A LOWER BOUND ON THE ESTIMATION ERROR FOR CERTAIN DIFFUSION PROCESSES

by

Ben Zion Bobrovsky and Moshe Zakai

Memorandum No. ERL-M490

October 1974

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

A LOWER BOUND ON THE ESTIMATION ERROR FOR CERTAIN DIFFUSION PROCESSES

Ben Zion Bobrovsky and Moshe Zakai*

ABSTRACT

A lower bound on the minimal mean square error in estimating non-linear diffusion processes is derived. The bound holds for causal and noncausal filtering.

*Faculty of Electrical Engineering, Technion-Israel Institute of Technology, Haifa, Israel. The work of the second author was supported in part by the U.S. Army Research Office, Grant DAHC04-74-G0087.

1. INTRODUCTION

The filtering problem for diffusion processes was considered by Stratonovich [1], Kushner [2] and others. It is well known that except for the linear - Gaussian case, the implementation of the optimal filter is impossible. Consequently, suboptimal approximations to the optimal filter have received considerable attention. Furthermore, there is no tractable solution for the performance of the optimal filter (except for a few special cases) which raises the problem of comparing the performance of the suboptimal solution with that of the optimal one.

A possible way to overcome this difficulty is to derive useful bounds on the mean square error associated with the optimal filter. Upper and lower bounds for this problem were first introduced in [3]. The lower bounds presented in [3] were based on information theoretic arguments. Lower bounds for the estimation of Gaussian processes from nonlinear measurements based on the Cramér-Rao bound, were derived in [4]. In a previous paper [5] the authors have presented a lower bound on the filtering error of certain Markov processes, the bound was based on the Van-Trees version of the Cramér-Rao bound [6]. A heuristic proof was given in [5] for the time continuous case and it seems difficult to construct a rigorous proof along the same lines. A rigorous proof, by a different approach is presented in this paper. In our special case of estimating a Gaussian process from nonlinear measurements the results coincide with those of Snyder and Rhodes [4].

The main results of the present paper are given in Sections 2 and 5. A general bound is derived in Section 2, it can be considered as an infinite dimensional extension of the Van-Trees version of the Cramér-

Rao bound [6]. The relation between the two is very roughly as follows, let $p(\underline{y}, \underline{\alpha})$ be a probability density in the vector $\underline{y}$, $\underline{\alpha}$ variables, then the mean square error of estimating $\underline{\alpha}$ from $\underline{y}$ is lower bounded in [6] by elements of the inverse of the matrix

$$E\left\{\frac{\partial \log p(\underline{y},\underline{\alpha})}{\partial \alpha_i} \quad \frac{\partial \log p(\underline{y},\underline{\alpha})}{\partial \alpha_j}\right\}$$

Consider now the case where $\underline{\alpha}$ is two dimensional and replace, $\partial \log p/\partial\alpha_1$ by

$$\frac{1}{\delta}\left(1 - \frac{p(\underline{y},\alpha_1+\delta,\alpha_2)}{p(\underline{y},\alpha_1,\alpha_2)}\right)$$

in this form the ratio $p(\underline{y},\alpha_1+\delta,\alpha_2)/p(\underline{y},\alpha_1,\alpha_2)$ can be replaced by a Radon-Nykodym derivative and a general bound can be derived for the case where $\underline{\alpha}$ is infinite dimensional. The bound of section 2 is also applicable to other problems and is, therefore, of independent interest.

The results of Section 2 are specialized in Section 3 to certain diffusion processes. The α_i components of $\underline{\alpha}$ being linear functionals of a diffusion process $\underline{x}_s$, $0 \leq s \leq T$. In Section 4, it is shown that for the linear case a limiting form of the bound is satisfied by equality. The results derived in Section 5, when specialized to a one dimensional time invariant diffusion are roughly as follows. Let

$$\begin{aligned} dx_t &= m(x_t)dt + dw_t \\ dy_t &= g(x_t)dt + \sqrt{N_0}\, d\nu_t \end{aligned} \tag{1}$$

$y_0 = 0$ and the density of x_0 is $p(x_0)$. Let

$$\sigma_0^2 = \left(\int_{-\infty}^{\infty} p(x)\left(\frac{\partial \log p(x)}{\partial x}\right)^2 dx\right)^{-1}$$

$dm(x)/dx = \dot{m}(x)$, $dg(x)/dx = \dot{g}(x)$. Define A_t, B_t by

$$A_t = E\ \dot{m}(x_t)$$

$$N_0^{-1}B_t^2 + A_t^2 = E\,(\dot{m}^2(x_t) + \dot{g}^2(x_t))$$

(note that B_t is real). Consider the linear system

$$du_t = A_t u_t dt + dw_t$$

$$dv_t = B_t u_t dt + \sqrt{N_0}\, d\nu_t \tag{2}$$

$E\ u_0 = 0$, $E\ u_0^2 = \sigma_0^2$, where A_t, B_t, σ_0^2 are as defined above. Then it is shown in Section 5 that the filtering error associated with the linear system (2) is a lower bound to that of the nonlinear system (1), namely if $\hat{x}_T = E(x_T | y_t,\ 0 \leq t \leq T)$, $\hat{u}_T = E\{u_T | v_t,\ 0 \leq t \leq T\}$ then

$$E(x_T - \hat{x}_t)^2 \geq E(u_T - \hat{u}_T)^2$$

and a similar result holds for the smoothing problem

2. A GENERAL LOWER BOUND

Let $(Y, \mathcal{B}_Y)$ be a measurable space and let $(R_i, \mathcal{B}_i)$, $i = 1,2,\ldots$ be replicas of the real line with the Borel σ-field on the real line. Let $\Omega = Y \times \prod_{i=1}^{\infty} R_i$, $\mathcal{A}$ will denote the minimal σ-field over $(\mathcal{B}_y, \mathcal{B}_i\ \ i = 1,2,\ldots)$. Consider now the probability space $(\Omega, \mathcal{A}, \mu)$, $\Omega = \{\omega\}$ where $\omega = (y, \alpha_1, \alpha_2, \ldots)$ where $y \in Y$, $\alpha_i \in R_i$.

For any $A \in \mathcal{A}$ and $\delta \neq 0$ we define $A(i,\delta)$ as follows

$$A(i,\delta) = \{\omega:\ (y, \alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \alpha_i + \delta, \alpha_{i+1}, \ldots) \in A\}$$

namely, $A(i,\delta)$ is obtained from A by shifting by $-\delta$ the i-th α coordinate of all the elements of A. Obviously $A(i,\delta) \in \mathcal{A}$. Define now a new measure $\mu_{i,\delta}$ on $(\Omega, \mathcal{A})$ as follows: for all $A \in \mathcal{A}$

$$\mu_{i,\delta}(A) = \mu(A(i,\delta))$$

Let $\mathcal{A}_i$ be the sub-σ-field of $\mathcal{A}$ induced by $\mathcal{B}_i$, namely the cylinders with base in $\mathcal{B}_i$. Let $\mathcal{A}_{-i}$ be the sub-σ-field of $\mathcal{A}$ induced by $\mathcal{B}_y$ and $\mathcal{B}_j$, $i \neq j$; then $\mathcal{A}_{-i}$ contains all the sets in $\mathcal{A}$ such that $A(i,\delta) = A$, $-\infty < \delta < \infty$. E will be used to denote expectations with respect to the μ measure.

Theorem 1 Assume that $\mu_{i,\delta}$ is absolutely continuous with respect to μ, ($\mu_{i,\delta} << \mu$) $i = 1,2,\ldots,N$, $E\,\alpha_i^2 < \infty$, $\int|\alpha_i|d\mu_{i,\delta} < \infty$, $i = 1,\ldots,N$. Let $\underline{J}(N,\delta)$ denote the N×N matrix

$$(\underline{J}(N,\delta))_{i,j} = \frac{1}{\delta^2}\,E\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)\left(1 - \frac{d\mu_{j,\delta}}{d\mu}\right) \tag{3}$$

Let $\hat{\alpha}_i = E(\alpha_i|\mathcal{B}_Y)$ and let $\underline{\varepsilon}_N^2$ denote the error matrix

$$(\underline{\varepsilon}_N^2)_{ij} = E(\alpha_i - \hat{\alpha}_i)(\alpha_j - \hat{\alpha}_j)$$

If $\underline{J}(N,\delta)$ is nonsingular then

$$\underline{\varepsilon}_N^2 \geq J_N^{-1}(N,\delta) \tag{4}$$

Furthermore, if as $\delta\to0$, $\lim \underline{J}(N,\delta)$ exists and is nonsingular then

$$\underline{\varepsilon}_N^2 \geq (\lim_{\delta\to0} \underline{J}(N,\delta))^{-1}. \tag{5}$$

Proof Note, first, that if $f(\omega)$ is $\mathcal{A}_{-i}$ measurable then

$$\int f(\omega)d\mu = \int f(\omega)d\mu_{i,\delta}$$

where the integration is over Ω. Therefore, since for $j \neq i$, α_j and $\hat{\alpha}_j$ are $\mathcal{A}_{-i}$ measurable

$$\frac{1}{\delta}\,E\{(\alpha_j - \hat{\alpha})\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)\} = 0 \tag{6}$$

Also not that $\int\alpha_i d\mu_{i,\delta} = \int(\alpha_i+\delta)d\mu$ therefore

$$\frac{1}{\delta}\,E\left\{(\alpha_i - \hat{\alpha}_i)\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)\right\} = -1 \tag{7}$$

Let

$$\rho_{i,\delta} = \frac{1}{\delta}\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right).$$

Setting

$$\underline{z}^T = ((\hat{\alpha}_1-\alpha), (\hat{\alpha}_2-\alpha), \ldots, (\hat{\alpha}_N-\alpha),\ \rho_{1,\delta}, \ldots, \rho_{N,\delta})$$

then $E\ \underline{z}\ \underline{z}^T \geq 0$, therefore, by (6) and (7)

$$\begin{pmatrix} \underline{\varepsilon}_N^2 & \underline{I} \\ \underline{I} & \underline{J}(N,\delta) \end{pmatrix} \geq 0$$

where I is the N×N identity matrix. Therefore, for any pair of N vectors $\underline{u}$, $\underline{v}$

$$(\underline{u}^T, \underline{v}^T)\begin{pmatrix} \underline{\varepsilon}_N^2 & \underline{I} \\ \underline{I}\ , & \underline{J} \end{pmatrix}\begin{pmatrix} \underline{u} \\ \underline{v} \end{pmatrix} \geq 0 \qquad (8)$$

Setting, in particular, $\underline{v} = -(\underline{J}(N,\delta))^{-1}\underline{u}$ we obtain that for any $\underline{u}$, $\underline{u}^T(\underline{\varepsilon}^2 - J^{-1})\underline{u} \geq 0$ which proves (4). Equation (5) follows by the same argument after taking the limit of (8) as $\delta \to 0$.

<u>Theorem 2</u> Under the conditions of Theorem 1, if $M < N$ then

$$\underline{\varepsilon}_M^2 \geq (\underline{J}^{-1}(N,\delta))_M$$

$$\underline{\varepsilon}_M^2 \geq ((\lim_{\delta\to 0} J(N,\delta))^{-1})_M$$

where $(A)_M$ is the M×M matrix obtained from the N×N matrix A by deleting the last (N-M) rows and columns. In particular $E(\alpha_1-\hat{\alpha}_1)^2 \geq (J^{-1}(N,\delta))_{11}$.

The proof is the same as that of Theorem 1 except that $\underline{u}^T = (u_1^T, 0, \ldots, 0)$ where $\underline{u}_1$ is an arbitrary M vector.

<u>Remark</u>: The matrix $\lim \underline{J}(N,\delta)$, $\delta \to 0$, will be called the information matrix.

3. THE INFORMATION MATRIX ASSOCIATED WITH THE NONLINEAR FILTERING PROBLEM

Let $\underline{x}_t$, $\underline{y}_t$ satisfy the stochastic differential equation

$$d\underline{x}_t = \underline{m}(\underline{x}_t,t)dt + \underline{C}_t\, d\underline{w}_t$$

$$d\underline{y}_t = \underline{g}(\underline{x}_t,t)dt + \underline{D}_t\, d\underline{\nu}_t \tag{9}$$

where $\underline{w}_t$, $\underline{\nu}_t$ are independent standard Brownian motions of dimensions m and g respectively. $\underline{x}_t$ and $\underline{y}_t$ are m and g dimensional vectors respectively. $\underline{y}_0 = 0$, assumptions on $\underline{x}_0$ will be made later. The elements of the m vector $\underline{m}(x,t)$ are assumed to be differentiable with respect to x_i, i=1,...,m and $\frac{\partial m_i(\underline{x},t)}{\partial x_j}$ is assumed to be bounded and continuous over $\|\underline{x}\| < \infty$, $0 \leq t \leq T$. A similar assumption is made on the elements of the g vector $\underline{g}(\underline{x},t)$. Denote

$$(\underline{\dot{m}}(\underline{x},t))_{ij} = \frac{\partial m_i(\underline{x},t)}{\partial x_j}$$

$$(\underline{\dot{g}}(\underline{x},t))_{ij} = \frac{\partial g_i(\underline{x},t)}{\partial x_j}$$

(the boundedness of $\dot{m}$, $\dot{g}$ implies that $\underline{m}$ ang $\underline{g}$ satisfy a uniform Lipschitz condition. The continuity of $\underline{\dot{m}}$, $\underline{\dot{g}}$ implies the continuity of E $\underline{\dot{m}}(\underline{x}_t,t)$, E $\underline{\dot{g}}(x_t,t)$ in t). $\underline{C}_t$ and $\underline{D}_t$ are m×m and g×g matrices respectively, non-singular with continuous entries for $0 \leq t \leq T$. The requirement that C_t be nonsingular can be relaxed as will be pointed out later. E will be used to denote expectations with respect to the measure induced by (9).

<u>Theorem 3</u> Let $\underline{x}_0 = \underline{0}$ and let μ be the measure induced by $\underline{x}_t, y_t$, $0 \leq t \leq T$ of (9). Let $f_i(t)$ be m dimensional vector valued function of t possessing continuous derivatives $\dot{f}_i(t) = d\, f_i(t)/dt$ on [0,T], $f_i(t) = 0$. Let $\mu_{i,\delta}$ be the measure induced by $\underline{x}_t'$, $\underline{y}_t'$, $0 \leq t \leq T$ where

$$\underline{x}'_t = \underline{x}_t + \delta\ f_i(t)$$

$$\underline{y}'_s = y_s$$

Then

$$\lim_{\delta\to 0} E\ \frac{1}{\delta^2}\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)\left(1 - \frac{d\mu_{j,\delta}}{d\mu}\right) = \tag{10}$$

$$= E\left\{\int_0^T (\dot{f}_i(t) + \underline{\dot{m}}(\underline{x}_t,t)\ f_i(t))^T(C_tC_t^T)^{-1}\ (\dot{f}_j(t) - \underline{\dot{m}}(\underline{x}_t,t)f_j(t))dt\right.$$

$$\left. + \int_0^T\ (\underline{\dot{g}}(\underline{x}_t,t)\ f_i(t))^T(D_tD_t^T)^{-1}\quad \dot{g}(x_t,t)\ f_j(t)dt\right\}$$

<u>Proof</u>: From (9) we have

$$d\underline{x}'_t = \delta\underline{\dot{\phi}}_i(t)dt + \underline{m}(\underline{x}_t - \delta\underline{\phi}_i(t),t)dt + \underline{C}_t\ d\underline{w}_t$$

$$d\underline{y}'_t = \underline{g}(\underline{x}_t - \delta\underline{\phi}_i(t))dt + \underline{D}_t\ d\underline{\nu}_t$$

Let

$$\underline{a}_{i,\delta}(s) = \underline{C}_s^1(-\underline{m}(\underline{x}_s,s) + \underline{m}(\underline{x}_s - \delta\underline{\phi}_i(s),s) + \delta\dot{\phi}_i(s))$$

$$\underline{b}_{i,\delta}(s) = \underline{D}_s^1(-\underline{g}(\underline{x}_s,s) + \underline{g}(\underline{x}_s - \delta\underline{\phi}_i(s),s)) \tag{11}$$

Note that $\underline{a}$ and $\underline{b}$ are bounded. By [7] and [8] the measures μ and $\mu_{i,\delta}$ are mutually absolutely continuous and

$$\frac{d\mu_{i,\delta}}{d\mu} = \text{Exp}\left\{\int_0^T \underline{a}^T(s)d\underline{w}_s - \frac{1}{2}\int_0^T \underline{a}^T(s)\underline{a}(s)ds + \right.$$

$$\left. + \int_0^T \underline{b}^T(s)d\underline{\nu}_s - \frac{1}{2}\int_0^T \underline{b}^T(s)\ \underline{b}(s)ds\right\} \tag{12}$$

We proceed through the following three lemma:

Lemma 1

$$\sup_{0<\delta\leq\delta_0} E\left(\frac{d\mu_{i,\delta}}{d\mu}\right)^n < \infty$$

Proof: By (12)

$$\left(\frac{d\mu_{i,\delta}}{d\mu}\right)^k = \text{Exp}\ \frac{n^2-n}{2} \int_0^T (\underline{a}^T\underline{a} + \underline{b}^T\underline{b})dt \cdot$$

$$\cdot\ \text{Exp}\left\{\int_0^T n\underline{a}^T\, d\underline{w}_t - \int_0^T \frac{n^2}{2}\, \underline{a}^T\underline{a}\, dt + \int_0^T n\underline{b}^T d\nu_t \right.$$

$$\left. - \int_0^T \frac{n^2}{2}\, \underline{b}^T\underline{b}\, dt\right\}$$

$$\leq \text{Exp}\ \frac{n^2-n}{2}\, K\, \delta^2\, T \cdot$$

$$\cdot\,\text{Exp}\left\{\int_0^T n\underline{a}^T\, d\underline{w}_t - \int_0^T \frac{n^2}{n}\, \underline{a}^T\underline{a}\, dt + \int_0^T n\, \underline{b}^T\, d\underline{\nu}_t \right.$$

$$\left. - \frac{1}{2}\int_0^T \frac{n^2}{n}\, \underline{b}^T\underline{b}\, dt\right\}$$

The expectation of the second factor in the right hand side of the last equation is 1 [7], therefore,

$$\sup_{0<\delta\leq\delta_0} E\left(\frac{d\mu_{i,\delta}}{d\mu}\right)^n \leq \text{Exp}\ \frac{n^2-n}{2}\ K\delta^2 T < \infty$$

Lemma 2 Let

$$\underline{a}_i(s) = -\underline{C}_s^{-1}(\underline{\dot{m}}(\underline{x}_s,s)\underline{\phi}(s) - \dot{\phi}_i(s))$$

$$\underline{b}_i(s) = -D_s^{-1}\, \underline{\dot{g}}(\underline{x}_s,s)\underline{\phi}_i(s) \qquad (13)$$

Then, as $\delta \to 0$

$$\frac{1}{\delta}\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right) \xrightarrow{\text{q.m.}} \int_0^T \underline{a}_i^T(s)\, d\underline{w}_s + \int_0^T \underline{b}_i^T(s) d\nu_s$$

<u>Proof</u>: Let

$$\Lambda_s^{i,\delta} = \text{Exp}\left\{\int_0^T \underline{a}_{i,\delta}^T(s) d\underline{w}_s - \frac{1}{2}\int_0^T \underline{a}_{i,\delta}^T(s)\underline{a}_{i,\delta}(s) ds + \right.$$
$$\left. + \int_0^T \underline{b}_{i,\delta}^T(s)\, d\underline{\nu}_s - \frac{1}{2}\int_0^T \underline{b}_{i,\delta}^T(s)\, \underline{b}_{i,\delta}(s) ds \right\} \tag{14}$$

Then $\Lambda_T^{i,\delta} = d\mu_{i,\delta}/d\mu$. By Itô's rule of differentiation

$$\frac{1}{\delta}\left(\frac{d\mu_{i,\delta}}{d\mu} - 1\right) = \int_0^T \frac{1}{\delta}\Lambda_s^{i,\delta}\, \underline{a}_{i,\delta}^T(ds) d\underline{w}_s +$$
$$+ \int_0^T \frac{1}{\delta}\Lambda_s^{i,\delta}\, \underline{b}_{i,\delta}^T(s) d\underline{\nu}_s \tag{15}$$

Now,

$$E\left(\int_0^T \Lambda_s^{i,\delta}\, \underline{a}_{i,\delta}^T(s)\, \frac{1}{\delta}\, d\underline{w}_s - \int_0^T \underline{a}_i^T(s) d\underline{w}_2\right)^2$$

$$= E\int_0^T \left|\Lambda_s^{i,\delta}\, \underline{a}_{i,\delta}(s)\, \frac{1}{\delta} - \underline{a}_i(s)\right|^2 ds$$

$$\leq 2E\int_0^T (\Lambda_s^{i,\delta} - 1)^2 \left|\frac{1}{\delta}\underline{a}_{i,\delta}\ (s)\right|^2 ds +$$

$$+ 2\int_0^T E\left|\frac{1}{\delta}\underline{a}_{i,\delta}(s) - \underline{a}_i(s)\right|^2 ds$$

The second integral tends to zero by the dominated convergence theorem since $\underline{a}_{i,\delta}(s)/\delta$ and $\underline{a}_i(s)$ are bounded. As for the first integral, since $\underline{a}_{i,\delta}/\delta$ is bounded and since $E\,\Lambda_s^{i,\delta} = 1$

$$E\int_0^T (\Lambda_s^{i,\delta} - 1)^2 \ |\underline{a}_{i,\delta}(s)/\delta|^2 \ ds \le K_1 \ E\int_0^T \{(\Lambda_s^{i,\delta})^2 - 2\Lambda_s^{i,\delta} + 1\}dt$$

$$\le K_1 \int_0^T (e^{\delta^2 KT} - 1)dt$$

Therefore, $\int_0^T \Lambda_s^{i,\delta} \frac{1}{\delta} \underline{a}_{i,\delta}^T(s)d\underline{w}_s$ converges in the mean to $\int_0^T \underline{a}_i^T(s)d\underline{w}_s$ as $\delta\to 0$. By the same argument $\int_0^T \Lambda_s^{i,\delta} \ \underline{b}_{i,\delta}^T(s)d\underline{\nu}_s$ converges in the mean to $\int_0^T \underline{b}_i^T(s)d\underline{w}_s$ which by (15) completes the proof of the lemma.

<u>Lemma 3</u> $\lim_{\delta\to 0} E(\ \ )$ in the left hand side of (10) is equal to $E \lim_{\delta\to 0} (\ \ )$ where the last limit is in probability.

<u>Proof</u> A sufficient condition for the interchange of the limit and the expectation (p.164 of [9]):

$$\sup_{0<\delta\le\delta_0} E\left(\frac{1}{\delta^2}\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)\right)^2 < \infty$$

This will follow from the proof that

$$\sup_{0<\delta\le\delta_0} E\,\frac{1}{\delta^4}\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)^4 < \infty \tag{16}$$

By (14)

$$\frac{1}{\delta^4}\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)^4 \le 2^4 \frac{1}{\delta^4}\left(\int_0^T \Lambda_s^{i,\delta} \ \underline{a}_{i,\delta}^T(s)d\underline{w}_s\right)^4 +$$

$$+ 2^4 \frac{1}{\delta^4}\left(\int_0^T \Lambda_s^{i,\delta} \ \underline{b}^T(s)d\underline{\nu}_s\right)^4$$

By [10] and since the elements of $\underline{a}_{k,\delta}(s)/\delta$ and $\underline{b}_{i,\delta}(s)/\delta$ are bounded

$$E \frac{1}{\delta^4}\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)^4 \le K_2\, E \int_0^T (\Lambda_s^{i,\delta})^4\, ds$$

By Lemma 1, (15) is true which proves the lemma. Returning now to the proof of Theorem 2, by the result of Lemma 2 there exists a sequence $\delta_n \to 0$ for which the convergence result of Lemma 2 holds also almost surely. On a subsequence of the δ_n sequence we will also have almost sure convergence for $\frac{1}{\delta_n{}'}\left(1 - \frac{d\mu_{j,\delta_n{}'}}{d\mu}\right)$. Since $x_n \xrightarrow{a.s.} x$, $y_n \xrightarrow{a.s.} y$ implies $x_n y_n \xrightarrow{a.s.} xy$ we get by Lemma 2

$$\frac{1}{\delta^2}\left(1 - \frac{d\mu_{i,\delta_n{}'}}{d\mu}\right)\left(1 - \frac{d\mu_{j,\delta_n{}'}}{d\mu}\right)$$

$$\xrightarrow{a.s.} \left(\int_0^T \underline{a}_i^T(s)d\underline{w}_s + \int_0^T \underline{b}_i^T(s)d\underline{\nu}_s\right)\left(\int_0^T \underline{a}_j^T(s)d\underline{w}_s + \int_0^T \underline{b}_j^T(s)d\underline{\nu}_s\right)$$

The result of Theorem 3 now follows by Lemma 3.

Theorem 4 Let $\underline{x}_0$ be a random variable with a density $p(\underline{x})$. μ and $\mu_{i,\delta}$ are as in Theorem 1 except that the condition $\underline{\phi}_i(0) = 0$ is not required. Let $\underline{e}$ be any unit m dimensional vector. The density $p(\underline{x})$ is assumed to satisfy

a) $p(\underline{x} + \delta\underline{e})$ is absolutely continuous with respect to $p(\underline{x})$ for all $0 < \delta < \delta_0$

b) $p(\underline{x})$ is differentiable in all it's variables

c) $\left|1 - \frac{p(\underline{x} + \delta\underline{e})}{p(\underline{x})}\right| \le \delta\, f(\underline{x},\delta)$

where $f(\underline{x},\delta)$ satisfies for some $\varepsilon > 0$

$$\sup_{0<\delta<\delta_0} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\underline{x})\; f^{2+\varepsilon}(\underline{x},\delta)\,dx_1 \ldots dx_m < \infty$$

Then

$$\lim_{\delta\to 0} E \frac{1}{\delta^2}\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right)\left(1 - \frac{d\mu_{j,\delta}}{d\mu}\right) =$$

$$= \underline{\phi}_i^T(0)\,(\underline{\sigma}_0^2)^{-1}\underline{\phi}_j(0) + E\Bigg\{\int_0^T (\dot{\underline{\phi}}_i(t) - \dot{\underline{m}}(\underline{x}_t,t)\;\underline{\phi}_i(t))^T (C_t C_t^T)^{-1}\cdot$$

$$\cdot(\dot{\underline{\phi}}_j - \dot{\underline{m}}(\underline{x}_t,t)\;\underline{\phi}_j(t))dt +$$

$$+ \int_0^T (\dot{\underline{g}}(x_t,t)\phi_i(t))^T (D_t D_t^T)^{-1}\dot{\underline{g}}(x_t,t)\cdot$$

$$\cdot\underline{\phi}_j(t)dt\Bigg\}$$

where $(\sigma_0^2)^{-1}$ is the m m matrix

$$(\underline{\sigma}_0^2)^{-1}_{i,j} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(\underline{x})\,\frac{\partial \log p(\underline{x})}{\partial x_i}\,\frac{\partial \log p(\underline{x})}{\partial x_j}\,dx_1,\ldots,dx_m$$

<u>Proof</u>: By [8]

$$\frac{d\mu_{i,k}}{d\mu} = \frac{p(\underline{x} + \delta\,\underline{\phi}_i(0))}{p(\underline{x})}\quad \Lambda_T^{i,\delta}$$

where $\Lambda_T^{i,\delta}$ is as defined by (14). Therefore

$$\frac{1}{\delta}\left(1 - \frac{d\mu_{i,\delta}}{d\mu}\right) = \frac{1}{\delta}\left(1 - \frac{p(\underline{x} + \delta\,\underline{\phi}_i(0))}{p(\underline{x})}\right) + \frac{p(\underline{x} + \delta\,\underline{\phi}_i(0))}{p(\underline{x})}\,\frac{1}{\delta}\,(1-\Lambda_T^{i,\delta})$$

The rest of the proof is essentially the same as in Theorem 2, taking first the conditional expectation given $\underline{x}_0$ and then taking the total expectation. The details are therefore omitted.

Remark The requirement that C_t be nonsingular can be relaxed as follows. Replace $\underline{C}_t^{-1}$ and $(\underline{C}_t\underline{C}_t^T)^{-1}$ by $\underline{C}_t^+$, $(\underline{C}_t,\underline{C}_t^T)^+$ respectively where $(\ )^+$ denotes the pseudo-inverse in all the equations of this section and require that the equation in $\underline{a}_{i,\delta}(s)$

$$-\underline{C}_s\ \underline{a}_{i,\delta}(s) = \underline{m}(\underline{x}_s,s) - \underline{m}(\underline{x}_s - \delta\ \underline{\phi}_i(s),s) - \delta\ \underline{\phi}_i(s)$$

has at least one solution which is continuous in $\underline{x}$, s and δ, $0<\delta<\delta_0$ and that the first line of equation (11) is uniformly bounded for all x, $s \in [0,T)$, $0<\delta\leq\delta_0$.

4. AN EXPRESSION FOR THE ERROR MATRIX ASSOCIATED WITH LINEAR FILTERING

Consider the linear system

$$\begin{aligned} d\underline{u}_t &= \underline{A}_t\ \underline{u}_t\ dt + \underline{C}_t\ d\underline{w}_t \\ d\underline{v}_t &= \underline{B}_t\ \underline{u}_t\ dt + \underline{D}_t\ d\underline{\nu}_t \end{aligned} \tag{17}$$

where $\underline{w}_t$ is a standard m dimensional Brownian motion, $\underline{\nu}_t$ is a standard g dimensional Brownian motion independent of $\underline{w}_t$. $\underline{A}_t$, $\underline{B}_t$, $\underline{C}_t$ $\underline{D}_t$ are m×m, g×m, m×m, g×g matrices respectively with continuous entries. $\underline{D}_t$ is nonsingular for all t in [0,T]. $\underline{v}_0 = 0$, $\underline{u}_0$ is assumed to be Gaussian $E\ \underline{u}_0 = 0$, $E\ \underline{u}_0\underline{u}_0^T = \underline{\sigma}_0^2$, $\underline{\sigma}_0^2$ is assumed to be nonsingular. All expectations in this section are with respect to the measure induced by (17).

Let $\underline{R}_{s,t} = E\ \underline{u}_s\underline{u}_t^T$ and let λ_i, $\underline{\phi}_i(t)$ be the eigenvalues and eigenfunctions associated with $\underline{R}_{s,t}$, $s,t \in [0,T]$. Let α_i denote the coefficients in the Karhunen-Loève expansion of $\underline{u}_t$, then $\underline{u}_t = \sum \alpha_i\ \underline{\phi}_i(t)$ in the mean and a.s., α_i are independent and Gaussian $(0,\lambda_i)$.

Lemma 4 [Parzen]:

Let

$$\underline{v}_t^N = \sum_{i=1}^{N} \alpha_i \, \underline{B}_t \, \underline{\phi}_i(t) + \int_0^t \underline{D}_s \, d\underline{\nu}_s$$

$$\hat{\alpha}_i^N = E(\alpha_i | \underline{v}_\theta^N, \ 0 \le \theta \le T)$$

then, $1 \le i, \ j \le N$

$$E(\alpha_i - \hat{\alpha}_i^N)(\alpha_j - \hat{\alpha}_j^N) = \left((\underline{\Gamma}_N^{-1} + \underline{K}_N)^{-1} \right)_{i,j} \tag{18}$$

where Γ_N is the N×N diagonal matrix $[\underline{\Gamma}_N]_{i,j} = \lambda_i \ \delta_{ij}, \delta_{ij}$ being the Kronecker δ, and K_N is the N×N matrix

$$(\underline{K}_N)_{i,j} = \int_0^T (B_t \ \underline{\phi}_i(t))^T \ (\underline{D}_t \underline{D}_t^T)^{-1} \ \underline{B}_t \ \underline{\phi}_j(t) dt$$

<u>Proof</u>: Equation (18) is a result of Parzen (Theorem 6.1) of [11]) where in [11]

$$(\underline{\Gamma}_N)_{i,j} = \langle \underline{\phi}_i(\cdot), \ \underline{\phi}_j(\cdot) \rangle_{RK1}$$

$$(\underline{K}_N)_{i,j} = \langle \int_0^o \underline{B}_\theta \ \underline{\phi}_i(\theta) d\theta, \int_0^o \underline{B}_0 \ \underline{\phi}_j(\theta) d\theta \rangle_{RK2}$$

where $\langle \cdot, \cdot \rangle_{RK1}$ is the reproducing kernel Hilbert space scalar product associated with the process $\underline{u}_t$ of (17) and $\langle \cdot, \cdot \rangle_{RK2}$ is the reproducing kernel Hilbert space scalar product associated with the process $\int_0^t \underline{D}_s d\underline{\nu}_s$. The explicit expressions for $\underline{\Gamma}_N$, $\underline{K}_N$ follow from Lemmas A-1 and A-2 (Appendix).

It follows from Lemmas A-1 and A-2 (Appendix) and (18) that

$$(\underline{\Gamma}_N^{-1} + \underline{K}_N)_{i,j} = \underline{\phi}_i^T(0) \ (\sigma_0^2)^{-1} \ \underline{\phi}_j(0) +$$

$$+ \int_0^T (\dot{\underline{\phi}}_i(t) - \underline{A}_t \ \underline{\phi}_i(t))(\underline{C}_t \underline{C}_t^T)^+ \ (\dot{\underline{\phi}}_j(t) - \underline{A}_t \ \underline{\phi}_j(t)) dt$$

$$+ \int_0^T (\underline{B}_t \ \phi_i(t))^T \ (D_t D_t^T)^{-1} \ \underline{B}_t \ \underline{\phi}_j(t) dt$$

Theorem 5

$$E(\alpha_i-\hat{\alpha}_i)(\alpha_j-\hat{\alpha}_j) = \lim_{N\to\infty} [(\underline{\Gamma}_N^{-1} + \underline{K}_N)^{-1}]_{i,j}$$

$$\hat{\alpha}_i = E(\alpha_i|\underline{v}_s, 0\leq s\leq T)$$

and $\underline{\Gamma}_N^{-1} + \underline{K}_N$ is as given by (19).

Proof: Consider, first, the case i=j=1. Let $\mathcal{B}_N$, N>1 denote the σ-field generated by $\{\underline{v}_\theta, 0\leq\theta\leq T; \alpha_{N+1}; \alpha_{N+2};\ldots\}$. Then $\hat{\alpha}_1^N = E\{\alpha_1|\mathcal{B}_N\}$ and $E\{\hat{\alpha}_1^N|\mathcal{B}_{N+1}\} = \hat{\alpha}_1^{N+1}$. Therefore, the random sequence $\hat{\alpha}_1^N$ forms a reverse martingale sequence in N and ([9] p.396) $\hat{\alpha}_1^\infty = \lim \hat{\alpha}_1^N$, $N\to\infty$ exist a.s.. Furthermore, $(\hat{\alpha}_1^N)^2$ is a reverse submartingale sequence, therefore, $E(\alpha_1-\hat{\alpha}_1^N)^2 = E\,\alpha_1^2 - E(\hat{\alpha}_1^N)^2$ is nondecreasing with N. Since the $\hat{\alpha}_1^N$ are Gaussian with zero mean on uniformly bounded variance, $E(\alpha_1-\hat{\alpha}_1^N)^2$ increases to $E(\alpha_1-\hat{\alpha}_1^\infty)^2$. It follows that $((\underline{\Gamma}_N^{-1} + \underline{K}_N)^{-1})_{11}$ tends to a limit as $N\to\infty$. It remains to be shown that this limit is $E(\alpha_1-\hat{\alpha}_1)^2$.

Let $\mathcal{B}_\infty$ be the intersection of all the $\mathcal{B}_n$ σ-fields then $\hat{\alpha}_1^\infty = E(\alpha_1|\mathcal{B}_\infty)$. Let $\sigma\{v_0^T\}$ denote the σ-field generated by $\underline{v}_\theta$, $0\leq\theta\leq T$, then obviously, $\mathcal{B}_\infty \subset \sigma\{\underline{v}_0^T\}$. It has to be shown that $\mathcal{B}_\infty = \sigma\{v_0^T\}$. Let $A\in\mathcal{B}_\infty$ then $A\in\mathcal{B}_N$ for all N. Let I_A be the indicator function of A then $I_A = G_{\underline{v}_0^T}(\alpha_{N+1},\alpha_{N+2},\ldots)$ where G is a measurable function of the $\underline{v}_0^T$, α_i, $i\geq N+1$ variables. By the theorem that every section of a measurable function is measurable ([9] p.134) it follows that for a fixed $\underline{v}_0^T$, $G_{\underline{v}_0^T}(\alpha_{N+1},\ldots)$ is a measurable function of α_i, $i\geq N+1$. Furthermore, $\underline{v}_0^T$ fixed, $G_{\underline{v}_0^T}(\alpha_{N+1},\ldots)$ is a tail function of the independent random variables α_i, therefore, by the Kolmogorov zero-one law ([9] p.229) G is trivial so that G is a function of v_0^T only therefore $\mathcal{B}_\infty = \sigma(\underline{v}_0^T)$.

The same proof holds for any $i=j$, $N>i$, the case $i \neq j$ follows along the same lines by considering $\alpha_i \pm \alpha_j$, $N>i,j$.

<u>Remark</u> If $\sigma_0^2 = 0$ then the results of Theorem 5 hold with $\phi_i^T(\underline{\sigma}_0^2)^{-1}\phi_i$ replaced by zero in Equation (19).

5. A LOWER BOUND FOR DIFFUSION PROCESSES

Let $\underline{x}_t$, $\underline{y}_t$ be as in Section 3. Let

$$\underline{A}_t = E\,\underline{\dot{m}}(\underline{x}_t,t) \tag{20}$$

and let $\underline{B}_t$ be a solution to the following equation

$$\underline{B}_t^T\,\underline{C}_t^T(\underline{D}_t\,\underline{D}_t^T)^{-1}\,\underline{C}_t\,\underline{B}_t = E\{(\underline{\dot{m}}(\underline{x}_t,t) - \underline{A}_t)^T(\underline{C}_t\,\underline{C}_t^T)^{-1}\,(\underline{\dot{m}}(x_t,t) - A_t)\} + E\{(\underline{\dot{g}}(x_t,t))^T\,(\underline{D}_t\,\underline{D}_t^T)^{-1}\,\underline{\dot{g}}(\underline{x}_t,t)\} \tag{21}$$

The right hand side of (21) is nonnegative definite matrix, therefore, it has a square root, this assures the existence of a solution to (21) for $\underline{B}_t$. Note that $\underline{A}_t$ and $\underline{B}_t$ are continuous in t.

<u>Theorem 6</u> Let $\underline{x}_t,\underline{y}_t$ satisfy the conditions of Theorem 3 (Theorem 4) consider the linear system (17) where $\underline{A}_t$, $\underline{B}_t$ are as defined by (20) and (21). $\underline{y}_0 = 0$, $\underline{x}_0 = 0$ ($E\,\underline{x}_0 = 0$, $E\,\underline{x}_0\underline{x}_0^T = \underline{\sigma}_0^2$ respectively) then

$$E^{(1)}(x_i(T) - \hat{x}_i(T))^2 \geq E^{(2)}(u_i(T) - \hat{u}_i(T))^2 \tag{22}$$

where $E^{(1)}$ and $E^{(2)}$ denote expectations with respect to the measures induced by (9) and (17) respectively. $x_i(T)$ denotes the i-th component of $\underline{X}(T)$, $\hat{x}_i(T) = E^{(1)}\{x_i(T)|y_t,\ 0\leq t\leq T\}$, $\hat{u}_i = E^{(2)}\{u_i(T)|v_t,\ 0\leq t\leq T\}$. Furthermore, let α be any functional on $L_2[0,T]$ then

$$E^{(1)}(\alpha-\hat{\alpha}^{(1)})^2 \geq E^{(2)}(\alpha-\hat{\alpha}^{(2)})^2 \tag{23}$$

where $\hat{\alpha}^{(1)} = E^{(1)}\{\alpha|y_t,\ 0\leq t\leq T\}$, $\hat{\alpha}^{(2)} = E^{(2)}\{\alpha|v_t,\ 0\leq t\leq T\}$.

Proof: Let $\underline{\phi}_i(t)$ denote the eigenfunctions associated with the $\underline{u}_t$, $0 \le t \le T$ process of Equation (17). By the result of Lemma A-2, the $\underline{\phi}_i(t)$ have differentiable components on [0,T]. Since the measures induced by (9) and (17) are equivalent [8], $\underline{\phi}_i(t)$, i=1,2,... are complete with respect to the $\underline{x}_t$ process of (9). Consider now the space ($\underline{y}_t$, $0 \le t \le T$, $\alpha_1, \alpha_2, \ldots$) where

$$\alpha_i = \int_0^T \underline{x}_s^T \, \underline{\phi}_i(s) ds.$$

Because of (20) and (21), the information matrix associated with the nonlinear system (Equation (10)) is equal to the information matrix associated with the linear system which is the same as Equation (10) but with $\underline{\dot{m}}$, $\underline{\dot{g}}$ replaced by $\underline{A}_t$ and $\underline{B}_t$ respectively. Therefore, by Theorems 1 and 5 the matrix

$$E^{(1)}(\alpha_i - \hat{\alpha}_i^{(1)})(\alpha_j - \hat{\alpha}_j^{(1)}) - E^{(2)}(\alpha_i - \hat{\alpha}_i^{(2)})(\alpha_j - \hat{\alpha}_j^{(2)})$$

is nonnegative. Let $\alpha = \sum_1^N \theta_i \, \alpha_i$ where θ_i, i = 1,...,N are constants then $E(\alpha - \hat{\alpha})^2 = \sum \sum \theta_i \, \theta_j \, E(\alpha_i - \hat{\alpha}_i)(\alpha_j - \hat{\alpha}_j)$. Since $\underline{\phi}_i(t)$ are complete with respect to x_s, $0 \le s \le T$, (23) and (22) follow.

APPENDIX

In this Appendix, we consider the random process defined by

$$d\underline{u}_t = \underline{A}_t \, \underline{u}_t \, dt + \underline{C}_t \, d\underline{w}_t \tag{A-1}$$

where $\underline{u}_0$, $\underline{A}_t$, $\underline{C}_t$, $\underline{w}_t$ are as in the first line of Equation (17). Two lemmas associated with Equation (A-1) will be stated and then proved. The reader is referred to [11] for the basic definitions and results on the reproducing kernel Hilbert space associated with random processes.

Lemma A-1 If C_t is nonsingular, $0 \le t \le T$ and $\underline{f}_i(t)$, i=1,2 are m dimensional vectors with components having bounded derivatives on [0,T], then the reproducing kernel scalar product $\langle \underline{f}_1, f_2 \rangle_{RK}$ associated with (A-1) is given by

$$\langle f_1, f_2 \rangle_{RK} = \underline{f}_1^T(0)(\underline{\sigma}_0^2)^{-1}\, \underline{f}_2(0) + \int_0^T (\dot{\underline{f}}_1(t) - \underline{A}_t \underline{f}_1(t))^T (\underline{C}_t \underline{C}_t^T)^{-1} (\dot{\underline{f}}_2(t) - \underline{A}_t\, f_2(t))dt \qquad \text{(A-2)}$$

and if $[\sigma_0^2] = 0$, $\underline{f}_1(0) = 0$, $\underline{f}_2(0) = 0$, the first term in the right hand side of the last equation is zero. $\dot{f}(t) = df(t)/dt$. Furthermore, if instead of requiring the $\underline{C}_t$ be nonsingular we require that $\underline{f}_1(t)$, $\underline{f}_2(t)$ be in the reproducing kernel Hilbert space associated with (A-1) the (A-2) remains true with $(\underline{C}_t \underline{C}_t^T)^{-1}$ replaced by the pseudoinverse $(\underline{C}_t \underline{C}_t^T)^+$.

Lemma A-2 Let λ_i, $\underline{\phi}_i(t)$ be the eigenvalues and eigenfunctions associated with the Karhunen-Loève expansion of $\underline{u}_t$, $0 \le t \le T$ then the components of $\underline{\phi}_i(t)$ possess continuous derivatives on [0,T] and

$$\frac{1}{\lambda_i}\, \delta_{ij} = \underline{\phi}_i^T(0)\ (\underline{\sigma}_0^2)^{-1}\, \underline{\phi}_j(0) + \int_0^T (\dot{\underline{\phi}}_i(t) - \underline{A}_t\ \phi_i(t))^T\ (\underline{C}_t \underline{C}_t^T)^+ \cdot (\dot{\underline{\phi}}_j(t) - \underline{A}_t\ \phi_j(t))dt \qquad \text{(A-3)}$$

where δ_{ij} is the Kronecker δ and $(\ )^+$ denotes the pseudoinverse. If $\underline{\sigma}_0^2 = 0$ then the same result holds with $\underline{\phi}_i^T(0)(\underline{\sigma}_0^2)^{-1}\, \underline{\phi}_j(0)$ replaced by zero.

Proof of Lemma A-1 Consider the process $\underline{u}_t' = \underline{f}(t) + \underline{u}_t$ then

$$d\underline{u}_t' = (\dot{\underline{f}}(t) - A_t\ \underline{f}(t) + A_t u_t')dt + \underline{C}_t\ d\underline{w}_t$$

The Radon-Nikodym derivative of the measure induced by $\underline{u}_t$ with respect to that of $\underline{u}_t'$ is given by [11]:

$$\operatorname{Exp}\{\langle \underline{u},\underline{f}\rangle_{RK} - \frac{1}{2}\langle \underline{f},\underline{f}\rangle_{RK}\} \quad \text{(A-4)}$$

provided that $\underline{f}(t)$ is in the domain of the reproducing Hilbert space of Equation (A-1). From [7] and [8], it follows that $\underline{f}(t)$ is in the reproducing kernel Hilbert space of (A-1) and by comparing the explicit expression for the R-N derivative given in [7] and [8] with (A-4), we get an explicit expression for $\langle \underline{f},\underline{f}\rangle_{RK}$ from which $\langle \underline{f}_1,\underline{f}_2\rangle_{RK}$ as given by (A-2) follows immediately. If $\underline{C}_t$ is singular for some t or for all t, then the condition that $\underline{f}$ is in the reproducing kernel Hilbert space of (A-1) assures that the measure induced by (A-1) is absolutely continuous with respect to that of $\underline{u}_t'$ [11] and the result of the final part of the lemma follows by comparing (A-4) with the results of [7] and [8].

Proof of Lemma A-2   Let $\underline{\Psi}_t$ be the m×m matrix which solves

$$\frac{d\underline{\Psi}_t}{dt} = \underline{A}_t \underline{\Psi}_t, \ \underline{\Psi}_0 = I .$$

It is easily verified that the solution to Equation (A-1) for t>s is given by

$$\underline{u}_t = \underline{\Psi}_t \underline{\Psi}_s^{-1} \underline{u}_s + \underline{\Psi}_t \int_s^t \underline{\Psi}_\theta^{-1} \underline{C}_\theta \, d\underline{w}_\theta$$

therefore

$$\underline{R}_{s,t} = E\,\underline{u}_s\underline{u}_t^T = E\,\underline{u}_s\underline{u}_s^T \cdot (\underline{\Psi}_s^{-1})^T \underline{\Psi}_t^T .$$

Furthermore, it is easily shown that $E\,\underline{u}_s\underline{u}_s^T = \underline{R}_{s,s}$ satisfies the equation

$$\frac{d\,\underline{R}_{ss}}{dt} = \underline{A}_s\underline{R}_{ss} + \underline{R}_{s,s}\underline{A}_s^T + \underline{C}_s\underline{C}_s^T$$

Note that $\underline{R}_{ss}$, $\underline{\Psi}_s$, Ψ_s^{-1} possess continuous derivatives over [0,T]. For all t in [0,T]

$$\int_0^T \underline{R}_{t,s}\, \underline{\phi}_i(s)ds + \int_t^T \underline{R}_{t,s}\, \underline{\phi}_i(s)ds = \lambda_i\, \underline{\phi}_i(t) \ ,$$

from the left hand side of this equation and the differentiability properties of $\underline{R}_{s,s}$, $\underline{\Psi}_s$, $\underline{\Psi}^{-1}$ it follows that the right hand side possesses a continuous derivative.

The eigenfunctions $\underline{\phi}_i(t)$ are in the reproducing kernel Hilbert space of (A-1) (for $\lambda_i \neq 0$) and [11] $\langle\underline{\phi}_i,\underline{\phi}_j\rangle_{RK} = \lambda_i^{-1}\, \delta_{ij}$. Therefore, (A-3) follows from (A-2).

REFERENCES

1) R. I. Stratonovich, "Conditional Markov Processes," Theory Probab. Appl. (USSR), vol. 5, pp. 156-178, 1960.

2) H. J. Kushmer, "Dynamical Equations for Optimal Nonlinear Filtering," J. Differ. Equations, vol. 2, pp. 179-190, 1967.

3) M. Zakai and J. Ziv, "Lower and Upper Bounds on the Optimal Filtering Error of Certain Diffusion Processes," IEEE Trans. Inf. Theory, vol. IT-18, no. 3, pp. 327-331, May 1972.

4) D. L. Snyder and I. B. Rhodes, "Filtering and Control Performance Bounds with Implication on Asymptotic Separation," Automatica, vol. 8, no. 6, pp. 747-753, November 1972.

5) B. Z. Bobrovsky and M. Zakai, "A Lower Bound on the Estimation Error for Certain Diffusion Processes," submitted for publication.

6) H. L. Van Trees, Detection Estimation and Modulation Theory, Part 1, Wiley, 1968.

7) J. V. Girsanov, "On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures," Theory Prob. Applications, vol. V, no. 3, pp. 285-301, 1960.

8) A. V. Skorokhod, "On the Differentiability of Measure Corresponding to Random Processes, II Markov Processes," Theory Prob. Applications, vol. V, no. 1, pp. 40-49, 1960.

9) M. Loève, Probability Theory, Van Nostrand, Princeton, 1963.

10) M. Zakai, "Some Moment Inequalities for Stochastic Integrals and for Solutions of Stochastic Differential Equations," Israel J. of Math., vol. 5, pp. 170-175, 1967.

11) E. Parzen, "A New Approach to the Synthesis of Optimal Smoothing and Prediction Systems," Mathematical Optimization Techniques, Univ. of California Press, pp. 75-108, 1963. Reprinted in E. Parzen, Time Series Analysis Papers, Holden Day, San Francisco, pp. 443-476, 1967.