ESTIMATION AND DETECTION OF SIGNALS
IN MULTIPLICATIVE NOISE

by

Alan S. Willsky*

Abstract

We define a class of detection-estimation problems on matrix Lie groups
in which the observation noise is multiplicative in nature. By examining
the differential versions of the hypotheses, which are bilinear in
nature, we are able to derive the relevant likelihood ratio formula
and the associated optimal estimation equations for the signal given
the observations and the assumption that the signal is present. These
estimation equations are of interest in their own right, in that they
represent a finite dimensional optimal solution to a nonlinear esti-
mation problem and can be viewed as consisting of a Kalman-Bucy filter
along with the on-line computation of the solution of the associated
Riccati equation, which is driven by the observations. The usefulness
of these results is illustrated via an example concerning the detection
of an actuator failure in a rigid body rotational control system.

I. Introduction

Kailath [1],[2] and Duncan [3] have derived rather general likelihood ratio equations for the detection of signals in additive noise. These equations explicitly involve the optimal least squares estimate of the signal given the observations and the assumption that the signal is present. In general, the optimal signal estimation equations are infinite dimensional in nature, and thus the practical implementations of the estimation-detection equations in the general case necessarily involves suboptimal, finite-dimensional approximations.

Thus it is of interest to find classes of signal and observation processes for which the optimal systems can be realized by finite-dimensional sets of equations. Of course the best known example of this type is the class of signals generated by linear systems driven by white noise and the class of observations that are linear in the signal and involve additive observation noise only. In this case, the optimal signal estimate is generated by a Kalman-Bucy linear filter [4],[5], and the detection equations can be implemented quite easily.

Recently, there have been several papers [6]-[14] that point out that estimation problems for certain (right- or left-invariant) bilinear observation processes can also be handled rather nicely. In this case, the tools of Lie theory [15]-[17] are of value in deriving equations for the optimal estimation system, which consists of a nonlinear preprocessor followed by a linear filter. The extension of these bilinear estimation results to the detection problem was carried out by Lo [13], who obtained finite dimensional estimation-detection results for certain right-invariant bilinear observation processes.

In this paper, we consider a somewhat different class of estimation-detection problems. As in Lo's case [13], our observation and signal processes evolve on certain matrix Lie groups, but in our case the observation noise enters <u>multiplicatively</u>. Such a model was first considered in [10] and [12] in relation to the estimation of the angular velocity of a rigid body. By considering the differential form of the observations, we are led to bilinear equations that differ from those of Lo [13] and those considered in [6]-[9] in a most significant way -- our equations are <u>neither left-nor right-invariant</u> (unless the underlying Lie group is abelian, in which case our results are essentially the same as Lo's). In this case, we cannot use the same trick that was so successful in [6]-[9], [11], and [13], but motivated by the results in [10],[12], we are able to obtain nonlinear finite dimensional optimal estimation-detection equations that are most interesting in that they include a Kalman-Bucy filter whose gain must be computed on-line, using the incoming values of the observation process in the integration of the associated Riccati equation.

In the next section we define several classes of processes on Lie groups, introduce our observation model, and compare it to the right-invariant bilinear model used in [6]-[9],[11], and [13]. Section III contains the derivation of the likelihood ratio for signal detection in multiplicative noise, and we also display the optimal nonlinear signal estimation equations. Several examples are included in Section IV. These include a problem formulation that may prove to be of value in detecting actuator and sensor failures in rigid body - inertial guidance systems.

II. Signal and Observation Processes on Matrix Lie Groups

As in [6]-[14], the basis for our generating random processes on matrix Lie groups is an injection procedure from the Lie algebra associated with the Lie group into the Lie group. This type of operation was first introduced by McKean [18],[19] and later extended by Willsky and Lo [6]-[14]. Let G be an n-dimensional matrix Lie group of N x N matrices with associated matrix Lie algebra L (for the relevant properties of matrix Lie groups, see [14]-[17]). Let $A_1, A_1, \ldots, A_n$ be a basis for L. Suppose we have an n-dimensional stochastic process x satisfying

$$dx(t) = f(x(t),t)dt + G(x(t),t)dw(t); \quad x(0) \text{ given} \tag{1}$$

where w is an m-dimensional Brownian motion process, independent of x(0), with

$$E[dw(t)dw'(t)] = Q(t)dt \tag{2}$$

Following [6],[13],[18]-[19], we inject x into G via the "product integral" in one of two ways:

$$X_1(t) = \bigcap_{s \leq t} \exp\left[\sum_{i=1}^{n} A_i dx_i(s)\right] \tag{3}$$

$$X_2(t) = \bigcap_{s \leq t} \exp\left[\sum_{i=1}^{n} A_i x_i(s)ds\right] \tag{4}$$

For the definition of the product integral and a discussion of the existence and properties of X_1 and X_2, see [13],[14],[19]. We only note that X_1 and X_2 satisfy the stochastic differential equations

$$dX_1(t) = \left\{ \sum_{i=1}^{n} A_i dx_i(t) + \sum_{i=1}^{n} \sum_{j=1}^{n} [G(x(t),t)Q(t)G'(x(t),t)]_{ij} A_i A_j dt \right\} X_1(t) \tag{5}$$

$$dX_2(t) = \left[\sum_{i=1}^{n} A_i x_i(t) dt \right] X_2(t) \tag{6}$$

(here x_i is the ith element of the vector x and D_{ij} denotes the ijth element of the matrix D). In (5) and (6) we see the inherently bilinear nature of these equations. In fact, they define right-invariant bilinear systems. By reversing the order of the products in the discrete approximation to the product integral, we obtain left-invariant bilinear systems (e.g. $dX_2(t) = X_2(t)\left[\sum_{i=1}^{n} A_i x_i(t) dt \right]$).

As discussed in [10],[11],[14] and proven in [13], if we assume that x(0) is known, the processes x, X_1, and X_2 are (almost surely) causally equivalent -- i.e. knowledge of $x^t \overset{\Delta}{=} \{x(s) | 0 \leq s \leq t\}$ is equivalent to knowledge of X_1^t or X_2^t. Intuitively, this is clear, since we can write ($X_1, X_2 \in G$ almost surely, which implies they are invertible a.s.)

$$\sum_{i=1}^{n} A_i dx_i(t) = [dX_1(t)]X_1^{-1}(t) - \sum_{i=1}^{n} \sum_{j=1}^{n} [G(x(t),t)Q(t)G'(x(t),t)]_{ij} A_i A_j dt \tag{7}$$

$$\sum_{i=1}^{n} A_i x_i(t) dt = [dX_2(t)]X_2^{-1}(t) \tag{8}$$

we can recover x from X_1 (assuming we can solve (7)) or X_2 because of the linear independence of the A_i (see [10],[13] for the details).

Thus we have the equivalence of the vector space process x and the Lie group processes X_1 and X_2 in that, given any of them we can construct the others, although the functional relationships among X_1, X_2 and x are, in general, quite complex (see [10], [14]).

We include the above general formulation to indicate the extent of the relationship between Lie group and vector space (Lie algebra) processes. For some further comments on and results for the general formulation, we refer the reader to [6], [13], and [14].

The value of the bijectivity of the algebra-to-group injection procedure is great, especially for the special class of linear-bilinear processes -- a setting in which we can solve detection and some estimation problems. In order to indicate this value, we will review a linear-bilinear problem formulation considered by Lo [13] (see also [6]-[12]). The extension of these techniques to the nonlinear case will be clear, although the general problem does not lead to finite dimensional solutions.

Let x be a k-dimensional process satisfying

$$dx(t) = F(t)x(t)dt + G(t)dw(t) \tag{9}$$

where w is an m-dimensional Brownian motion independent of the normally distributed initial condition with

$$E[x(0)] = 0 \quad E[x(0)x(0)'] = P_0 \tag{10}$$

$$E[w(t)] = 0 \quad E[dw(t)dw'(t)] = Q(t)dt \tag{11}$$

Let C(t) be an n x k matrix of continuous functions. We now write down a pair of hypotheses on the Lie Group G:

$$H_{1G}: \quad Z(t) = \bigcap_{s \leq t} \exp\left\{ \sum_{i=1}^{n} A_i [C(t)x(t)dt + dv(t)]_i \right\} \tag{12}$$

$$H_{0G}: \quad Z(t) = \bigcap_{s \leq t} \exp\left\{ \sum_{i=1}^{n} A_i dv_i(t) \right\} \tag{13}$$

where v is an n-dimensional Brownian motion, independent of w and $x(0)$, with

$$E[v(t)] = 0 \qquad E[dv(t)dv'(t)] = R(t)dt \tag{14}$$

$$R(t) > 0 \tag{15}$$

or, in differential form,

$$H_{1G}: \quad dZ(t) = \left\{ \sum_{i=1}^{n} A_i [C(t)x(t)dt + dv(t)]_i + \sum_{i=1}^{n}\sum_{j=1}^{n} R_{ij}(t)A_i A_j dt \right\} Z(t) \tag{16}$$

$$H_{0G}: \quad dZ(t) = \left\{ \sum_{i=1}^{n} A_i dv_i(t) + \sum_{i=1}^{n}\sum_{j=1}^{n} R_{ij}(t)A_i A_j dt \right\} Z(t) \tag{17}$$

Using the bijective property relating processes on the Lie algebra and the Lie group, we have the completely equivalent hypothese on the Lie algebra L:

$$H_{1L}: \quad dz(t) = C(t)x(t)dt + dv(t) \tag{18}$$

$$H_{0L}: \quad dz(t) = dv(t) \tag{19}$$

Note that if we identify the pair of hypotheses H_{1G} and H_{1L} and the pair H_{0G} and H_{0L}, we have

$$z(t) = \bigcap_{s \leq t} \exp\left\{ \sum_{i=1}^{n} A_i dz_i(t) \right\} \tag{20}$$

For the Lie algebra problem, we have the standard linear-Gaussian estimation-detection result. That is, let [0, T] be the time interval of interest and let C_ℓ be the space of n-dimensional continuous functions on [0, T] that are zero at t = 0. Also, let B_ℓ be the Borel field on C_ℓ under the uniform topology and let μ_1 and μ_0 be the measures induced by y on (C_ℓ, B_ℓ) under H_{1L} and H_{0L}, respectively. We then have [1]-[3] that the likelihood ratio for signal detection is

$$LR = \frac{d\mu_1}{d\mu_0}(z) = \exp\left\{ -\frac{1}{2} \int_0^T \hat{x}'(t|t)C'(t)R^{-1}(t)C(t)\hat{x}(t|t)dt \right.$$
$$\left. + \int_0^T \hat{x}'(t|t)C'(t)R^{-1}(t)dz(t) \right\} \qquad (21)$$

where $\int$ denotes the Ito integral and

$$\hat{x}(t|t) = E[x(t)|z^t, H_{1L}] \qquad (22)$$

This is computed by the following Kalman-Bucy filter [4],[5]:

$$d\hat{x}(t|t) = F(t)\hat{x}(t|t)dt + P(t)C'(t)R^{-1}(t)[dz(t) - C(t)\hat{x}(t|t)dt] \qquad (23)$$

$$\dot{P}(t) = F(t)P(t) + P(t)F'(t) - P(t)C'(t)R^{-1}(t)C(t)P(t) + G(t)Q(t)G'(t) \qquad (24)$$

$$P(0) = P_0 \qquad (25)$$

Using the bijectivity of the injection procedure, we would expect the following, which is in fact proven in [13]: let C_g be the family of continuous matrix-valued functions on [0,T] with values in G that are equal to I at 0, and let B_g be the Borel field on C_g under the uniform

topology. Denoting by ν_1 and ν_0 the measures induced on $(C_g, \mathcal{B}_g)$ by the process Z under the hypotheses H_{1G} and H_{0G}, respectively, we have the following representation of the likelihood ratio

$$LR = \frac{d\nu_1}{d\nu_0}(Z) = \exp\left\{-\frac{1}{2}\int_0^T \bar{x}'(t|t)C'(t)R^{-1}(t)C(t)\bar{x}(t|t)\,dt\right.$$

$$\left. + \int_0^T \bar{x}'(t|t)C'(t)R^{-1}(t)\,dz(t)\right\} \qquad (26)$$

where

$$\bar{x}(t|t) = E[x(t)|z^t, H_{1G}] = \hat{x}(t|t) \qquad \text{a.s.} \qquad (27)$$

Here $\bar{x}(t|t)$ is computed as follows:

$$d\bar{x}(t|t) = F(t)\bar{x}(t|t)\,dt + P(t)C'(t)R^{-1}(t)[dz(t) - C(t)\bar{x}(t|t)\,dt] \qquad (28)$$

where P is given by (26),(27) and we recover dz from Z and dZ from

$$\sum_{i=1}^{n} A_i\,dz_i(t) = [dZ(t)]Z^{-1}(t) - \sum_{i=1}^{n}\sum_{j=1}^{n} R_{ij}(t)A_iA_j\,dt \qquad (29)$$

Since $A_0,\ldots,A_{n-1}$ form a basis for L, we can write

$$B = \sum_{i=1}^{n} B^i A_i \qquad \forall\, B \in L, \; B_i \in \mathbb{R} \qquad (30)$$

(see [13] for details) and thus

$$dz'(t) = \left[\left(\sum_{i=1}^{n} A_i\,dz_i(t)\right)^1, \ldots, \left(\sum_{i=1}^{n} A_i\,dz_i(t)\right)^n\right] \qquad (31)$$

Thus, for this problem, the optimal estimation-detection system consists of a nonlinear preprocessor to recover dz from Z and dZ, followed by a Kalman-Bucy linear filter to compute $\hat{x}(t|t)$. This is followed by a system that takes $\hat{x}(t|t)$ and dz(t) as inputs and computes the LR from (21) or (26). Note that as mentioned earlier, it is crucial in the above development that dZ have the right-invariant representation (16) or (17). In fact, it is precisely this point that leads to the design of the nonlinear preprocessor followed by a linear filter with pre-computable gains. We note that assuming that $\hat{x}(0) \neq 0$ or that the Lie algebra hypotheses are

$$H_{1L}: \quad dz(t) = f(t)dt + C(t)x(t)dt + dv(t) \tag{32}$$

$$H_{0L}: \quad dz(t) = f(t)dt + dv(t) \tag{33}$$

where f is a deterministic term, causes no difficulty in the above analysis nor in the analysis described in the rest of the paper. Such a term can be thought of as a "carrier frequency" (see [6],[8]).

As discussed in [10] and [12], there are several physically important problems, including some inertial navigation and optical communication applications, in which the observation noise process is inherently multiplicative in nature. Based on this physical motivation and the results in [10] and [12], in the next section we formulate a multiplicative noise detection problem. As we shall see, this development will lead to bilinear equations that are neither right- nor left-invariant and for which the optimal detection-estimation system takes a rather striking form. The techniques introduced here are potentially useful in such problems as sensor, actuator, and plant failure

detection in linear and bilinear systems (which will be discussed in

subsequent papers; see also Example 1 in Section IV).

Let x,v, and C be as before, and let

$$y(t) = C(t)x(t) \tag{34}$$

We inject y into G via the usual product integral

$$Y(t) = \bigcap_{s \leq t} \exp\left[\sum_{i=1}^{n} A_i y_i(s) ds\right] \tag{35}$$

$$dY(t) = \left[\sum_{i=1}^{n} A_i y_i(t) dt\right] Y(t) = \left[\sum_{i=1}^{n} A_i (C(t)x(t))_i dt\right] Y(t) \tag{36}$$

We also inject v into G via a second product integral

$$V(t) = \prod_{s \leq t} \exp\left[\sum_{i=1}^{n} A_i dv_i(s)\right] \tag{37}$$

which is to be interpreted as corresponding to the left-invariant

bilinear stochastic equation

$$dV(t) = V(t)\left[\sum_{i=1}^{n} A_i dv_i(t) + \sum_{i=1}^{n} \sum_{j=1}^{n} R_{ij}(t) A_i A_j dt\right] \tag{38}$$

We note that (37) corresponds to reversing the order of the products in

the limiting expression for the product integral (see [10],[12]).

In the next section we will consider a detection problem in-

volving an observation process of the form

$$M(t) = Y(t)V(t) \tag{39}$$

Here Y is the signal process and V should be interpreted as observation

noise. For physical motivation for the models (38) and (39), see [10], [12]. The differential form of (39) is

$$dM(t) = \left[\sum_{i=1}^{n} A_i [C(t)x(t)]_i dt \right] M(t) + M(t) \left[\sum_{i=1}^{n} A_i dv_i(t) \right.$$
$$\left. + \sum_{i=1}^{n} \sum_{j=1}^{n} R_{ij}(t) A_i A_j dt \right] \tag{40}$$

This is <u>neither left-nor right-invariant</u> (in fact, it is the sum of a left-invariant and a right-invariant term), and thus, we <u>cannot</u> use the same nonlinear preprocessing trick that was successful in the previous problem in reducing the problem to a linear-Gaussian one. That is, in general, when we multiply through in (33) by $M^{-1}(t)$, the right-hand side is <u>not independent of M</u>, so we cannot obtain the filter form of nonlinear preprocessor followed by a linear filter with precomputed gains. As we shall see in the next section, the optimum filter-detector is highly nonlinear in nature and possesses a rather distinctive form.

In closing this section, we note that the observation process (40) is of the <u>same</u> form as that in (12) if the underlying Lie group G is abelian [15]. In this case, elements of L and G commute, which implies that right- and left-invariant bilinear systems are the same. Thus, our results will reduce to those of Lo [13] and Willsky and Lo [6]-[8] in the abelian case.

III. Estimation-Detection with Multiplicative Observation Noise

Let Y and V be given by (35)-(38). We define two hypotheses on G

$$H_{1G}: M(t) = Y(t)V(t) \tag{41}$$

$$H_{0G}: M(t) = V(t) \tag{42}$$

or, in differential form,

$$H_{1G}: dM(t) = \left[\sum_{i=1}^{n} A_i Y_i(t) dt \right] M(t) + M(t) \left[\sum_{i=1}^{n} A_i dv_i(t) + \sum_{i=1}^{n} \sum_{j=1}^{n} R_{ij}(t) A_i A_j dt \right] \tag{43}$$

$$H_{0G}: dM(t) = M(t) \left[\sum_{i=1}^{n} A_i dv_i(t) + \sum_{i=1}^{n} \sum_{j=1}^{n} R_{ij}(t) A_i A_j dt \right] \tag{44}$$

The problem is to determine the likelihood ratio for these two hypotheses and to display the associated filtering equations that arise.

As in the right-invariant case discussed in the preceding section, we will find it useful to transform the hypotheses (41),(43) and (42), (44) into completely equivalent hypotheses on the Lie algebra via a particular bijective mapping. We do this as follows: multiply both sides of (43) and (44) on the left by $M^{-1}(t)$ (which exists w.p. 1). Recall [15],[16] that if L is the matrix Lie algebra associated with the matrix Lie group G, then

$$X^{-1}A X \in L \qquad \forall A \in L \qquad \forall X \in G \tag{45}$$

Thus, we have the hypotheses on the Lie algebra L (almost surely):

$$H_{1L}: \quad M^{-1}(t)[dM(t)] - \sum_{i=1}^{n} \sum_{j=1}^{n} R_{ij}(t)A_iA_j dt$$

$$= M^{-1}(t)\left[\sum_{i=1}^{n} A_iy_i(t)dt\right] M(t) + \sum_{i=1}^{n} A_i dv_i(t) \tag{46}$$

$$H_{0L}: \quad M^{-1}(t)[dM(t)] - \sum_{i=1}^{n} \sum_{j=1}^{n} R_{ij}(t)A_iA_j dt = \sum_{i=1}^{n} A_i dv_i(t) \tag{47}$$

which are easily seen to be completely equivalent to H_{1G} and H_{0G}, respectively, since the mapping from M^t to Z^t, where Z is defined by

$$dZ(t) = M^{-1}(t)[dM(t)] - \sum_{i=1}^{n} \sum_{j=1}^{n} R_{ij}(t)A_iA_j dt \tag{48}$$

is seen to be a bijection of the same general type as that used in [13].

To simplify the hypotheses (46),(47), we coordinatize Z using the basis $A_1,\ldots,A_n$. That is, if we write

$$Z(t) = \sum_{i=1}^{n} A_iz_i(t) \tag{49}$$

our Lie algebra hypotheses become

$$H_{1L}: \quad dz(t) = H(M(t),t)x(t)dt + dv(t) \tag{50}$$

$$H_{0L}: \quad dz(t) = dv(t) \tag{51}$$

where $H(M(t),t)$ is an $n \times k$ matrix that depends on $M(t)$ (it is clear that the right-hand side of (46) is linear in $x(t)$, since $y(t) = C(t)x(t)$). This matrix can be computed as follows: write

$$M^{-1}(t)A_iM(t) = \sum_{j=1}^{n} \gamma_{ij}(M(t))A_j \tag{52}$$

Then

$$M^{-1}(t) \left[\sum_{i=1}^{n} A_i y_i(t) \right] M(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij}(M(t)) A_j y_i(t)$$

$$= \sum_{j=1}^{n} \left[\sum_{i=1}^{n} \gamma_{ij}(M(t)) y_i(t) \right] A_j$$

$$= \sum_{j=1}^{n} [H(M(t),t) x(t)]_j A_j \tag{53}$$

where

$$H(M(t),t) = \Gamma'(M(t)) C(t) \tag{54}$$

and the ijth element of Γ is γ_{ij}. Note that if G is _abelian_,

$$M^{-1}(t) A_i M(t) = A_i \qquad \forall i \tag{55}$$

which implies that $\Gamma = I$ and $H = C$, and thus, as discussed at the end

of the preceding section, our hypotheses (50),(51) reduce to the right-

invariant hypotheses (18),(19). However, in the general case H depends

on M(t). We refer the reader to [13] for a mechanization of a procedure

for finding the coordinate functions γ_{ij}. The procedure involves simple

linear algebra and is easy to mechanize, and thus for our purposes, we

assume that we have two black boxes that can take dZ(t) and M(t) as

their respective inputs and produce as outputs dz and H(M(t),t), re-

spectively (see the examples in the next section in which we explicitly

display the details of these black boxes in several specific cases).

By the bijectivity-equivalence arguments discussed earlier, our

detection problem now reduces to considering the two hypotheses (50),

(51) with the knowledge that z^t and M^t are causally equivalent (this

is true because z^t and M^t are, and z is simply a particular coordinati-

zation of Z). Thus, at time t, M(t) and therefore H(M(t),t) are

conditionally known. Intuitively, we then see that (50) represents a

"conditionally" linear observation process. Thus one would expect the

optimal filter for x(t) given z^t to be a Kalman-Bucy filter with the

optimal gains and associated covariance computed on line using incoming

values of M or z! In addition, it would then be natural to expect the

likelihood ratio to have a form very similar to the ones in the previous

section but again incorporating incoming values of M or z.

These intuitive ideas are, in fact, correct, and a discussion

of how this result follows from several results in the literature is

presented in the appendix. Using the equations in the appendix and

noting that, since M(t) is a deterministic function of z^t, we have

$H(M(t),t) = \tilde{H}(z^t,t)$, we can derive the optimal estimation-detection

system depicted in Figure 1. The incoming observation dM(t) is inte-

grated, and the value of M(t), along with the known values of C(t) and

R(t) are used to compute H(M(t),t) and dz(t) as described earlier. The

conditional density for x(t) given z^t is the normal density $N(x;\hat{x}(t|t),$

$P(t|t))$ where both the conditional mean $\hat{x}(t|t)$ and the conditional

covariance $P(t|t)$ are computed on-line using incoming values of M and

z in the integration of the equations:

$$d\hat{x}(t|t) = F(t)\hat{x}(t|t)dt + K(t|t)[dz(t) - H(M(t),t)\hat{x}(t|t)dt] \qquad (56)$$

$$\hat{x}(0|0) = 0 \qquad (57)$$

$$\dot{P}(t|t) = F(t)P(t|t) + P(t|t)F'(t) + G(t)Q(t)G'(t) - K(t|t)R(t)K'(t|t)$$
$$\qquad (58)$$

$\int$ M(t)

COORDINATIZATION

H(M(t),t)

$\dot{P} = FP+PF'+GOG'-KRK'$

$K = PHR^{-1}$

K(t/t)

INVERSE

dM(t)

$M^{-1}(t)$ $M^{-1}(t)dM(t)$ $+$ $\otimes$ Σ dZ(t)   COORDINATIZATION   dz(t) $d\hat{x} = F\hat{x}dt+K\left[dz - H\hat{x}dt\right]$

$\hat{x}(t/t)$

$\displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n}R_{ij}(t)A_iA_j dt$

$da = -\dfrac{1}{2}\hat{x}'H'R^{-1}H\hat{x}dt+\hat{x}'H'R^{-1}dz$ $a(t/t)$ exp $(\cdot)$ LR(t/t)
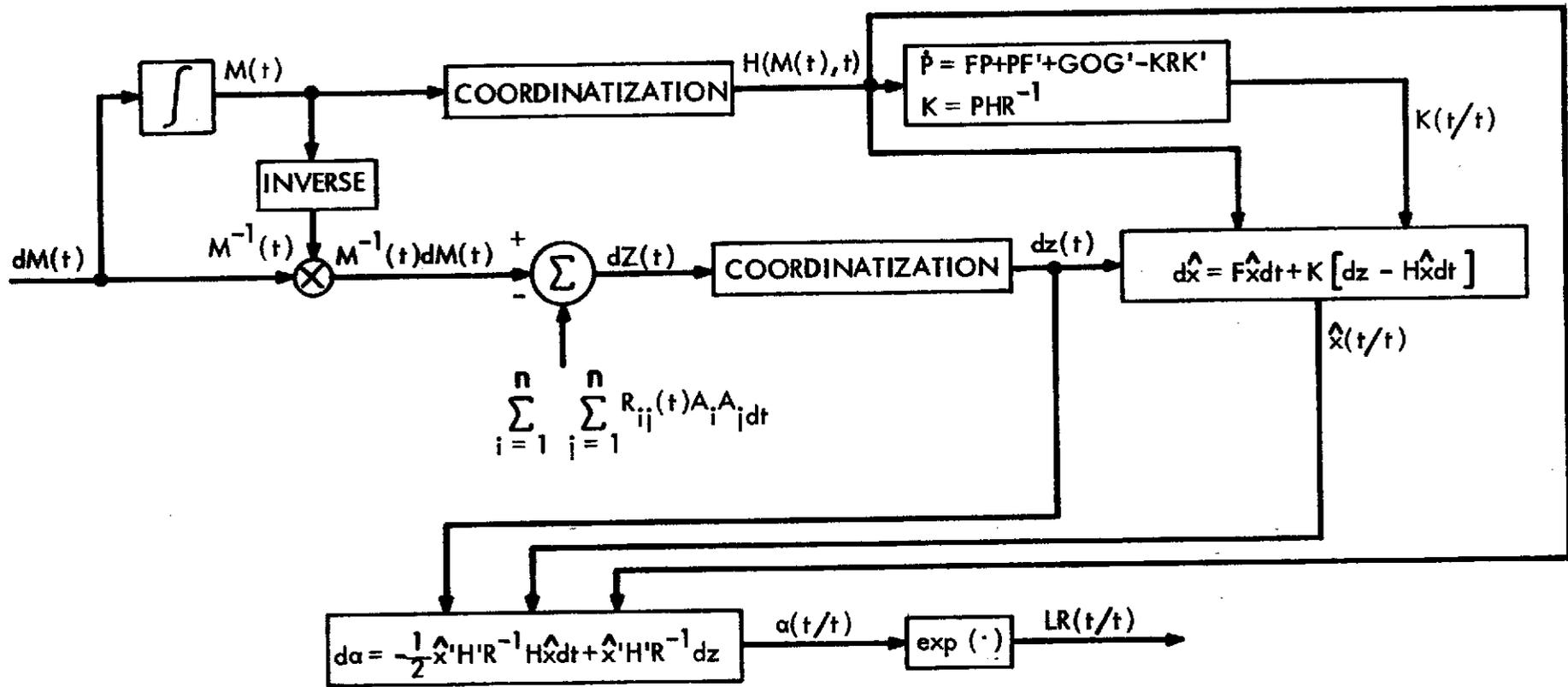
Fig. 1:   Illustrating the Optimal Estimation–Detection System for Multiplicative Observation Noise

$$P(0|0) = P_0 \tag{59}$$

$$K(t|t) = P(t|t)H'(M(t),t)R^{-1}(t) \tag{60}$$

The likelihood ratio $LR(t|t)$ for hypothesis H_1 over H_0, given observations up to time t, is then given by

$$LR(t|t) = \exp\left\{-\frac{1}{2}\int_0^t \hat{x}'(s|s)H'(M(s),s)R^{-1}(s)H(M(s),s)\hat{x}(s|s)\,ds\right.$$

$$\left. + \int_0^t \hat{x}'(s|s)H'(M(s),s)R^{-1}(s)\,dz(s)\right\} \tag{61}$$

(here (H_1,H_0) can be thought of as either (H_{1G},H_{0G}) or (H_{1L},H_{0L}), since they are equivalent).

The optimal estimation-detection system is quite distinctive in form, as it should be viewed as an optimal linear estimation system, augmented by the on-line integration of the nonlinear Riccati equation using the incoming values of the observations, followed by a likelihood ratio evaluation that again is identical in form to the usual linear-Gaussian one, but that also incorporates new values of the observations into the gains. It should also be noted that a simple example of a discrete time system for which the optimal filter has this type of form was reported by Åström [20, p. 236].

We note that as in [1], by using the chain rule one can readily extend these likelihood ratio results to problems such as detection in colored as well as white noise. For instance, we can consider the case in which $Y(t)$ may be generated in one of two hypothesized ways:

$$H_0: \quad dY(t) = \left(\sum_{i=1}^{n} A_i [C_1(t)x(t)]_i dt \right) Y(t) \qquad (62)$$

with x a k_1-vector satisfying

$$dx(t) = F_1(t)x(t)dt + G_1(t)dw(t) \qquad (63)$$

where $F_1, G_1,$ and C_1 are given matrix functions of appropriate dimensions.

The second hypothesis is

$$H_1: \quad dY(t) = \left(\sum_{i=1}^{n} A_i [C_2(t)\xi(t)]_i dt \right) Y(t) \qquad (64)$$

with ξ a k_2-vector satisfying

$$d\xi(t) = F_2(t)\xi(t)dt + G_2(t)dw(t) \qquad (65)$$

where $F_2, G_2,$ and C_2 are also given matrix functions.

In this case the likelihood ratio for H_1 and H_0 given the observation process

$$M(t) = Y(t)V(t) \qquad (66)$$

is obtained as the ratio of the LR for H_1 and H_2 and the LR for H_0 and H_2, where H_2 is the hypothesis

$$H_2: \quad Y(t) \equiv I \qquad (67)$$

Thus, it is easy to see that the system that computes the desired LR consists of two linear filters with on-line gain computations -- i.e. one filter estimating x assuming H_0 holds and one estimating ξ assuming H_1.

It is clear that there are a number of variations on this theme--

e.g. we can hypothesize different V processes, etc. Thus, we see that

the techniques developed in this section are potentially useful in

identifying the underlying dynamics of the system under consideration

and in detecting abrupt changes in the dynamics. Example 1 in the next

section is a simplified version of a very important practical problem,

and it indicates how our results may be applied.

We now make a few comments on optimal Lie group estimation. The

likelihood ratio formula (61) explicitly uses only the optimal (least

squares, maximum likelihood, etc.) estimate of the vector-valued quantity

$x(t)$. Referring to the definitions of y (34) and Y (35),(36), we see

that we can directly compute the optimal estimate of $y(t)$:

$$\hat{y}(t|t) = C(t)\hat{x}(t|t) \tag{68}$$

By identifying $\mathbb{R}^n$ with L via

$$u \longmapsto \sum_{i=1}^{n} A_i u_i \tag{69}$$

we see that we are essentially computing optimal estimates of processes,

such as y, on the Lie algebra. What about the estimation of a Lie

group-valued process such as Y? This is, in general, a very difficult

(in fact, unsolved) problem, since there is no simple relationship

between Y^t and y^t. In fact, the optimal estimate of Y(t) would seem

in general to require smoothing our estimates of the entire trajectory

y^t, and even having this it is not clear what to do! These difficulties

do not arise in the abelian case, which is studied in [6]-[9] and [14],

and explicit solutions can be found in other special cases. A first result is reported in [10] and [11], and further results will be presented in later papers.

Finally, in closing this section, we make a few comments about a slight generalization of the results of this section. So far we have assumed that $A_1, \ldots, A_n$ form a basis for L. Suppose instead we simply assume that $A_1, \ldots, A_n$ are linearly independent and generate L, which we assume is $\rho(>n)$-dimensional. Find $A_{n+1}, \ldots, A_\rho$ so that $A_1 \ldots, A_\rho$ form a basis for L. In this case, we must replace the coordinitization in (49) by

$$z(t) = \sum_{i=1}^{\rho} A_i z_i(t) \tag{70}$$

and the Lie algebra hypotheses (50),(51) become

$$H_{1L}: \quad dz(t) = H(M(t),t)x(t)dt + S(t)dv(t) \tag{71}$$

$$H_{0L}: \quad dz(t) = S(t)dv(t) \tag{72}$$

where $H(M(t),t)$ is now a $\rho \times k$ matrix (computed in precisely the same fashion) and $S(t)$ is the $\rho \times n$ matrix given by

$$S(t) = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{array}{l} \uparrow \\ n \\ \downarrow \\ \uparrow \\ \rho-n \\ \downarrow \end{array} \tag{73}$$
$$|\leftarrow n \rightarrow| \leftarrow \rho-n \rightarrow|$$

In this case (71),(72) include several perfect observations, and the

optimal estimation system becomes an observer-estimator [25],[26]. Thus, there are no conceptual difficulties introduced by this generalization.

IV. Examples

In this section we will present two examples illustrating the techniques developed in the preceding section.

Example 1: Consider the Lie group SO(3), consisting of all 3 x 3 orthogonal matrices with positive determinat. Such a matrix can be thought of as representing the orientation of a rigid body in $\mathbb{R}^3$ -- i.e. it is a "direction cosine" matrix [27] representing the orientation of an object with respect to a (possibly inertial) reference frame. The Lie algebra so(3) associated with SO(3) has the basis

$$A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \qquad A_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (74)$$

Note that SO(3) is nonabelian. For further discussions of the properties and physical significance of SO(3), we refer the reader to [10],[12],[14], [27]-[29].

We now suppose that we have a stochastic process x in $\mathbb{R}^3$ that is given by one of the two hypotheses

$$H_0: dx(t) = f(t)dt + dw(t) \qquad (75)$$

$$H_1: dx(t) = \xi dt + f(t)dt + dw(t) \qquad (76)$$

where f(t) is a deterministic 3-dimensional time function, and ξ is a random (constant) vector with normal distribution

$$E(\xi) = 0 \qquad E(\xi\xi') = P \qquad (77)$$

Also, $x(0)$ is a normally distributed random vector, independent of ξ, with

$$E[x(0)] = 0 \qquad E[x(0)x'(0)] = P_0 \qquad (78)$$

and w is a three-dimensional Brownian motion process, independent of ξ and $x(0)$, with

$$E[w(t)] = 0 \qquad E[w(t)w'(t)] = Q(t)dt \qquad (79)$$

The x process is injected into SO(3) via the equation

$$dX(t) = \left[\sum_{i=1}^{3} A_i x_i(t) dt\right] X(t) \qquad (80)$$

If we think of X as the direction cosine matrix of a rigid body, x has the physical interpretation of being an angular velocity vector, representing the angular velocity of the rigid body with respect to a reference frame (the coordinatization of these quantities depends upon the particular application; see [27] for details). In this case we can interpret physically the two hypotheses: the term $f(t)$ represents known torques that we apply to the body and the Brownian motion term represents random disturbances. The random term ξ represents a possible actuator failure in the control system of the craft -- e.g. a jammed reactor jet on a spacecraft or a failed control surface on an aircraft. Thus, the problem of distinguishing between these hypotheses can be viewed as a failure detection problem.

Before discussing the relevant observation process and associated detection system, we comment on the above dynamical model. Note that the angular velocity equations we have postulated are simpler than the

usual nonlinear Euler equations [27]. Equations (75) and (76) or somewhat more complicated linear equations can be viewed as reasonable approximations if: (1) the rigid body is "nearly" spherically symmetric; or (2) we linearize Euler's equations about a nominal (which might be included in the $f(t)$ term); or (3) we make $Q(t)$ large enough so that the nonlinear effects can be viewed as process noise. Also, there is no difficulty in considering more general linear dynamics -- e.g. if we linearize about a nominal, or if ξ is taken to be time-varying.

We now describe the observation process of interest to us. We assume that X represents the relative orientation of the body with respect to inertial space, and we suppose that the rigid body is equipped with an inertial platform that is to be kept fixed in inertial space (see [27] for a detailed discussion). Because of drifts in the gyroscopes used to sense rotation of the rigid body, the platform drifts relative to inertial space. As discussed in [12] and [14], a possible model for this drift is to take $V(t)$, the orientation of inertial space with respect to the platform, to be a left-invariant Brownian motion

$$dV(t) = V(t) \left\{ \sum_{i=1}^{3} A_i dv_i(t) + \frac{1}{2} \sum_{i,j=1}^{3} R_{ij}(t) A_i A_j dt \right\} \tag{81}$$

where v is a 3-dimensional Brownian motion, independent of $x(0)$, ξ, and w, with

$$E[v(t)] = 0 \qquad\qquad E[v(t)v'(t)] = R(t) dt \tag{82}$$

$$R(t) > 0 \tag{83}$$

Our observation process in the orientation M(t) of the rigid body with respect to the platform, which can be determined by reading off gimbal angles and is given by

$$M(t) = X(t)V(t) \tag{84}$$

As discussed in the preceding section, the incremental change in M(t) is given by

$$dM(t) = \left\{ \sum_{i=1}^{3} A_i x_i(t) \right\} M(t) dt$$

$$+ M(t) \left\{ \sum_{i=1}^{3} A_i dv_i(t) + \frac{1}{2} \sum_{i,j=1}^{3} R_{ij}(t) A_i A_j dt \right\} \tag{85}$$

(see [12],[14], and [27] for a discussion of how one obtains pulse-like or incremental information in such systems).

Performing the type of transformation used in the previous section (note that $M^{-1}(t) = M'(t)$ a.s.), we have

$$dZ(t) = M'(t) dM(t) - \frac{1}{2} \sum_{i,j=1}^{3} R_{ij}(t) A_i A_j dt$$

$$= M'(t) \sum_{i=1}^{3} A_i x_i(t) M(t) dt + \sum_{i=1}^{3} A_i dv_i(t) \tag{86}$$

Also, we obtain the following expression for z(t), the $\mathbb{R}^3$-coordinatization of Z, and its differential:

$$z'(t) = [Z_{32}(t), Z_{13}(t), Z_{21}(t)] \tag{87}$$

$$dz(t) = H(M(t)) x(t) dt + dv(t) \tag{88}$$

where

$$H(M) = \begin{bmatrix} M_{22}M_{33} - M_{23}M_{32} & M_{13}M_{32} - M_{12}M_{33} & M_{12}M_{23} - M_{13}M_{22} \\ M_{23}M_{31} - M_{21}M_{33} & M_{11}M_{33} - M_{13}M_{31} & M_{13}M_{21} - M_{11}M_{23} \\ M_{21}M_{32} - M_{22}M_{31} & M_{12}M_{31} - M_{11}M_{32} & M_{11}M_{22} - M_{12}M_{21} \end{bmatrix} \quad (89)$$

Having these expressions, we have the following equation for the likelihood ratio for the two hypotheses:

$$LR(t|t) = \frac{\exp\left\{-\frac{1}{2}\int_0^t \hat{x}_1{}'(s|s)H'(M(s))R^{-1}(s)H(M(s))\hat{x}_1(s|s)\,ds \right.}{\exp\left\{-\frac{1}{2}\int_0^t \hat{x}_0{}'(s|s)H'(M(s))R^{-1}(s)H(M(s))\hat{x}_0(s|s)\,ds \right.}$$

$$\frac{\left. +\int_0^t \hat{x}_1{}'(s|s)H'(M(s))R^{-1}(s)\,dz(s)\right\}}{\left. +\int_0^t \hat{x}_0{}'(s|s)H'(M(s))R^{-1}(s)\,dz(s)\right\}} \qquad (90)$$

where $\hat{x}_1(t|t)$ is the conditional mean of $x(t)$ given z^t, assuming H_1 holds, while to compute $\hat{x}_0(t|t)$, we assume H_0 holds. The stochastic differential equations for these quantities are

$$d\hat{x}_0(t|t) = f(t)dt + K_0(t|t)[dz(t) - H(M(t))\hat{x}_0(t|t)dt] \qquad (91)$$

$$\dot{P}_0(t|t) = Q(t) - K_0(t|t)R(t)K_0{}'(t|t) \qquad (92)$$

$$K_0(t|t) = P_0(t|t)H'(M(t))R^{-1}(t) \qquad (93)$$

$$\begin{bmatrix} d\hat{x}_1(t|t) \\ d\hat{\xi}(t|t) \end{bmatrix} = \begin{bmatrix} f(t)+\hat{\xi}(t|t) \\ 0 \end{bmatrix} dt + K_1(t|t)[dz(t) - H(M(t))\hat{x}_0(t|t)dt] \qquad (94)$$

$$\dot{P}_1(t|t) = \tilde{F}P_1(t|t) + P_1(t|t)\tilde{F}'(t) + Q(t) + K_1(t|t)R(t)K_1'(t|t)$$

(95)

$$K_1(t|t) = P_1(t|t)\tilde{H}(M(t))R^{-1}(t)$$

(96)

Here P_1 is a 6 x 6 matrix, and

$$\tilde{F} = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \qquad \tilde{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \qquad \tilde{H} = [H \quad 0]$$

(97)

We also note that sensor failure detection can be considered by hypothesizing several different forms for V.

Example 2: Consider GL(2,$\mathbb{R}$), the group of 2 x 2 invertible matrices. Its Lie algebra consists of all 2 x 2 matrices and has the basis

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

(98)

Let x be the k-dimensional process satisfying

$$dx(t) = F(t)x(t)dt + G(t)dw(t)$$

(99)

and let y be the 4-dimensional process

$$y(t) = C(t)x(t)$$

(100)

We also take v to be a 4-dimensional Brownian motion independent of w with

$$E(dv(t)dv'(t)) = Idt$$

(101)

We inject y and v into GL(2,$\mathbb{R}$) via

$$dY(t) = \left(\sum_{i=1}^{4} A_i y_i(t) \right) Y(t)dt$$

(102)

$$dV(t) = V(t) \left[\sum_{i=1}^{4} A_i dv_i(t) + (A_1+A_4)dt \right] \qquad (103)$$

and define the two hypotheses

$$H_1: M(t) = Y(t)V(t) \qquad (104)$$

$$H_0: M(t) = V(t) \qquad (105)$$

As discussed in Section III, we can define Z via

$$dZ(t) = M^{-1}(t)dM(t) - (A_1+A_4)dt \qquad (106)$$

and, defining the 4-vector

$$z'(t) = [Z_{11}(t), Z_{12}(t), Z_{21}(t), Z_{22}(t)] \qquad (107)$$

the two hypotheses become

$$H_1: dz(t) = H(M(t),t)x(t)dt + dv(t) \qquad (108)$$

$$H_0: dz(t) = dv(t) \qquad (109)$$

where we compute $H(M(t),t)$ from

$$H(M(t),t) = \Gamma'(M(t))C(t) \qquad (110)$$

where γ_{ij}, the ij element of Γ, is given by

$$\gamma_{i1}(M) = (M^{-1}A_iM)_{11} \qquad \gamma_{i2}(M) = (M^{-1}A_iM)_{12}$$

$$\qquad (111)$$

$$\gamma_{i3}(M) = (M^{-1}A_iM)_{21} \qquad \gamma_{i4}(M) = (M^{-1}A_iM)_{22}$$

For instance,

$$\gamma_{11}(M) = \frac{M_{11}M_{22}}{M_{11}M_{22} - M_{12}M_{21}} \qquad (112)$$

Having these terms, we can apply the results of Section III to obtain explicit optimal estimation and likelihood ratio equations.

V. Conclusions

In this paper we have considered a class of optimal estimation-detection problems involving multiplicative observation noise. By considering the differential form of the observation process, we 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 for on-line system identification and 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 rotations and inertial navigation systems.

APPENDIX: The Computation of a Likelihood Ratio and a Conditional Density

In Section III we were confronted with a signal detection problem of the form

$$H_1: \quad dz(t) = H(z^t,t)x(t)dt + dv(t) \tag{113}$$

$$H_2: \quad dz(t) = dv(t) \tag{114}$$

where v is an n-dimensional Brownian motion

$$E[v(t)] = 0 \qquad E[dv(t)dv'(t)] = R(t)dt \tag{115}$$

$$R(t) > 0 \tag{116}$$

and x is a k-dimensional process satisfying

$$dx(t) = F(t)x(t)dt + G(t)dw(t) \tag{117}$$

Here x(0) is assumed to be a Gaussian random variable with

$$E[x(0)] = \hat{x}_0 \qquad E[(x(0) - \hat{x}_0)(x(0) - \hat{x}_0)'] = P_0 \tag{118}$$

and w is an m-dimensional Brownian motion with

$$E[w(t)] = 0 \qquad E[dw(t)dw'(t)] = Q(t)dt \tag{119}$$

It is assumed that v,w, and x(0) are mutually completely independent. Also, we note that H is allowed to be a function of the past observations z^t, and we define the "signal" process

$$s(t) = H(z^t,t)x(t) \tag{120}$$

We now note that future values of $v(\cdot)$ are independent of past values of $z(\cdot)$ and $s(\cdot)$. Also, for the particular case of interest in Section III, it can be shown by a tedious but straightforward calculation that if [0,T] is the time interval of interest,

-29-

$$\int_0^T Es^2(t)\,dt < \infty \qquad (121)$$

These facts enable us to use the vector version of the likelihood

ratio formula derived in [2]. Let C_ℓ be the space of n-dimensional

continuous functions on [0,T] with associated Borel field B_ℓ under the

uniform topology. Letting μ_1 and μ_0 denote the measures induced on

(C_ℓ, B_ℓ) by z under H_1 and H_0, respectively, we have

$$LR = \frac{d\mu_1}{d\mu_0}(z) = \exp\left\{-\frac{1}{2}\int_0^T \hat{s}'(t|t)R^{-1}(t)\hat{s}(t|t)\,dt + \int_0^T \hat{s}'(t|t)\bar{R}^{-1}(t)\,dz(t)\right\}$$

$$(122)$$

where

$$\hat{s}(t|t) = E[s(t)|z^t, H_1] = H(z^t,t)E[x(t)|z^t, H_1] = H(z^t,t)\hat{x}(t|t) \qquad (123)$$

Thus, it remains to derive a method for computing $\hat{x}(t|t)$. We

first note that for any $0 \le t_1 < t_2 < .. < t_r \le t$, the variables

$z_{t_1},\ldots,z_{t_r}$ and x(t) are not jointly Gaussian. However, as we shall

see, the conditional density for x(t) given z^t is Gaussian with mean

and covariance that depends on z^t! To see this, we refer to the work of

Kailath [21],[22], and Frost and Kailath [23] on the innovations

approach to least squares estimation. In particular, in [23] a partial

differential equation for the conditional density is derived. The

derivation assumes the complete independence of $v(\cdot)$ and $s(\cdot)$ which we

do not have in our case. However, as Frost and Kailath [23] suggest,

if we use the weaker innovations representation of Fujisaki, Kallianpur,

and Kunita [24], which requires only that future values of $v(\cdot)$ be independent of past $s(\cdot)$ and $v(\cdot)$ plus the integrability condition (121) (actually, a weaker condition will do), we can obtain a partial differential equation of essentially the same form. That is, if we let $\rho(x,t)$ denote the conditional density for $x(t)$ given z^t evaluated at x, we have

$$d\rho(x,t) = L(\rho)(x,t)dt$$
$$+ [x-\hat{x}(t|t)]'H'(z^t,t)R^{-1}(t)[dz(t)-H(z^t,t)\hat{x}(t|t)dt]\rho(x,t) \tag{124}$$

where L is the Fokker-Planek operator for (117)

$$L(\rho)(x,t) = -\rho(x,t)\,\mathrm{tr}\,F(t) - \left[\frac{\partial\rho}{\partial x}(x,t)\right]'F(t)x$$
$$+ \frac{1}{2}\,\mathrm{tr}\left[G(t)Q(t)G'(t)\frac{\partial^2\rho}{\partial x^2}(x,t)\right] \tag{125}$$

A straightforward computation shows that the solution to (124) is

$$\rho(x,t) = N(x;\hat{x}(t|t),\,P(t|t)) \tag{126}$$

where $N(x;\alpha,P)$ is the (multi-dimensional) normal density with mean α and covariance P, and we compute $\hat{x}(t|t)$ and $P(t|t)$ from

$$d\hat{x}(t|t) = F(t)\hat{x}(t|t)dt + P(t|t)H'(z^t,t)R^{-1}(t)[dz(t)-H(z^t,t)\hat{x}(t|t)dt \tag{127}$$

$$\hat{x}(0|0) = \hat{x}_0 \tag{128}$$

$$\dot{P}(t|t) = F(t)P(t|t) + P(t|t)F'(t) + G(t)Q(t)G'(t)$$
$$- P(t|t)H'(z^t,t)R^{-1}(t)H(z^t,t)P(t|t) \tag{129}$$

$$P(0|0) = P_0 \qquad\qquad (130)$$

Note that P depends on z^t.

We also note that one can compute the infinite set of conditional moments of $x(t)$ directly from the stochastic differential equations derived in [24], and one finds that the moments of $N(x;\hat{x}(t|t),P(t|t))$ satisfy these equations.

References

1. T. Kailath, "A General Likelihood-Ratio Formula for Random Signals in Gaussian Noise," IEEE Trans. on Information Theory, Vol. IT-15, No. 3, May 1969, pp. 350-361.

2. T. Kailath, "A Further Note on a General Likelihood Formula for Random Signals in Gaussian Noise," IEEE Trans. on Inf.Th., Vol. IT-16, No. 4, 1970, pp. 393-396.

3. T.E. Duncan, "Evaluation of Likelihood Functions," Information and Control, Vol. 13, 1968.

4. R.E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," Trans. ASME, Ser. D, J. Basic Eng., Vol. 82, 1960, pp. 35-45.

5. R.E. Kalman and R.S. Bucy, "New Results in Linear Filtering and Prediction Theory," Trans. ASME, Ser. D, J. Basic Eng., 1961, pp. 95-108.

6. J.T. Lo and A.S. Willsky, "Estimation for Rotational Processes with One Degree of Freedom I: Introduction and Continuous Time Processes," submitted to IEEE Trans. on Aut. Cont.

7. A.S. Willsky and J.T. Lo, "Estimation for Rotational Processes with One Degree of Freedom II: Discrete Time Processes," submitted to IEEE Trans. on Aut. Cont.

8. A.S. Willsky and J.T. Lo, "Estimation for Rotational Processes with One Degree of Freedom III: Applications and Implementation," submitted to IEEE Trans. on Aut. Cont.

9. J.T. Lo and A.S. Willsky, "Stochastic Control of Rotational Processes with One Degree of Freedom," submitted to SIAM J. Control.

10. A.S. Willsky, "Some Estimation Problems on Lie Groups," Proc. of the NATO Advanced Study Institute on Geometric and Algebraic Methods for Nonlinear Systems, Imperial College, London, England, August 27-Sept. 7, 1973.

11. A.S. Willsky, "A Representation and Estimation Result for a Certain Class of Random Bilinear Systems," to be submitted for publication.

12. A.S. Willsky, "Some Results on the Estimation of the Angular Velocity and Orientation of a Rigid Body," to be submitted for publication.

13. J.T. Lo, "Signal Detection on Lie Groups," Proc. of the NATO Advanced Study Institute on Geometric and Algebraic Methods for Nonlinear Systems, Imperial College, London, England, August 27-Sept. 7, 1973; also submitted to the IEEE Trans. on Inf. Th.

14. A.S. Willsky, _Dynamical Systems Defined on Groups: Structural Properties and Estimation_, Ph.D. Thesis, Dept. of Aeronautics and Astronautics, M.I.T., Cambridge, Mass., June 1973.

15. C. Chevalley, _Theory of Lie Groups_, Princeton University Press, Princeton, 1946.

16. H. Samelson, _Notes on Lie Algebras_, Van Nostrand Reinhold Co., New York, 1969.

17. R.W. Brockett,"System Theory on Group Manifolds and Coset Spaces," _SIAM J. Control_, Vol. 10, No. 2, May 1972, pp. 265-284.

18. H.P. McKean, Jr., "Brownian Motions on the 3-Dimensional Rotation Group," _Mem. Coll. Sci. Kyoto Univ._, Vol. 33, 1960, pp. 25-38.

19. H.P. McKean, Jr., _Stochastic Integrals_, Academic Press, N.Y., 1969.

20. K.J. Aström, _Introduction to Stochastic Control Theory_, Academic Press, New York, 1970.

21. T. Kailath, "An Innovations Approach to Least-Squares Estimation, Part I: Linear Filtering in Additive White Noise," _IEEE Trans. Aut. Control_, Vol. AC-13, 1968, pp. 646-655.

22. T. Kailath, "A Note on Least Squares Estimation by the Innovations Method," _SIAM J. Control_, Vol. 10, No. 3, Aug. 1972, pp. 477-486.

23. P. Frost and T. Kailath, "An Innovations Approach to Least-Squares Estimation, Part III: Nonlinear Least-Squares Estimation," _IEEE Trans. Aut. Control_, Vol. AC-16, 1971, pp. 217-226.

24. M. Fujisaki, G. Kallianpur, H. Kunita, "Stochastic Differential Equations for the Nonlinear Filtering Problem," _Osaka J. Math_, Vol. 9, 1972, pp. 19-40.

25. E. Tse, and M. Athans, "Observer Theory for Continuous-Time Linear Systems," _Information and Control_, Vol. 22, 1973, pp. 405-434.

26. B.J. Uttam and W.F. O'Halloran,Jr., "On Observers and Reduced-Order Optimal Filters for Linear Stochastic Systems," _Proc. of the 1972 JACC_, California, August 1972.

27. W. Wrigley, W. Hollister, and W. Denhard, _Gyroscopic Theory, Design, and Instrumentation_, The M.I.T. Press, Cambridge, Mass., 1969.

28. B.W. Stuck, "A New Method for Attitude Estimation," to appear.

29. J.T. Lo, "Signal Detection of Rotational Processes and Frequency Demodulation," presented at the 7th Annual Princeton Conference on Information Sciences and Systems, March 1973.