SOME EXAMPLES OF LIE ALGEBRAIC STRUCTURE IN NONLINEAR ESTIMATION

Michiel Hazewinkel
Econometric Institute
Erasmus University
Rotterdam, The Netherlands

Chan-Huan Liu

Steven I. Marcus
Department of Electrical Engineering
The University of Texas
Austin, Texas

ABSTRACT

Several examples of the application of Lie algebraic techniques to nonlinear estimation problems are presented. The method relates the computation of the (unnormalized) conditional density and the computation of statistics with finite dimensional estimators to the structure of a certain Lie algebra $L(\Sigma)$. The general method is explained; for a particular example, the structures of the Lie algebras associated with the unnormalized conditional density equation and the finite dimensionally computable conditional moment equations are analyzed in detail. Two general classes of examples are also discussed, and the implications of their Lie algebraic structures are explored.

I. INTRODUCTION

This paper is motivated by the problem of recursively filtering the state x_t of a nonlinear stochastic system, given the past observations $z^t = \{z_s,\ 0 \leq s \leq t\}$. The systems we consider satisfy the Ito stochastic differential equations

$$\begin{aligned} dx_t &= f(x_t)dt + G(x_t)dw_t \\ dz_t &= h(x_t)dt + R_t^{\frac{1}{2}} dv_t \end{aligned} \qquad (\Sigma)$$

where w and v are independent unit variance vector Wiener processes, f and h are vector-valued functions, G is a matrix-valued function, and $R > 0$. The optimal (minimum-variance) estimate of x_t is of course the conditional mean $\hat{x}_t \triangleq E[x_t|z^t]$ (also denoted $\hat{x}_{t|t}$ or $E^t[x_t]$); $\hat{x}_t$ satisfies the (Ito) stochastic differential equation [1]-[3]

$$d\hat{x}_t = \left[\hat{f}(x_t) - (\widehat{x_t h^T - \hat{x}_t \hat{h}^T}) R^{-1}(t)\hat{h}\right]dt + (\widehat{x_t h^T - \hat{x}_t \hat{h}^T}) R^{-1}(t)dz_t \quad (1)$$

This research was supported in part by the Air Force Office of Scientific Research (AFSC) under grant AFOSR-79-0025, in part by the National Science Foundation under grant ENG 76-11106, and in part by the Joint Services Electronics Program under Contract F49620-77-C-0101. A portion of this research was conducted at the Econometric Institute, Erasmus University, Rotterdam, The Netherlands.

where ^ denotes conditional expectation given z^t and h denotes $h(x_t)$. Also, the conditional probability density $p(t,x)$ of x_t given z^t (we will assume that $p(t,x)$ exists) satisfies the stochastic partial differential equation [3],[4]

$$dp(t,x) = Lp(t,x)dt + (h(x)-\hat{h}(x))^T R^{-1}(t)(dz_t - \hat{h}(x)dt)p(t,x) \quad (2)$$

where

$$L(\cdot) = -\sum_{i=1}^{n} \frac{\partial(\cdot f_i)}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2(\cdot(GG^T)ij)}{\partial x_i \partial x_j} \quad (3)$$

is the forward diffusion operator.

Notice that the differential equation (1) is not recursive, and indeed appears to involve an infinite dimensional computation in general. Aside from the linear-Gaussian case in which the Kalman filter is optimal, there are very few known cases in which the optimal estimator is finite dimensional (a number of these are summarized in [5]). In [6]-[8] it is shown that, for certain classes of nonlinear stochastic systems, the conditional mean (and all conditional moments) of x_t given z^t can be computed with recursive filters of finite dimension.

Brockett [9],[10] and Mitter [11] have recently shown that Lie algebras play an important role in nonlinear estimation theory; the perspective of Brockett [9] is the following (we assume for simplicity that z is a scalar). Instead of studying the equation (2) for the conditional density, we consider the Zakai equation for an unnormalized conditional density $\rho(t,x)$ [12]:

$$d\rho(t,x) = L\rho(t,x)dt + h(x)\rho(t,x)dz_t \quad (4)$$

where $\rho(t,x)$ is related to $p(t,x)$ by the normalization

$$p(t,x) = \rho(t,x)\cdot\left(\int\rho(t,x)dx\right)^{-1} \quad (5)$$

The Zakai equation (4) is much simpler than (2); indeed, (4) is a bilinear differential equation [13] in ρ, with z considered as the input. This is the first clue that the Lie algebraic and differential geometric techniques developed for finite dimensional systems of this type may be brought to bear here. Suppose that some statistic of the conditional distribution of x_t given z^t can be calculated with a finite dimensional recursive estimator of the form

$$d\eta_t = a(\eta_t)dt + b(\eta_t)dz_t \tag{6}$$

$$E[c(x_t)|z^t] = \gamma(\eta_t) \tag{7}$$

where η evolves on a finite dimensional manifold, and a, b and γ are analytic. Of course, this statistic can also be obtained from $\rho(t,x)$ by

$$E[c(x_t)|z^t] = \int c(x)\rho(t,x)dx(\int\rho(t,x)dx)^{-1} \tag{8}$$

For the rest of the development, it is more convenient to write (4) and (6) in Fisk-Stratonovich form (so that they obey the ordinary rules of calculus and so that Lie-algebraic calculations involving differential operators can be performed as usual):

$$d\eta_t = \tilde{a}(\eta_t)dt + b(\eta_t)dz_t \tag{9}$$

$$d\rho(t,x) = [L - \frac{1}{2}h^2(x)]\rho(t,x)dt + h(x)\rho(t,x)dz_t \tag{10}$$

where the i^{th} component

$$\tilde{a}_i(\eta) = a_i(\eta) - \frac{1}{2}\sum_j b_j(\eta)\frac{\partial b_i}{\partial \eta_j}(\eta).$$

The two systems (9),(7) and (10),(8) are thus two representations of the same mapping from "input" functions z to "outputs" $E[c(x_t)|z^t]$: (10),(8) via a bilinear infinite dimensional state equation, and (9),(7) via a nonlinear finite dimensional state equation. Generalizing the results of [14],[15] to infinite dimensional state equations, the major result of [9] is that, under appropriate hypotheses, the Lie algebra F generated by $\tilde{a}$ and b (under the commutator $[\tilde{a},b] = \frac{\partial b}{\partial \eta}\tilde{a} - \frac{\partial \tilde{a}}{\partial \eta}b$) is a homomorphic image (quotient) of the Lie algebra $L(\Sigma)$ generated by $A_0 = L - \frac{1}{2}h^2(x)$ and $B_0 = h(x)$ (under the commutator $[A_0,B_0] = A_0B_0 - B_0A_0$). Conversely, any homomorphism of $L(\Sigma)$ onto a Lie algebra generated by two complete vector fields on a finite dimensional manifold allows the computation of some information about the conditional density with a finite dimensional estimator of the form (9). It is not known in what generality such results are valid, especially for cases in which $L(\Sigma)$ is infinite dimensional, and much work remains to be done. However, it is clear that there is a strong relationship between $L(\Sigma)$ and the existence of finite dimensional filters. In this paper, we discuss the properties of $L(\Sigma)$ for some interesting classes of examples. These Lie algebraic calculations give some new insights into certain nonlinear estimation problems and some guidance in the search for finite dimensional estimators; however, to actually construct the finite dimensional filters, one must usually use other, more direct, methods (see, e.g., [6]-[8],[16]).

In [9], this approach is explicitly carried out and analyzed for the problem in which f, G and h are all linear. In that case, the Lie algebra $L(\Sigma)$ of the Zakai equation is finite dimensional and the unnormalized conditional density can in fact be computed with a finite dimensional estimator, the Kalman filter. In this paper, we first carry out a similar analysis for the simplest example of the class considered in [6]-[8]. For this example, all conditional moments of the state can be computed wit finite dimensional filters; the Lie algebra $L(\Sigma)$ is infinite dimensional but has many finite dimensional quotients (the Lie algebras of the finite dimensiona filters). In the remainder of the paper, a summary of results for other classes of systems is presented

II. AN EXAMPLE WITH FINITE DIMENSIONALLY COMPUTABLE CONDITIONAL MOMENTS

Consider the system with state equations

$$\begin{aligned} dx_t &= dw_t \\ dy_t &= x_t^2 dt \end{aligned} \tag{11}$$

and observations

$$dz_t = x_t dt + dv_t \tag{12}$$

where v and w are unit variance Wiener processes, $\{x_0,y_0,v,w\}$ are independent, and x_0 is Gaussian. Th computation of $\hat{x}_t$ is of course straightforward by means of the Kalman filter, but the computation of $\hat{y}$ requires a nonlinear estimator.

For the system (11)-(12), the Zakai equation (1 in Fisk-Stratonovich form is

$$d\rho(t,x) = (-x^2\frac{\partial}{\partial y} + \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}x^2)\rho(t,x)dt + x\rho(t,x)dz_t \tag{13}$$

so the Lie algebra $L(\Sigma)$ is generated by

$A_0 = -x^2\frac{\partial}{\partial y} + \frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}x^2$ and $B_0 = x$.

The following theorem is straightforward to prove.

Theorem 1:

(i) The Lie algebra $L(\Sigma)$ generated by A_0 and B_0 has as basis the elements A_0 and $B_i \triangleq x(\partial^i/\partial y^i)$, $C_i \triangleq \partial/\partial x(\partial^i/\partial y^i)$, $D_i \triangleq \partial^i/\partial y^i$, $i=0,1,2,\ldots$.

(ii) The commutation relations are given by

$[A_0,B_i] = C_i, \quad \forall\, i$

$[A_0,C_i] = B_i + 2B_{i+1}, \quad \forall\, i$

$[A_0,D_j] = [B_i,D_j] = [C_i,D_j] = [B_i,B_j] = [C_i,C_j] = 0 \quad \forall\, i,j$

$[B_i,C_j] = -D_{i+j}, \quad \forall\, i,j$

(iii) The center of $L(\Sigma)$ is $\{D_i,\ i=0,1,2,\ldots\}$.

(iv) Every ideal of $L(\Sigma)$ has finite codimension, i.e., for any ideal I, the quotient $L(\Sigma)/I$ is finite dimensional.

(v) Let I_j be the ideal generated by B_j, with basis $\{B_i,C_i,D_i;\ i\geq j\}$. Then $I_0 \supset I_1 \supset \ldots$ and $\bigcap_j I_j = \{0\}$, s that the canonical map $\pi: L(\Sigma) \to \bigoplus_j L(\Sigma)/I_j$ is injective.

(vi) $L(\Sigma)$ is the semidirect sum of A_0 and the nilpotent ideal I_0; hence $L(\Sigma)$ is solvable.

In light of the remarks in the previous section it should be expected that many statistics of the conditional distribution can be computed with finite dimensional estimators, since there are an infinite number of finite dimensional quotients (homomorphic images) $L(\Sigma)$. By Ado's theorem, these can be realized by bilinear systems. However, we will present a slightly different realization of the sequence of quotients in (v) above: $L(\Sigma)/I_1$ is realized by the Kalman filter for $\hat{x}_t$ ($L(\Sigma)/I_1$ is the oscillator algebra [9]-[11]), and $L(\Sigma)/I_j$ ($j \geq 2$) is realized by the estimator which computes $\hat{x}_t$ and $y_t^i = E[y_t^i | z^t]$ ($i=1,2,\ldots,j-1$). Of course, the dimension of $L(\Sigma)/I_j$ increases with j, so we will only present the estimator equations for $j=4$. Other sequences of quotients possessing the property (v) can also be realized (e.g., those generated by the $\{C_j\}$), but those realizations do not have as natural an interpretation in terms of conditional moments.

The properties (iv) and (v) of the structure theorem are useful for an "estimation algebra" to possess, in the following sense: they basically say that $L(\Sigma)$ has enough finite dimensional quotients that it is determined by their direct sum. Translating this into an estimation context via the reasoning of the previous section, if we can realize all the quotients with finite dimensionally computable statistics, then these properties give us hope of being able to approximate the conditional density (or conditional characteristic function) with a convergent series of functions of these statistics, even if the conditional density <u>cannot</u> be computed exactly by a finite dimensional estimator.

The method of [6] for computing the finite dimensional estimator for $\hat{y}_t$ systematically uses the estimation equation (1) and the fact that the conditional density of x given z is Gaussian to express higher order moments in terms of lower. This procedure can also be applied to obtain equations for higher order conditional moments of y for the estimation problem (11)-(12). The first three conditional moments of y_t, together with $\hat{x}_t$ and the necessary auxiliary filter states are computed recursively by the finite dimensional estimator (in Fisk-Stratonovich form, with explicit time-dependent notation omitted):

$$
d\begin{bmatrix} \hat{x} \\ \hat{\xi} \\ \hat{y} \\ \hat{\theta} \\ \widehat{y^2} \\ \\ \hat{\phi} \\ \widehat{y^3} \\ \\ \\ t \end{bmatrix}
=
\begin{bmatrix}
-\hat{x}P \\
\hat{x}(1-P_{12}) - \hat{\xi}P^{-1} \\
\hat{x}^2 - 2\hat{x}\hat{\xi}P + P - PP_{12} \\
\hat{x}(P_{12}-P_{13}) + \hat{\xi}P(1-P_{12}) - \hat{\theta}P^{-1} \\
2\hat{x}^2\hat{y} + 2\hat{y}P + 8\hat{x}\hat{\xi}P + 4PP_{12} - 4\hat{x}\hat{\xi}\hat{y}P - 8\hat{x}\hat{\theta}P \\
\quad - 2\hat{y}PP_{12} - 4\hat{\xi}^2P^2 - 4PP_{13} \\
\hat{x}(P_{13}-P_{14}) + \hat{\xi}P(P_{12}-P_{13}) + \hat{\theta}(P-PP_{12}) - \hat{\phi}P^{-1} \\
3\hat{x}^2\widehat{y^2} + 3\widehat{y^2}P + 24\hat{x}\hat{\xi}\hat{y}P + 48\hat{x}\hat{\theta}P + 24\hat{\xi}^2P^2 + 12\hat{y}PP_{12} \\
\quad + 24PP_{13} - 3\widehat{y^2}PP_{12} - 48\hat{\xi}\hat{\theta}P^2 - 12\hat{y}PP_{13} \\
\quad - 12\hat{\xi}^2\hat{y}P^2 - 24PP_{14} - 6\hat{x}\hat{\xi}\widehat{y^2}P - 24\hat{x}\hat{\theta}\hat{y}P - 48\hat{x}\hat{\phi}P \\
1
\end{bmatrix} dt +
\begin{bmatrix}
P \\ P_{12} \\ 2\hat{\xi}P \\ P_{13} \\ 4\hat{\xi}\hat{y}P + 8\hat{\theta}P \\ \\ P_{14} \\ 6\hat{\xi}\widehat{y^2}P + 24\hat{\theta}\hat{y}P + 48\hat{\phi}P \\ \\ \\ 0
\end{bmatrix} dz;
\quad
\begin{bmatrix} \hat{x}_0 \\ \hat{\xi}_0 \\ \hat{y}_0 \\ \hat{\theta}_0 \\ \widehat{y_0^2} \\ \\ \hat{\phi}_0 \\ \widehat{y_0^3} \\ \\ \\ t_0 \end{bmatrix}
=
\begin{bmatrix} E[x_0] \\ 0 \\ E[y_0] \\ 0 \\ E[y_0^2] \\ \\ 0 \\ E[y_0^3] \\ \\ \\ 0 \end{bmatrix}
\qquad (14)
$$

where we write the right-hand side of (14) more compactly as $a_0 dt + b_0 dz$. The nonrandom conditional covariance equations are

$$
\begin{aligned}
\dot{P} &= 1 - P^2 \\
\dot{P}_{12} &= P - (P+P^{-1})P_{12} \\
\dot{P}_{13} &= 2PP_{12} - PP_{12}^2 - (P+P^{-1})P_{13} \\
\dot{P}_{14} &= 2PP_{13} + PP_{12}^2 - 2PP_{12}P_{13} - (P+P^{-1})P_{14} \\
P(0) &= \mathrm{cov}(x_0) \neq 0;\ P_{12}(0) = P_{13}(0) = P_{14}(0) = 0
\end{aligned}
\qquad (15)
$$

The estimator (14) is obtained by first augmenting the state x with auxiliary states ξ, θ, and ϕ; then the Kalman filter for the linear system with states $[x,\xi,\theta,\phi]$ and observations z computes $[\hat{x},\hat{\xi},\hat{\theta},\hat{\phi}]$. In addition, $[P,P_{12},P_{13},P_{14}]$ is the first row of the Kalman filter error covariance matrix; (15) is obtained by selecting the corresponding components of the Riccati equation. Then $\hat{y}$, $\widehat{y^2}$, and $\widehat{y^3}$ are seen, after tedious calculations, to be computed by the given equations. The filter state is augmented with t in order to make (14) time-invariant, thus facilitating the use of Lie algebraic techniques. The filter (14) can be viewed as a cascade of linear filters [18]: $[\hat{x},\hat{\xi},\hat{\theta},\hat{\phi},t]$ satisfies a linear equation; some of these states then feed forward and can be viewed as parameters in a linear equation for $\hat{y}$; the states $\hat{x},\hat{\xi},\hat{\theta},\hat{y},t$ then feed forward as parameters into a linear equation for $\widehat{y^2}$; etc. This structure is typical of the class of finite dimensional estimators derived in [6]-[8].

In order to study the structure of the estimation problem as discussed in Section I, we must analyze the Lie algebra F generated by a_0 and b_0 in (14). The structure of the class of problems of [6] is analyzed from a different point of view in [22].

Theorem 2:
(i) F has as basis the elements a_0; b_i, c_i, i=0,1,2,3; d_i, i=1,2,3, where a_0 and b_0 are given in (14) and

$$c_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 0 \\ 1 \\ 2\hat{x} \\ P_{12} \\ 4\widehat{xy} + 8\hat{\xi}P \\ P_{13} \\ 6\widehat{xy^2} + 24\hat{\xi}\hat{y}P + 48\hat{\theta}P \\ 0 \end{bmatrix}, \quad c_1 = \begin{bmatrix} 0 \\ p^{-1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$b_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 8\hat{x} \\ P_{12} \\ 24\widehat{xy} + 48\hat{\xi}P \\ 0 \end{bmatrix}, \quad c_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ p^{-1} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad b_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 48\hat{x} \\ 0 \end{bmatrix},$$

$$c_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ p^{-1} \\ 0 \\ 0 \end{bmatrix}, \quad d_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 2\hat{y} \\ 0 \\ 3y^2 \\ 0 \end{bmatrix}, \quad d_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 3\hat{y} \\ 0 \end{bmatrix}, \quad d_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

(ii) The commutation relations are given by

$[a_0, b_i] = c_i$, i=0,1,2,3

$[a_0, c_i] = b_i - b_{i+1}$, i=0,1,2

$[a_0, c_3] = b_3$

$$[b_i, c_j] = \begin{cases} -2d_{i+j}, & i+j=1 \\ -8d_{i+j}, & i+j=2 \\ -48d_{i+j}, & i+j=3 \\ 0, & \text{otherwise} \end{cases}$$

$[a_0, d_j] = [b_i, d_j] = [c_i, d_j] = 0, \quad \forall\, i,j$

(iii) Let $\tilde{I}_4$ be the ideal in $L(\Sigma)$ with basis B_i, C_i, D_i, $i \geq 4$ and D_0. Then F is isomorphic to $L(\Sigma)/\tilde{I}_4$; hence, F is also solvable.
(iv) The isomorphism ϕ between $L(\Sigma)/\tilde{I}_4$ and F is given by:

$\phi(A_0) = a_0$; $\phi(B_i) = (-\tfrac{1}{2})^i b_i$,
$\phi(C_i) = (-\tfrac{1}{2})^i c_i$, i=0,1,2,3; $\phi(D_i) = (-1)^i (i!) d_i$, i=1,2,3;
$\phi(E) = 0$, $E \in \tilde{I}_4$.

(v) F is the semidirect sum of a_0 and the nilpotent ideal generated by b_0.

Remarks:
(i) The estimator (14) is not quite a realization of $L(\Sigma)/I_4$, since D_0 is also in the kernel of the homomorphism (i.e., the ideal $\tilde{I}_4$). However, a finite dimensional estimator realizing $L(\Sigma)/I_4$ (or $L(\Sigma)/I_j$, for any j) is easily obtained by augmenting (14) with the equation for the normalization factor α_t for $\rho(t,x)$ (the denominator of (5)) which satisfies (in Fisk-Stratonovich form)

$$d\alpha_t = -\frac{1}{2}(\hat{x}_t^2 + P_t)\,\alpha_t\,dt + \hat{x}_t \alpha_t dt \tag{16}$$

If (16) is augmented at the end of (14), the Lie algebra generated by a_0 and b_0 has the same commutation relations as in (ii) above, except that $[b_0, c_0] = [0, \ldots, 0, \alpha]' \triangleq d_0$, and d_0 commutes with all the other elements. Thus, a realization of $L(\Sigma)/I_4$ is an easy modification of (14).
(ii) The property (v) is typical of a cascade of linear systems.
(iii) One of the conditions in [9] for the existence of a Lie algebra homomorphism from L to the Lie algebra of a finite dimensional estimator is that the estimator be a "minimal" realization in a certain sense. If we consider the output of (14) to be $\widehat{y^3}$ and consider this realization of the input-output map from z to $\widehat{y^3}$, then it can be verified by the methods of [15] that the realization is locally weakly controllable and locally weakly observable. This implies that there is no other realization with lower dimension; it is in this sense that the statistics $\hat{\xi}$, $\hat{\theta}$, $\hat{\phi}$ are necessary for the computation of $\widehat{y^3}$.
(iv) An even more detailed analysis of the Lie algebraic structure of this example is carried out in [17], and we have also done a similar analysis for systems of the form (11), with x_t^2 replaced by a general monomial x_t^p; for $p > 2$, a similar but more complex Lie algebraic structure is exhibited.

III. PRO-FINITE DIMENSIONAL FILTERED LIE ALGEBRAS

A Lie algebra L is defined to be a pro-finite dimensional filtered Lie algebra if L has a decreasing sequence of ideals $L = L_{-1} \supset L_0 \supset L_1$ ... such that

(a) $\cap L_i = 0$

(b) L/L_i is a finite dimensional Lie algebra for all i.

The terminology is analogous to that of pro-finite groups [19], and this property is possessed by $L(\Sigma)$ for the example of the previous section. For a more general estimation problem with $L(\Sigma)$ having this property, if each of the quotients can be realized with a recursively filterable statistic, then the injectivity of the map makes it reasonable to conjecture that these statistics represent some type of power series expansion of the conditional density. Of course, many other difficult technical questions such as moment determinacy will be relevant to this problem, but the structure of the Lie algebra should provide some guidance as to possible successful approaches to the problem.

The class of estimation problems considered in [6]-[8] has the property that $L(\Sigma)$ is a pro-finite dimensional filtered Lie algebra, as illustrated by the previous example. In fact, the sequence of finite dimensional filters for the conditional moments realize the finite dimensional quotients. Another class of examples is the following.

Example 1 (degree increasing operators and bilinear systems): Consider a system of the form (Σ), and suppose that f, G, and h are analytic with $f(0)=0$ and $G(0)=0$, and that the power series expansions of f and G around zero are of the form

$$f(x) = \sum_{|\alpha|\geq 1} f_\alpha(x)x^\alpha, \quad G(x) = \sum_{|\alpha|\geq 1} G_\alpha(x)x^\alpha,$$

where α ranges over the multiindices $\alpha=(\alpha_1,\ldots,\alpha_n)$ and $|\alpha| = \alpha_1+\ldots+\alpha_n$. An example of such systems is the class of bilinear systems

$$\begin{aligned} dx_t &= Ax_t + \sum_{i=1}^{k} B_i x_t dw_t^i \\ dz_t &= Cx_t dt + dv_t \end{aligned} \tag{17}$$

Another example is

$$\begin{aligned} dx_t &= x_t dt + (e^{-x_t} - 1 + x_t)dw_t \\ dz_t &= h(x_t)dt + dv_t \end{aligned} \tag{18}$$

with h analytic; in general, a wide variety of examples can be found.

Let $M = \mathbb{R}[[x_1,\ldots,x_n]]$ be the module of all (formal) power series in $x_1,\ldots,x_n$, and define the submodules

$$M_i = \{\sum \alpha x^\alpha | a_\alpha = 0 \text{ for } |\alpha| < 1\}.$$

It is easy to see that, for this class of examples,

$$(L - \tfrac{1}{2}h^2(x))\ M_i \subset M_i$$

and

$$h(x)\ M_i \subset M_i$$

and thus $L(\Sigma)M_i \subset M_i$. It can be shown that $L(\Sigma)$ is a pro-finite dimensional filtered Lie algebra, with filtration given by

$$L_i \triangleq \{X \in L(\Sigma) | X M \subset M_i\}.$$

IV. THE WEYL ALGEBRAS W_n

The Weyl algebra W_n [20] is the algebra of all polynomial differential operators; i.e., $W_n = \mathbb{R} < x_1,\ldots,x_n; \frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n} >$. A basis for W_n consists of all monomial expressions

$$e_{\alpha,\beta} \triangleq x^\alpha \frac{\partial^\beta}{\partial x^\beta} \triangleq x_1^{\alpha_1} \cdots x_n^{\alpha_n} \frac{\partial^{\beta_1}}{\partial x_1^{\beta_1}} \cdots \frac{\partial^{\beta_n}}{\partial x_n^{\beta_n}} \tag{19}$$

where α,β range over all multiindices $\alpha=(\alpha_1,\ldots,\alpha_n)$, $\beta=(\beta_1,\ldots,\beta_n)$, $\alpha,\beta \in \mathbb{N} \cup \{0\}$ (the non-negative integers). The center of W_n (i.e., the ideal of all elements $Z \in W_n$ such that $[X,Z]=0$ for all $X \in W_n$) is the one-dimensional space $\mathbb{R}\cdot 1$ with basis $\{1\}$.

Theorem 3: The Lie algebra $W_n/\mathbb{R}\cdot 1$ is simple -- i.e., it has no ideals other than $\{0\}$ and the whole Lie algebra. Equivalently, W_n has no ideals other than $\{0\}$, $\mathbb{R}\cdot 1$, and W_n.

Let $\hat{V}_n$ be the Lie algebra of vector fields $\hat{V}_m \triangleq \{\sum_{i=1}^{m} f_i(x_1,\ldots,x_m)\frac{\partial}{\partial x_i}\}$ with (formal) power series coefficients $f_i \in \mathbb{R}[[x_1,\ldots,x_m]]$, and let $V(M)$ be the Lie algebra of C^∞-vector fields on a C^∞-manifold M. The following results can be proved [21].

Theorem 4: Fix $n \neq 0$. Then there are no non-zero homomorphisms from W_n to $\hat{V}_m$ or from $W_n/\mathbb{R}\cdot 1$ to $\hat{V}_m$ for any m.

Theorem 5: Fix $n \neq 0$. Then there are no non-zero homomorphisms from W_n to $V(M)$ or $W_n/\mathbb{R}\cdot 1$ to $V(M)$ for any finite dimensional C^∞-manifold M.

These results suggest (assuming the appropriate analog of the results of [9]) that if a system Σ has estimation algebra $L(\Sigma) = W_n$ for some n, then neither the conditional density of x_t given z^t nor any statistic of the conditional density can be computed with a finite dimensional filter of the form (9) with a and b C^∞ or analytic. Typical examples of such systems are given below.

Example 2 (the cubic sensor problem [22]): Consider the system

$$\begin{aligned} dx &= dw \\ dz &= x^3 dt + dv \end{aligned} \tag{20}$$

$L(\Sigma)$ is generated by $\frac{1}{2}\frac{\partial^2}{\partial x^2} - \frac{1}{2}x^6$ and x^3; it is shown that $L(\Sigma) = W_1$.

Example 3 (mixed linear-bilinear type): Consider the system

$$\begin{aligned} dx &= y\,dt + y\,dw_1 \\ dy &= dw_2 \\ dz &= x\,dt + dv \end{aligned} \tag{21}$$

$L(\Sigma)$ is generated by $-y \frac{\partial}{\partial x} + \frac{1}{2} y^2 \frac{\partial^2}{\partial x^2} + \frac{1}{2} \frac{\partial^2}{\partial y^2} - \frac{1}{2} x^2$ and x; it is shown that $L(\Sigma) = W_2$.

Example 4: Consider the system with state equations (11) and observations

$$dz_{1t} = x_t dt + dv_{1t}$$

$$dz_{2t} = y_t dt + dv_{2t}.$$

$L(\Sigma)$ is generated by $\frac{1}{2} \frac{\partial^2}{\partial x^2} - x^2 \frac{\partial}{\partial y} - \frac{1}{2} x^2 - \frac{1}{2} y^2$, x, and y; it is easily shown that $L(\Sigma) = W_2$. This is precisely the example of Section II, but the addition of the observation z_2 transforms $L(\Sigma)$ into W_2.

We note that a number of the calculations in this section are similar to those of [24] for Lie algebras of polynomials under the Poisson bracket.

REFERENCES

1. H. J. Kushner, "Dynamical equations for optimal nonlinear filtering," *J. Diff. Equations,* Vol. 3, 1967, pp. 179-190.
2. M. Fujisaki, G. Kallianpur, and H. Kunita, "Stochastic differential equations for the nonlinear filtering problem," *Osaka J. Math.,* Vol. 1, 1972, pp. 19-40.
3. R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes I.* New York: Springer-Verlag, 1977.
4. H. J. Kushner, "On the dynamical equations of conditional probability functions with application to optimal stochastic control theory," *J. Math. Anal. Appl.,* Vol. 8, 1964, pp. 332-344.
5. J. H. Van Schuppen, "Stochastic filtering theory: A discussion of concepts, methods and results," in *Stochastic Control Theory and Stochastic Differential Systems,* M. Kohlmann and W. Vogel, eds., New York: Springer-Verlag, 1979.
6. S. I. Marcus and A. S. Willsky, "Algebraic structure and finite dimensional nonlinear estimation," *SIAM J. Math. Anal.,* Vol. 9, April 1978, pp. 312-327.
7. S. I. Marcus, "Optimal nonlinear estimation for a class of discrete-time stochastic systems," *IEEE Trans. Automatic Control,* Vol. AC-24, April 1979, pp. 297-302.
8. S. I. Marcus, S. K. Mitter, and D. Ocone, "Finite dimensional nonlinear estimation for a class of systems in continuous and discrete time," *Proc. Int. Conf. on Analysis and Optimization of Stochastic Systems,* Oxford, Sept. 6-8, 1978.
9. R. W. Brockett, "Remarks on finite dimensional nonlinear estimation," *Journees sur l'Analyse des Systèmes,* Bordeaux, 1978.
10. R. W. Brockett, "Classification and equivalence in estimation theory," *Proc. 1979 IEEE Conf. on Decision and Control,* Ft. Lauderdale, Dec. 1979.
11. S. K. Mitter, "Modeling for stochastic systems and quantum fields," *Proc. 1978 IEEE Conf. on Decision and Control,* San Diego, January 1979.
12. M. Zakai, "On the optimal filtering of diffusion processes," *Z. Wahr. Verw. Geb.,* Vol. 11, 1969, pp. 230-243.
13. R. W. Brockett, "System theory on group manifolds and coset spaces," *SIAM J. Control,* Vol. 10, 1972, pp. 265-284.
14. H. J. Sussmann, "Existence and uniqueness of minimal realizations of nonlinear systems," *Math. Systems Theory,* Vol. 10, 1976/1977.
15. R. Hermann and A. J. Krener, "Nonlinear controllability and observability," *IEEE Trans. Automatic Control,* Vol. AC-22, October 1977, pp. 728-740.
16. V. E. Benes, "Exact finite dimensional filters for certain diffusions with nonlinear drift," presented at the 1979 IEEE Conf. on Decision and Control, Ft. Lauderdale, December 1979.
17. C.-H. Liu and S. I. Marcus, "The Lie algebraic structure of a class of finite dimensional nonlinear filters," in *Filterdag Rotterdam 1980,* M. Hazewinkel (ed.), Report of the Econometric Inst., Erasmus University, Rotterdam, 1980.
18. E. D. Sontag, *Polynomial Response Maps.* New York: Springer-Verlag, 1979.
19. J.-P. Serre, *Cohomologie Galoisienne,* Lecture Notes in Math. 5. New York: Springer-Verlag, 1964.
20. J. Dixmier, *Enveloping Algebras.* Amsterdam: North Holland, 1977.
21. M. Hazewinkel and S. I. Marcus, "On Lie algebras and finite dimensional filtering," to appear.
22. R. S. Bucy and J. Pages, "A priori error bounds for the cubic sensor problem," *IEEE Trans. Automatic Control,* Vol. AC-23, February 1978, pp. 88-91.
23. S. D. Chikte and J. T.-H. Lo, "Optimal filters for bilinear systems with nilpotent Lie algebras," *IEEE Trans. Automatic Control,* Vol. AC-24, December 1979, pp. 948-953.
24. A. Avez and A. Heslot, "L'algèbre de Lie des polynômes en les coordonnées canoniques munie du crochet de Poisson," *C. R. Acad. Sci. Paris A,* T.288, 7 Mai 1979, pp. 831-833.