Smoothness of the Nonlinear White Noise Filter

by

Hans-Peter Hucke
Purdue University

Technical Report #86-5

Department of Statistics
Purdue University

March 1986

Smoothness of the Nonlinear White Noise Filter

by

Hans-Peter Hucke
Purdue University

ABSTRACT

The nonlinear filtering problem in the finitely additive white noise model is studied and the smooth dependence of the filter on the observations is proved. A Taylor series expansion for the filter is given as well as finite expansions. Orders of convergence of such finite expansions are derived and an application to the discrete sampling problem is discussed.

Key Words and Phrases: Nonlinear filter, white noise model, robustness, series expansion, approximations.

1. Introduction.

The aim of this paper is to continue the study of the robustness properties of the nonlinear filter in the finitely additive white noise model, which was begun in [2].

In [2] it was shown that the white noise filter is a locally Lipschitz continuous function of the observations.

Our main result, which will be proved in section 3 shows that the filter is in fact a smooth function of the observed sample paths, i.e. infinitely differentiable is the Fréchet-sense.

As a consequence, we can give a Taylor series expansion for the filter which in turn can be used to give bounds on the approximation error when using only a finite series expansion.

The latter results can be regarded as a 'deterministic' counterpart to theorems of Ocone [9], who used series of multiple stochastic integrals to express the nonlinear filter in the conventional, stochastic calculus model.

In the following we shall always assume that we are dealing with the following (white noise) model for the signal and observation processes.

Let $\mathbf{T} = [0, T]$ be a finite time interval, also let S be a complete, separable metric space and denote by $\mathcal{S}$ its Borel σ-field.

We take the signal process $(X_t)_{t \epsilon \mathbf{T}}$ to be a progressively measurable S-valued Markov process, defined on a probability space $(\Omega, \mathcal{A}, \Pi)$.

By $\mathbb{K}$ we denote a (possibly infinite dimensional) real Hilbert space with inner product $< \cdot, \cdot >$ and norm $|\cdot|$.

$\mathbb{K}$ will be the state space of the observation process and for most practical applications can be taken as a finite-dimensional Euclidean space $\mathbb{R}^m$.

Suppose h is a nonlinear function mapping the Cartesian product $\mathbf{T} \times S$ into $\mathbb{K}$. We assume further that h is measurable and that is satisfies the condition

$$E \int_0^T |h_t(X_t)|^2 dt < \infty \tag{1.1}$$

Next we define a map ξ from Ω into $L^2(\mathbf{T}, \mathbb{K}) = \left\{ f : \mathbf{T} \to K : \int_0^T |f(t)|^2 \, dt < \infty \right\}$ by

$$\xi_t(\omega) := h_t(X_t(\omega)) \tag{1.2}$$

This map will play the role of the signal in the white noise model.

The observation noise e is modeled as a random element in the Hilbert space L^2 which has the canonical Gaussian distribution. By this we mean that the distribution of e corresponds to the characteristic functional

$$C_e(\rho) = \exp\left\{ \frac{1}{2} (\rho, \rho) \right\} \tag{1.3}$$

defined for all $\rho \epsilon L^2$.

Here and in the sequel $(\cdot, \cdot)$ will denote the inner product in L^2, given by

$$(\rho, \psi) = \int_0^T < \rho(u), \ \psi(u) > \ du \tag{1.4}$$

for $\rho, \psi \epsilon L^2$. The associated norm will henceforth be denoted by $\|\cdot\|$.

It is now easily seen from (1.3) that the measure induced on L^2 by e is not countably additive. Thus writing the observation model as

$$y = \xi + e \tag{1.5}$$

it is apparent that y is not defined on a countably additive probability space.

To this end Kallianpur and Karandikar [4] have introduced the concept of quasi-cylindrical measures and mappings, and have given a rigorous definition of (1.5) in this context. For further details see also [7].

Equation (1.5) is generally referred to as the white noise model and is to be viewed as a model in L^2.

The white noise model is to be interpreted in the following way: Observations (y) are obtained in the form of a 'useful' signal which is contaminated by an additive Gaussian noise.

On the basis of this model we now want to estimate functions of the signal proccess (X_t) given observations $y \epsilon L^2$.

This problem was considered in [4]–[7] and we state here for reference only the Bayes formula, which gives the fundamental solution to the estimation problem.

Let $H_t := \left\{ \rho \epsilon L^2 \colon \int_t^T |\rho(u)|^2 du = 0 \right\} \subset L^2$ and denote by $Q_t y$ the orthogonal projection of y onto H_t.

Theorem (1.6): Let $t \epsilon \mathbf{T}$ be fixed and let g be a random variable, defined on $(\Omega, \mathcal{A}, \Pi)$ such that $E|g| < \infty$.

Then the conditional expectation $E[g|Q_t y]$ exists and is given by

$$E[g|Q_t y] = \frac{\sigma_t(g, Q_t y)}{\sigma_t(1, Q_t y)} \quad \text{(Bayes Formula)} \tag{1.7}$$

where

$$\sigma(g, Q_t y) = \int g(\omega) \cdot q_t(y, \omega) \, \Pi(d\omega) \tag{1.8}$$

with

$$q_t(y, \omega) = \exp\left\{ (Q_t y, Q_t \xi(\omega)) - \frac{1}{2} \parallel Q_t \xi(\omega) \parallel^2 \right\}. \tag{1.9}$$

$E[g|Q_t y]$ is the optimal mean square estimator of g given observations $Q_t y = (y_u)_{0 \le u \le t}$.

Remarks:

(i) The conditional expectation $E[g|Q_t y]$ is not a random variable in the ordinary sense of the word, but rather a so-called quasi cylindrical mapping. For details see [7].

(ii) The term $\sigma_t(g, Q_t y)$ is called the unnormalized conditional expectation and is frequently more convenient to work with.

(iii) The problems of nonlinear filtering, prediction and smoothing are contained in theorem (1.6). These correspond to the cases when $g = f(X_s)$ for some function $f \colon S \to \mathbb{R}$ and $s = t$ (respectively $s > t$ and $s < t$).

Formula (1.7) is a white noise analogue to the so-called Kallianpur-Striebel formula of the stochastic calculus approach to nonlinear filtering, see e.g., [3], [8].

We now want to study the properties of the estimator $E[g|Q_t y]$ as a function of $y \epsilon L^2$.

This will be done in section 3, but before we recall some definitions and lemmas from the differential calculus on Banach spaces is the next section.

2. Some Auxilary Results

Since it is our aim to prove the strong differentiability of $E[g|Q_t y]$ with respect to $y \epsilon L^2$, we recall in this section some results from the theory of differential calculus in Banach spaces. All of these definitions and lemmas are stated without proof and can be found in [1].

Definition (2.1): Let E and F be real, separable Banach spaces with norms $\|\cdot\|$ and $|\cdot|$ respectively.

(i) $L(E,F)$ is defined as the (Banach-) spaces of all bounded linear operators between E and F.

The operator norm on $L(E,F)$ will again be denoted by $\|\cdot\|$.

(ii) For any $n \geq 1$ we denote by E^n the n-fold Cartesian product of E.

$L_n(E,F)$ is then defined as the (Banach-) space of all multilinear continuous maps from E^n into F.

Definition (2.2): Let U be an open subset of E and let $f\colon U \to F$.

The map f is called (Fréchet-) differentiable at a point $f'(a) \epsilon L(E,F)$ such that

$$\lim_{x \to a} \frac{|f(x) - f(a) - f'(a) \cdot (x-a)|}{\| x - a \|} = 0 \tag{2.3}$$

If f is differentiable at a, then the operator $f'(a)$ is called the derivative of f at a. f is called differentiable in U if (2.3) holds for all $a \epsilon U$.

We also need to work with derivatives of higher order which can be recursively defined in the following way.

Definition (2.4): Let V be an open neighborhood of a in E. The map f is called n times differentiable at a if f is $n-1$ times differentiable at a and the map $x \to f^{(n-1)}(x)$ from V into $L_{(n-1)}(E,F)$ is differentiable at the point a.

We denote the derivative of order n by $f^{(n)}$; it can be identified with an element of $L_n(E,F)$. If $f^{(n)}$ exists and the map $f^{(n)}\colon V \to L_n(E,F)$ is continuous we will say that f is of class C^n in V.

When f is of class C^n for all $n \geq 1$ we call f a C^∞ map.

The following two lemmas will be particularly useful for our purposes.

Lemma (2.5): Let E, F and G be Banach spaces and U, V be open sets in E and F respectively.

Let

$$f\colon\ U \to F$$

and

$$g\colon\ V \to G$$

If $a \epsilon U$ is such that $f(a) \epsilon V$ and f is differentiable at a and g is differentiable at $f(a)$ then

$$h\colon = g \circ f\colon\ U \to G$$

is differentiable at a with derivative

$$h'(a) = g'(f(a)) \circ f'(a) \tag{2.6}$$

Lemma (2.7):

(i) If f is the restriction of a bounded linear operator, i.e. $f(x) = Ax$ for $x \epsilon U$ with $A \epsilon L(E,F)$, then f is differentiable in U with derivative $f'(a) = A$ for all $a \epsilon U$.

(ii) If f is constant on U, then the derivative of f vanishes identically.

The last result we state here provides the means to interchange the order of differentiation and integration and is formulated with the intended application to the filtering problem in mind.

Lemma (2.8): Let $f\colon\ E \times \Omega \to \mathbb{R}$ be a measurable function, such that for all $\omega \epsilon \Omega$ the derivative

$$f'(x,\omega) = \frac{\partial}{\partial y}\ f(x,\omega)$$

exists and is locally bounded for almost every ω. Then the the map $g\colon\ E \to \mathbb{R}$

$$g(x) := \int f(x,\omega)\ \Pi(d\omega)$$

is differentiable and the derivative at $a \epsilon E$ is the operator defined by

$$g'(a) \circ x = \int f'(a,\omega) \circ x\ \Pi(d\omega) \qquad \text{for } x \epsilon E \tag{2.9}$$

Proof: Let $U_r(a)$ be the open ball of radius r with center at $a \epsilon E$.

The local boundedness of f' then gives

$$\| f'(x,\omega) \| \leq K$$

for all $x \epsilon U_r(a)$, almost every ω and some constant K. By the mean value theorem this implies

$$|f(x,\omega) - f(y,\omega)| \leq K \cdot \| x - y \|$$

for $x, y \epsilon U_r(a)$.

Hence for $x \epsilon U_r(a)$

$$\int \frac{|f(x,\omega) - f(a,\omega) - f'(a,\omega)\ (x-a)|}{\| x - a \|}\ \Pi(d\omega) \leq \int 2 \cdot K\ \Pi(d\omega) < \infty$$

Thus we can apply the dominated convergence theorem to arrive at

$$\begin{aligned} &\lim_{x \to a} \frac{|\int f(x,\omega)\Pi(d\omega) - \int f(a,\omega)\Pi(d\omega) - \int f'(a,\omega)(x-a)\Pi(d\omega)|}{\| x - a \|} \\ (2.10) \qquad &\leq \int \lim_{x \to a} \frac{|f(x,\omega) - f(a,\omega) - f'(a,\omega)(x-a)|}{\| x - a \|}\ \Pi(d\omega) = 0 \end{aligned}$$

since f is differentiable at a.

This establishes the differentiability of g and at the same time shows that the derivative is of the form (2.9). □

3. Smoothness of the Nonlinear Filter

Although we are mainly interested in the nonlinear filter $E[f(X_t)|Q_t y]$ we will phrase the result for the more general estimator $E[g|Q_t y]$, when g is an integrable, real-valued random variable.

We can now state our main result concerning the smoothness of the estimator $E[g|Q_t y]$ in the white noise model.

Theorem (3.1): If g: $\Omega \to \mathbb{R}$ and h: $\mathbf{T} \times S \to \mathbb{K}$ are bounded, then $E[g|Q_t y]$ is of class C^∞ with respect to $y \epsilon L^2$.

We will prove (3.1) in three steps: First we will show the C^∞-property for $q_t(y,\omega)$, then using (2.8) we prove that $\sigma_t(g, Q_t y)$ is of class C^∞. Finally the Bayes formula for $E[g|Q_t y]$ will allow the application of (2.5) and complete the proof.

Lemma (3.2): The map $q_t(y,\omega)$ is of class C^∞ for all $\omega \epsilon \Omega$ and $t \epsilon \mathbf{T}$.

Proof: Let $t \epsilon \mathbf{T}$, $\omega \epsilon \Omega$ be arbitrary. Since $(Q_t y, Q_t \xi(\omega))$ is a bounded, linear map from L^2 into $\mathbb{R}$, its derivative at any point $a \epsilon L^2$ is the linear operator defined by

$$(Q_t \cdot, Q_t \xi(\omega)).$$

By the composition rule (2.5) the derivative $q_t'(a,\omega)$ is therefore the operator

$$(3.3) \qquad q_t'(a,\omega) = \exp\left\{(Q_t a, Q_t \xi(\omega)) - \frac{1}{2}\ \| Q_t \xi(\omega) \|^2\right\} \cdot (Q_t \cdot, Q_t \xi(\omega))$$

From this form of the derivative it is easy to see that for any $n \geq 1$ the n-th derivative of q at $a \epsilon L^2$ is given by

$$q_t^{(n)}(a,\omega) = \exp\left\{(Q_t a, Q_t\xi(\omega)) - \frac{1}{2} \parallel Q_t\xi(\omega) \parallel^2\right\} \cdot \underbrace{(Q_t\cdot, Q_t\xi(\omega)) \cdot \cdots \cdot (Q_t\cdot, Q_t\xi(\omega))}_{\text{n-times}} \tag{3.4}$$

which is an element of $L_n(L^2, \mathbb{R})$.

This completes the proof. □

Lemma (3.5):

(i) Let $n \geq 1$ and assume that

$$E[|g(\omega)| \cdot \parallel \xi(\omega) \parallel^n] < \infty \tag{3.6}$$

then the unnormalized conditional expectation $\sigma_t(g, Q_t y)$ is of class C^n in y.

(ii) If condition (3.6) is replaced by

$$E[|g(\omega)| \cdot \exp\{\parallel \xi(\omega) \parallel\}] < \infty \tag{3.7}$$

then $\sigma_t(g, Q_t y)$ is of class C^∞ in y.

Proof: From lemma (3.2) it follows that the n-th derivative of $g(\omega) \cdot g_t(y,\omega)$ at $a \epsilon L^2$ is given by

$$f^{(n)}(a,\omega) := g(\omega) \cdot q_t(a,\omega) \cdot (Q_t\cdot, Q_t\xi(\omega)) \cdot \cdots \cdot (Q_t\cdot, Q_t\xi(\omega)) \tag{3.8}$$

Now let $r > 0$ and $x \epsilon U_r(a)$ then

$$|f^{(n)}(x,\omega) \circ y| \leq \exp\left\{\frac{1}{2} \parallel x \parallel^2\right\} \cdot |g(\omega)| \cdot \parallel y \parallel^n \cdot \parallel \xi(\omega) \parallel^n \tag{3.9}$$

for all $y \epsilon L^2$.

By virtue of (3.6) the random variable $|g(\omega)| \cdot \parallel \xi(\omega) \parallel^n$ is a.s. finite so we can deduce from (3.9) that the derivative $f^{(n)}(a,\omega)$ is for almost all ω locally bounded.

We can thus apply lemma (2.8) to the map

$$\sigma_t(g, Q_t y) = \int g(\omega) \cdot g_t(y,\omega)\, \Pi(d\omega)$$

to conclude that its n-th order derivative $\sigma_t^{(n)}(y, Q_t a)$ exists for every $a \epsilon L^2$ and is given by the operator

$$\int g(\omega) \cdot g_t(a,\omega) \cdot (Q_t\cdot, Q_t\xi(\omega)) \cdot \cdots \cdot (Q_t\cdot, q_t\xi(\omega)) \Pi(d\omega). \tag{3.10}$$

Since $\sigma_t^{(n)}(g, Q_t a)$ obviously depends continuously on $a, \sigma_t(g, Q_t y)$ is of class C^n.

The proof of (ii) proceeds in the same way. We only have to note that condition (3.7) implies (3.6) for all $n \geq 1$. □

The Bayes formula provides us now with the obvious means to prove the smoothness of $E[g|Q_t y]$ by using lemma (3.5) for the numerator and denominator separately.

We now restate theorem (3.1) in a slightly more general form.

Theorem (3.11):

(i) Let $n \geq 1$ and assume that (3.6) holds. If in addition

(3.12) $$E[\| \xi(\omega) \|^n] < \infty$$

then $E[g|Q_t y]$ is of class C^n in y.

(ii) If (3.7) holds and furthermore

(3.13) $$E[\exp\{\| \xi(\omega) \|\}] < \infty$$

then $E[g|Q_t y]$ is of class C^∞.

Proof: We only prove (i), since part (ii) follows by the same arguments outlined in the proof of lemma (3.5). Conditions (3.6) and (3.12) imply that both numerator and denominator are of class C^n.

Setting

$$u(y) := \frac{1}{\sigma_t(1, Q_t y)}$$

and

$$v(y) = \sigma_t(g, Q_t y)$$

We can express the n-th derivative of $E[g|Q_t y]$ at a point a by means of the Leibniz-formula

$$(v \cdot u)^{(n)}(a) = \sum_{j=0}^{n} \binom{n}{j} v^{(j)}(a) \cdot u^{(n-j)}(a)$$

which is valid, since $\sigma_t(1, Q_t y)$ is bounded away from 0 (cf. [2]).

Here we define $v^{(0)} = v$ and $u^{(0)} = u$. □

Remark: The proof of theorem (3.1) now follows from (3.11). Since the boundedness of g and h imply conditions (3.7) and (3.13).

An explicit computation of higher derivatives of $E[g|Q_t y]$ is not warranted, because of their complicated structure.

It is more convenient to work only with the unnormalized estimator $\sigma_t(g, Q_t y)$.

Our goal now is to utilize the smoothness of $E[g|Q_t y]$ to develop a series representation for the filter and establish orders of convergence of finite expansions.

4. Taylor Series Expansion

From now on we assume that the assumptions of theorem (3.1) are satisfied.

First we need to establish the following result:

Lemma (4.1): For every $n \geq 1$ the derivative $\sigma_t^{(n)}(g, Q_t y)$ is locally bounded.

Proof: Set $f(y) := \sigma_t(g, Q_t y)$ and let $a \epsilon L^2$ and $r > 0$ be fixed. Define $U_r(a) := \{x \epsilon L^2 : \| x - a \| < r\}$. Evaluating the n-th order derivative (3.10) at $x_1, \ldots, x_n \epsilon L^2$ we observe that

$$|\int g(\omega) \cdot q_t(a, \omega) \cdot (Q_t x_1, Q_t \xi(\omega)) \cdot \cdots \cdot (Q_t X_{n1} Q_t \xi(\omega)) \, \Pi(d\omega)|$$

$$\leq \exp\left\{\frac{1}{2} \| a \|^2\right\} \cdot \| x_1 \| \cdot \cdots \cdot \| x_n \| \cdot \int |g(\omega)| \cdot \| Q_t \xi(\omega) \|^n \; \Pi(d\omega)$$

Due to the boundedness of g and h the latter integral can be bounded by a constant C. If we furthermore define

$$K_r := \sup_{x \epsilon U_r(a)} \exp\left\{\frac{1}{2} \| x \|^2\right\}$$

the relationship

$$|f^{(n)}(x)(x_1, \ldots, x_n)| \leq C \cdot K_r \cdot \| x_1 \| \cdot \cdots \cdot \| x_n \|$$

holds for all $x \epsilon U_r(a)$.

Considering $f^{(n)}$ as an element of $L_n(L^2, \mathbb{R})$ this means that

$$\| f^{(n)}(x) \| \leq C \cdot K_r \tag{4.2}$$

for all $x \epsilon U_r$ and thus completing the proof. □

We are now in a position to state the general Taylor series expansion for the unnormalized conditional expectation.

Theorem (4.3): Let $a \epsilon L^2$ be given and let U be a convex neighborhood of a containing the point $a + \delta$, $\delta \epsilon L^2$.

Under the conditions of theorem (3.1) the unnormalized conditional expectation $f(y) := \sigma_t(g, Q_t y)$ has the Taylor series expansion

$$f(a + \delta) = f(a) + \sum_{n=1}^{\infty} \frac{1}{n!} f^{(n)}(a) \circ (\delta)^n$$

Here $(\delta)^n$ denotes the vector $\underbrace{(\delta,\cdots,\delta)}_{\text{n-times}} \epsilon \bigotimes_{i=1}^{n} L^2$

Proof: By theorem (3.1) f is of class C^∞. From lemma (4.1) we can deduce that all derivatives of f are locally bounded by

$$\| f^{(n)}(y) \| \leq K \cdot C^n \qquad \text{for all } y \epsilon U$$

with

$$K = \sup_{y \epsilon U} \exp\left\{\frac{1}{2} \| y \|^2\right\} \cdot \int |g(\omega)| \, \Pi(d\omega)$$

and

$$C = \sup\{|h_t(x)|: \; t\epsilon \mathbf{T}, x \epsilon S\}.$$

We can thus apply the Taylor formula with Lagrange remainder (c.f., theorem 5.6.2 of [1]) which yields

$$|f(a+\delta) - f(a) - \sum_{n=1}^{m} \frac{1}{n!} f^{(n)}(a) \circ (\delta)^n| \leq K \cdot C^{m+1} \frac{\| \delta \|^{m+1}}{(m+1)!}.$$

But in this form it is obvious, that as $m \to \infty$ the Taylor series converges and formula (4.4) holds, which completes the proof. □

Since the function $g \equiv 1$ is clearly bounded, we can thus give an expression of the conditional expectation as the ratio of two convergent Taylor series. To avoid messy notation we define $k(y) := \sigma_t(1, Q_t y)$.

Theorem (4.5): Under the same assumptions as theorem (4.3) the representation

$$E[g|Q_t(a+\delta)] = \frac{f(a) + \sum_{n=1}^{\infty} \frac{1}{n!} f^{(n)}(a) \circ (\delta)^n}{k(a) + \sum_{n=1}^{\infty} \frac{1}{n!} k^{(n)}(a) \circ (\delta)^n} \tag{4.6}$$

for the estimator of g in the white noise model holds.

Remark (4.7): It is helpful for applications to phrase (4.6) in a slightly different form. Let y and z be two elements of L^2 and suppose that U is a convex set that includes both y and z. Setting $\delta = y - z$ we can then write (4.6) in the form

$$E[g|Q_t y] = \frac{f(z) + \sum_{n=1}^{\infty} \frac{1}{n!} f^{(n)}(z) \circ (y-z)^n}{k(z) + \sum_{n=1}^{\infty} \frac{1}{n!} k^{(n)}(z) \circ (y-z)^n} \tag{4.8}$$

We also remark at this point that there has been related work on series expansions of the optimal nonlinear filter in the stochastic calculus approach to filtering by Ocone [9].

The series representations derived in [9] are however expansions in terms of multiple stochastic integrals and are therefore inappropriate for pathwise computations of the filter.

Formula (4.8) on the other hand is clearly in a pathwise form, allowing the computation for every sample path $y \epsilon L^2$. This is due to the inherent pathwise nature of the white noise model.

5. Applications

Of particular interest for applications will be the approximation of the estimates by a finite term Taylor series expansion.

In this respect we have the following result:

Theorem (5.1): Let U be a convex set containing the points y and $z \epsilon L^2$ and suppose that condition (3.6) holds. Then

$$f(y) = f(z) + \sum_{j=1}^{n} \frac{1}{j!} f^{(j)}(z) \circ (y-z)^j + o(\| y - z \|^n) \tag{5.2}$$

here we have again defined $f(y) = \sigma_t(g, Q_t y)$. The asymptotic term $o(\| y - z \|^n)$ is to be understood for $\| y - z \| \to 0$.

The proof is an immediate consequence of the fact that $\sigma_t(g, Q_t y)$ is of class C^n and the Taylor formula with Lagrange remainder (c.f., theorem 5.6.3 of [1]).

To illustrate the application of (5.1) for the nonlinear filtering problem we now give the explicit expresion for $n = 1$ and $g(\omega) = f(X_t(\omega))$, for some function of f.

Corollary (5.3): If $E\left[|f(x_t)| \cdot \int_0^T |h_u(X_u)|^2\, du\right] < \infty$ then

$$\begin{aligned} \sigma_t(g(X_t), Q_t y) &= \sigma_t(f(X_t), Q_t z) + \\ &+ \int \left[f(X_t(\omega)) \cdot \exp\left\{ \int_0^t < z_u, h_u(X_u(\omega)) > - \frac{1}{2} |h_u(X_u(\omega))|^2 du \right\} \right. \\ &\left. \cdot \int_0^t < y_u - z_u, h_u(X_u(\omega)) > du \right] \Pi(d\omega) + o(\| y - z \|) \end{aligned} \tag{5.4}$$

The merit of (5.4), and its more general version (5.2), is that it allows a qualitative asessment of the error incurred by approximating the unnormalized conditional expectation via a finite Taylor series expansion around a given point z.

Establishing a result analogous to theorem (5.1) for the normalized estimator $E[g|Q_t y]$ is in principle no difficulty.

In fact if assumptions (3.6) and (3.12) hold, in which case $E[g|Q_t y]$ is of class C^n, the same relation (5.2) will be valid with $f(y) = E[g|Q_t y]$.

However, as mentioned in the proof of theorem (3.11), the derivatives of $E[g|Q_t y]$ are far too complicated to make a finite expansion useful.

We only give here the version for $g = f(X_t)$ and $n = 1$ corresponding to corollary (5.3).

Corollary (5.5): If $\max\left\{E\left[\int_0^T |h_u(X_u)|^2 du\right]; E\left[|f(X_y)| \cdot \int_0^T |h_u(X_u)|^2 du\right]\right\} < \infty$, then

$$E[f(X_t)|Q_t y] = E[f(X_t)|Q_t z] +$$

$$+ \frac{1}{\sigma_t(1, Q_t z)} \cdot \left\{ \int \left[f(X_t(\omega)) \cdot q_t(z,\omega) \cdot \int_0^t < y_u - z_u, h_u(X_u(\omega)) > du \right] \Pi(d\omega) \right.$$

$$\left. - E[f(X_t)|Q_t z] \cdot \int \left[q_t(z,\omega) \cdot \int_0^t < y_u - z_u, h_u(X_u(\omega)) > du \right] \Pi(d\omega) \right\}$$

$$+ o(\| y - z \|)$$

For practical applications it may therefore be better to only expand the numerator of the Bayes formula in a finite Taylor series around z and compute the denominator $\sigma_t(1, Q_t y)$ for the original observation.

This way the order of convergence will be maintained, but the approximation will be in a more tractable form.

In closing we point out that a prominent application of theorem (5.1) is its usefulness for the discrete sampling problem.

In most practical situations, especially when working with digital equipment, it is not possible to obtain the whole sample path y itself. Rather values of y at discrete sampling points $0 = t_0, t_1, \ldots, t_N = T$ are obtainable. Denote the values of y at these points by $y_0, y_1, \ldots, y_N$.

On the basis of this information one can then construct 'natural' approximations to y of the form

(5.6) $$z_1(t) = \sum_{j=0}^{N} I_{[t_j, t_{j+1})}(t) \cdot y_j + I_{\{t_N\}}(t) \cdot y_N$$

(step function approximation), or

(5.6) $$z_2(t) = y_{j-1} + \frac{y_j - y_{j-1}}{t_j - t_{j-1}} \cdot (t - t_{j-1}) \qquad \text{for } t \epsilon [t_{j-1}, t_j]$$

(polygonal approximation).

Since both z_1 and z_2 are clearly L^2 functions, we can apply the previous results to this case and give approximations to $\sigma_t(f(X_t), Q_t y)$ by a finite Taylor expansion around z_1 or z_2.

6. References:

[1] CARTAN, H. (1971). Differential calculus. Herman, Paris.

[2] HUCKE, H.-P. (1985). Estimation of continuous time Markov processes in a finitely additive white noise model. CSP Technical Report No. 121, Dept. of Statistics, Univ. of North Carolina.

[3] KALLIANPUR, G. (1980). Stochastic filtering theory. Springer, Berlin.

[4] KALLIANPUR, G. and KARANDIKAR, R. L. (1983). A finitely additive white noise approach to nonlinear filtering. *Appl. Math. and Optimization* (10), pp. 159–185.

[5] KALLIANPUR, G. and KARANDIKAR, R. L. (1984). Measure valued equations for the optimal filter in finitely additive nonlinear filtering theory. *Z. W.* (66), pp. 1–17.

[6] KALLIANPUR, G. and KARANDIKAR, R. L. (1984). The non-linear filtering problem for the unbounded case. *Stoch. Proc. and Applications* (18), pp. 57–66.

[7] KALLIANPUR, G. and KARANDIKAR, R. L. (1985). White noise calculus and nonlinear filtering theory. *Annals of Probability* (13), pp. 1033–1107.

[8] LIPSTER, R. S. and SHIRYAYEV, A. N. (1977, 1978). Statistics of random processes, vls. I and II. Springer, New York.

[9] OCONE, D. (1983). Multiple integral expansions for nonlinear filtering. *Stochastics* (10), pp. 1–30.