation, Dep. Aeronautics and Astronautics, M.I.T., Cambridge, June 1973.

[6] A. S. Willsky and S. I. Marcus, "Estimation for bilinear stochastic systems," Rep. ESL-R-544, M.I.T. Electron. Syst. Lab., M.I.T., Cambridge, May 1974.

[7] S. I. Marcus and A. S. Willsky, "A class of finite dimensional optimal nonlinear filters," presented at the 5th Symp. Nonlinear Estimation and Its Applications, San Diego, Calif., Sept. 1974.

[8] A. S. Willsky, "Estimation and detection of signals in multiplicative noise," Rep. ESL-R-521, M.I.T. Electron Syst. Lab., Cambridge, Oct. 1973.

[9] W. Wrigley, W. Hollister, and W. Denhard, *Gyroscopic Theory, Design, and Instrumentation*. Cambridge, Mass.: M.I.T. Press, 1969.

[10] B. Etkin, *Dynamics of Atmospheric Flight*. New York: Wiley, 1972.

[11] J. T. Lo and A. S. Willsky, "Estimation for rotational processes with one degree of freedom—Part I: Introduction and continuous-time processes," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 10–21, Feb. 1975.

[12] K. J. Åström, *Introduction to Stochastic Control Theory*. New York: Academic, 1970.

System Error Bounds for Lagrange Polynomial Estimation of Band-Limited Functions

JOHN J. KNAB, MEMBER, IEEE

***Abstract*—Several recent articles [1]–[3] have discussed the use of Lagrange polynomials for band-limited signal estimation. We find error bounds when using Lagrange polynomials for interpolation and extrapolation of finite-power band-limited signals if a finite number of regularly spaced noisy samples are used.**

I. Introduction

Suppose $f(t)$ has a $(2N + 1)$st derivative and the samples $f(t_k)$ are given, for $0 \le k \le 2N$. The Lagrange polynomial and error term that pass through the $f(t_k)$ points and used to estimate $f(t)$ for points other than t_k are given by [4]

$$f(t) = \sum_{k=0}^{2N} f(t_k)\frac{w(t)}{w'(t_k)(t - t_k)} + \frac{w(t)}{(2N + 1)!} f^{(2N+1)}(u) \tag{1}$$

where $w(t) \equiv (t - t_0)(t - t_1)\cdots(t - t_{2N})$ and u is some point in the smallest interval containing t and $[t_0,t_{2N}]$. If each sample $f(t_k)$ contains an error ε_k, then the system error $e(t)$ in reconstructing $f(t)$ from the Lagrange polynomial is

$$e(t) = -\sum_{k=0}^{2N} \varepsilon_k \frac{w(t)}{w'(t_k)(t - t_k)} + \frac{w(t)}{(2N + 1)!} f^{(2N+1)}(u). \tag{2}$$

The first term of (2) is called the channel error while the second term of (2) is called the truncation error.

We shall consider as a deterministic signal class those $f(t)$, bounded by M, which can be represented by

$$f(t) = \frac{1}{2\pi}\int_{-\omega_0}^{\omega_0} e^{i\omega t}\,dG(\omega)$$

where $G(\omega)$ is of bounded variation. We call this signal class $B_{\omega_0}(M)$; $f(t)$ is said to be bandlimited to ω_0 if it is in this class. We shall analyze the system error (2) if $f(t)$ is in $B_{\omega_0}(M)$ and the ε_k are deterministic errors with $\eta \equiv \max|\varepsilon_k|$; we call this the deterministic case. We shall also consider a random signal class in which the $f(t)$ are from a wide-sense-stationary process whose autocorrelation function $R(\tau)$ is in $B_{\omega_0}(P)$ with $P = R(0)$. We shall analyze the system error when this random signal class is reconstructed using (1) and the ε_k are random errors of zero mean and variance σ^2 which are uncorrelated with each other and uncorrelated with $f(t_k)$, for all k; we call this the random case.

Throughout our discussion, we shall assume that the samples $f(t_k)$ are uniformly spaced at the following t points: $-NT$, $-(N-1)T,\cdots,0,T,\cdots,NT$. We let $T = \pi(1 - \delta)/\omega_0$, where δ is between zero and unity; if $\delta = 0$, this corresponds to sampling at the Nyquist rate and as δ approaches unity, the time between samples T approaches zero. We also normalize the time: $x \equiv t/T$.

II. Results

If $|x| \le \frac{1}{2}$, we have the case of central interpolation since there are approximately an equal number of samples to the left and right of x. For the deterministic case a bound of the system error

Manuscript received May 11, 1974; revised January 23, 1975.
The author is with the Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio.