

isid/ms/2003/08

January 28, 2003

<http://www.isid.ac.in/~statmath/eprints>

# Linear Filtering with Ornstein Ulhenbeck Process As Noise

ABHAY G. BHATT

Indian Statistical Institute, Delhi Centre  
7, SJSS Marg, New Delhi-110 016, India

# Linear Filtering with Ornstein Ulhenbeck Process As Noise

ABHAY G. BHATT

*Indian Statistical Institute, New Delhi.*

## ABSTRACT

We consider a linear filtering model (with feedback) when the observation noise is an Ornstein-Ulhenbeck process with parameter  $\beta$ . The coefficients appearing in the model are all assumed to be bounded. In addition, the coefficients appearing in the observation equation are also assumed to be differentiable. We consider the genral case when the OU noise is also correlated with the signal. Under these conditions we derive the filtering equations for the optimal filter.

---

*AMS 2000 subject classification:* Primary 60G35 Secondary 60G15, 60H15.

*Key words and phrases:* Linear filtering, Ornstein-Ulhenbeck Process, Optimal Filter.

# 1 Introduction

The theory of filtering can be explained as follows. The process of interest  $X$ , called the Signal, is not directly observable. However, it is assumed that at time  $t$  a (known) function of  $\{X(s) : s \leq t\}$  is observable in the presence of an additive noise.

The question of interest is to *estimate* the signal  $X(t)$  given the observations  $\{Y(s) : 0 \leq s \leq t\}$ . This is known as filtering the noise (to recover the signal). The best estimate, called the *optimal filter* is the conditional distribution of  $X(t)$  given the observation  $\sigma$ -field  $\mathcal{F}_t^Y = \sigma\{Y(s) : 0 \leq s \leq t\}$ . We will denote this conditional distribution by  $\pi(t)$ .

In the classical theory of filtering, the observation noise is modelled to be a Brownian motion. Then, under fairly general conditions on the signal and on the observation function, the filter  $\pi$  solves an infinite dimensional stochastic differential equation called the Fujisaki-Kallianpur-Kunita (FKK) equation (Fujisaki *et. al.* (1972)). Uniqueness of solution to the FKK equation is also known (under some integrability conditions) when the signal-observation pair  $(X, Y)$  is uniquely determined either via a martingale problem (Bhatt *et.al.* (1995)) or as a solution of a (non-anticipative) stochastic differential equation (Bhatt and Karandikar (1999)).

The assumption of the observation noise being a Brownian motion is in some sense a natural one to make. (see *e.g.* the Introduction in Bhatt and Karandikar (2002)). At the same time it has given rise to a very rich theory as mentioned in the previous paragraph. (See also Elliott (1982) and Kallianpur (1980)).

However this model has been objected to by engineers from a practical point of view. The actual observed paths of the accumulative observation process  $Y$  are smooth. But the classical model gives zero probability to all such smooth paths. (See Kallianpur and Karandikar (1988) for a detailed discussion on this.)

Recently, several authors have considered the filtering model when the observation noise is a process other than a Brownian motion. In Kunita (1993), the author initiated the study of filtering theory with general Gaussian noise processes. This was also followed up in Gawarecki and Mandrekar (2000) and Mandrekar and Mandal (2000).

A special case of the above is when the noise is an Ornstein-Uhlenbeck process (OUP) with parameter  $\beta > 0$ . Such a filtering model answers the criticism by engineers of the classical model. Moreover, when  $\beta$  is large, the OUP approximates Brownian motion. Thus a filtering model with OUP noise approximates the classical model.

The optimal filter  $\pi^\beta$  for the model with OUP noise has been studied. This is done by looking at the Bayes' formula for the filter in Mandrekar and Mandal (2000) and Bhatt and Karandikar (2002). The filtering equation has been studied in Gawarecki and Mandrekar (2000), Bhatt *et.al.* (2002) and Kouritzin and Xiong (2002). However each of these articles use different assumptions on the model, in particular, on the function  $h$ . *e.g.* in Gawarecki and Mandrekar (2000) and Mandrekar and Mandal (2000) the authors require that  $h(X(t))$  is almost surely differentiable in  $t$ , (a very stringent condition). A perturbed model (with a small time-lag in recording the observations) is used in Bhatt and Karandikar (2002). Bhatt

*et.al.* (2002) and Kouritzin and Xiong (2002) require the observation function  $h$  to be bounded and smooth. Also, a point to note is that all the above mentioned articles assume that the observation noise and the signal are independent.

In this article we will consider a very important special case - linear filtering. In the classical set-up this was first considered by Kalman and Bucy (Kalman (1960), Kalman and Bucy (1961)). The Kalman-Bucy filter is also the most widely used filter because of its simplicity. In this case the optimal filter is *finite dimensional* in the sense that  $\pi$  is characterized as the unique solution of a system of *two* SDE's (as opposed to infinitely many in the general non-linear filtering problem).

We will consider the linear filtering model with OUP as noise. The model (given by (3.7)–(3.8)) is very general in the sense that it allows feedback (from the observations into the signal). Also we consider the case when the observation noise is correlated with the noise driving the signal. We however assume that the (non-random) coefficients that appear in the model are all bounded and the ones appearing in (3.8) are also differentiable. Under these conditions the optimal filter continues to be finite dimensional and we derive the corresponding filtering equations.

The paper is organised as follows. In the next section we consider the corresponding classical linear filtering model and write down the filtering equations (Theorem 2.1). Also, for later use, we note down a result on FKK equations from the general theory of nonlinear filtering theory (Theorem 2.2).

In Section 3, we introduce the linear filtering model for the signal-observation process with OUP as the observation noise. We derive the equations of filtering for the optimal filter (Theorem 3.2). This is done by effecting a series of transformations on the model to recast it as a classical model of filtering (*i.e.* one where the observation noise is a certain Brownian motion). However, in doing so, the terms appearing in the SDE for the *transformed* signal-observation pair process depend on the entire past of the process. This necessitates the use of results from general non-linear filtering theory. We also show that the filtering equations admit a unique solution.

For notational simplicity, we will consider the one dimensional case.

## 2 The Classical Model

We start with a general linear filtering model (allowing feedback) in the classical set-up. The signal process  $X$  and the observation process  $Y$  are given by the system of stochastic differential equations

$$dX(t) = [a_0(t) + a_1(t)X(t) + a_2(t)Y(t)] dt + b(t)dW^S(t) \quad (2.1)$$

$$dY(t) = [h_0(t) + h_1(t)X(t) + h_2(t)Y(t)] dt + dW^N(t) \quad (2.2)$$

for  $0 \leq t \leq T$  where  $X(0)$  is a Gaussian random variable independent of  $(W^S, W^N)$  and  $Y(0) = 0$ . The coefficients  $a_0, a_1, a_2, h_1, h_2, h_3$  and  $b$  are all assumed to be non-random, bounded

and measurable. The observation noise  $W^N$  and the noise  $W^S$  driving the SDE for the signal  $X$  are assumed to be standard Brownian motions with

$$\langle W^S, W^N \rangle_t = \int_0^t \rho(u) du.$$

The following facts are well-known. Such a system of SDE's admits a unique solution  $(X, Y)$ . By virtue of the linear nature of the coefficients, it follows that the pair  $(X, Y)$  is jointly Gaussian. Then the optimal filter  $\pi$ , which is the conditional distribution of the signal given the observations, is defined by

$$\int f d\pi(t) = E[f(X(t)) | \mathcal{F}_t^Y] \quad \text{for all bounded, continuous functions } f. \quad (2.3)$$

Further,  $\pi(t)$  is also a Gaussian measure (see *e.g.* Elliott (1982), Kallianpur (1980)). Hence  $\pi(t)$  is completely determined by its first two moments  $\hat{X}(t)$  and  $P(t)$  which are defined by

$$\hat{X}(t) = E[X(t) | \mathcal{F}_t^Y]. \quad (2.4)$$

$$P(t) = E[(X(t) - \hat{X}(t))^2 | \mathcal{F}_t^Y]. \quad (2.5)$$

Moreover, the conditional variance  $P(t)$  is non-random. (See *e.g.* page 522 of Rao (1985).)

The additional information got by observing  $Y$  at the instant  $t$  is given by the *innovations process*  $I(t)$ , which is defined by

$$I(t) = Y(t) - \int_0^t [h_0(u) + h_1(u)\hat{X}(u) + h_2(u)Y(u)] du.$$

It is well-known that  $I$  is a standard Brownian motion with respect to the observation  $\sigma$ -field  $\mathcal{F}_t^Y$ .

We have the following theorem (see Theorem 10.5.1. of Kallianpur (1980)) regarding the optimal filter.

**Theorem 2.1.** *The Kalman filter  $(\hat{X}, P)$  for the signal-observation model (2.1)–(2.2) is the unique solution of the following system of equations.*

$$\begin{aligned} \hat{X}(t) = E[X_0] + \int_0^t [a_0(u) + a_1(u)\hat{X}(u) + a_2(u)Y(u)] du \\ + \int_0^t [b(u)\rho(u) + P(u)h_1(u)] dI(u) \end{aligned} \quad (2.6)$$

and

$$P(t) = V[X_0] + \int_0^t [2a_1(u)P(u) + (b(u))^2 - (b(u)\rho(u) + h_1(u)P(u))^2] du \quad (2.7)$$

**Remark 2.1.** The above result is also true for unbounded coefficients under some appropriate integrability conditions, see Theorem 10.5.1. of Kallianpur (1980).

To end this section we give the general FKK equation (Theorem 8.4.4 from Kallianpur (1980), also Theorem 18.11 from Elliott (1982)) which will be used in the next section. Suppose that the observation model is given by

$$Z(t) = \int_0^t H(u, X, Z)du + \int_0^t \alpha(u)dW(u)$$

where  $W$  is a standard Brownian motion, the observation function  $H$  is a non-anticipating functional of  $(X, Z)$  and  $\alpha$  is a deterministic function bounded away from zero. We assume the energy condition

$$\int_0^T |H(u, X, Z)|^2 du < \infty \text{ a.s.} \quad (2.8)$$

Suppose that the signal process  $X$  is such that

$$M^f(t) = f(X(t)) - \int_0^t \widetilde{A}f(u)du$$

is a martingale with

$$\langle M^f, W \rangle(t) = \int_0^t \widetilde{D}f(u)du.$$

where  $\widetilde{A}f$  and  $\widetilde{D}f$  are some (non-anticipative) functionals of  $X$  and  $Z$ . Let  $\pi(t, f)$  denote the conditional distribution of  $f(X(t))$  given the observation  $\sigma$ -field  $\mathcal{F}_t^Z = \sigma\{Z(u) : 0 \leq u \leq t\}$ . Let

$$I(t) = \int_0^t (\alpha(u))^{-1} dY(u) - \int_0^t (\alpha(u))^{-1} \pi(u, H)du$$

be the innovations process. Then we have the following theorem.

**Theorem 2.2.** *Under the above setup,  $\pi(t, f)$  satisfies*

$$\begin{aligned} \pi(t, f) &= \pi(0, f) + \int_0^t \pi(u, \widetilde{A}f)du \\ &+ \int_0^t \left[ \pi(u, \widetilde{D}f) + (\alpha(u))^{-1} (\pi(u, Hf) - \pi(u, H)\pi(u, f)) \right] dI(u). \end{aligned} \quad (2.9)$$

### 3 Linear Signal-Observation Model with OUP noise

Let  $W^N$  and  $W^S$  be standard Brownian motions with

$$\langle W^S, W^N \rangle(t) = \int_0^t \rho(u)du. \quad (3.1)$$

We will assume that  $\rho$  is a bounded measurable function on  $[0, T]$  with

$$\rho(u) > -1, \quad 0 \leq u \leq T. \quad (3.2)$$

Fix a  $\beta > 0$ . Consider the Ornstein-Uhlenbeck *velocity* process

$$\begin{aligned} dV^\beta(t) &= -\beta V^\beta(t)dt + dW^N(t), \quad 0 \leq t \leq T, \\ V^\beta(0) &= 0. \end{aligned} \quad (3.3)$$

Then  $V^\beta$  is given by

$$V^\beta(t) = \beta \int_0^t e^{-\beta(t-u)} dW^N(u), \quad 0 \leq t \leq T. \quad (3.4)$$

Let  $O^\beta$  denote the Ornstein-Uhlenbeck *displacement* process defined by

$$O^\beta(t) = \int_0^t V^\beta(u) du, \quad 0 \leq t \leq T. \quad (3.5)$$

Then  $O^\beta$  has the form

$$O^\beta(t) = \int_0^t (1 - e^{-\beta(t-u)}) dW^N(u), \quad 0 \leq t \leq T. \quad (3.6)$$

It is well-known that  $O^\beta$  converges to  $W^N$  in  $L^2$  as  $\beta \rightarrow \infty$ . (See Nelson (1967)).

We will consider the following linear signal-observation model. Let  $X_0$  be a Gaussian random variable independent of  $(W^S, W^N)$ . Throughout the article, the coefficients  $a_0, a_1, a_2, h_0, h_1, h_2$  and  $b$  will be assumed to be non-random, bounded and measurable functions on  $[0, T]$ . Further the coefficients  $h_0, h_1, h_2$  will be assumed to be continuously differentiable on  $[0, T]$ . Let  $C < \infty$  denote a common bound for all the coefficients and their derivatives. The model is then given by

$$dX^\beta(t) = \left[ a_0(t) + a_1(t)X^\beta(t) + a_2(t)Y^\beta(t) \right] dt + b(t)dW^S(t) \quad (3.7)$$

$$dY^\beta(t) = \left[ h_0(t) + h_1(t)X^\beta(t) + h_2(t)Y^\beta(t) \right] dt + dO^\beta(t) \quad (3.8)$$

for  $0 \leq t \leq T$  with  $X^\beta(0) = X_0$  and  $Y^\beta(0) = 0$ . This is a SDE with linear coefficients driven by the Gaussian semi-martingale  $(W^S, O^\beta)$ . Existence and uniqueness of solution to such an equation is quite well-known, (see *e.g.* Karandikar (1989), Métivier (1982)). We also have the following observation which is crucially used in the rest of the article.

**Theorem 3.1.** *Let the processes  $W^S, W^N, O^\beta$ , the random variable  $X_0$  and the coefficients  $a_0, a_1, a_2, h_0, h_1, h_2$  and  $b$  be as above. Then the solution of the SDE (3.7)–(3.8) is a Gaussian process.*

*Proof.* We note that  $(W^S, O^\beta, X_0)$  is jointly Gaussian. We will now construct a Gaussian solution for (3.7)–(3.8). Let

$$X^{\beta,0}(t) = X_0, \quad Y^{\beta,0}(t) = 0 \quad 0 \leq t \leq T,$$

and define successively for  $m \geq 1$ ,

$$X^{\beta,m}(t) = \int_0^t \left[ a_0(u) + a_1(u)X^{\beta,m-1}(u) + a_2(u)Y^{\beta,m-1}(u) \right] du + \int_0^t b(u)dW^S(u)$$

$$Y^{\beta,m}(t) = \int_0^t \left[ h_0(u) + h_1(u)X^{\beta,m-1}(u) + h_2(u)Y^{\beta,m-1}(u) \right] du + O^\beta(t).$$

Note that  $(X^{\beta,m}, Y^{\beta,m})$  is a Gaussian process for every  $m \geq 1$ . Also

$$\begin{aligned} & |(X^{\beta,m} - X^{\beta,m-1})(t)|^2 \\ &= \left| \int_0^t \left[ a_1(u)(X^{\beta,m-1} - X^{\beta,m-2})(u) + a_2(u)(Y^{\beta,m-1} - Y^{\beta,m-2})(u) \right] du \right|^2 \\ &\leq 2C^2T \int_0^t \left[ |(X^{\beta,m-1} - X^{\beta,m-2})(u)|^2 + |(Y^{\beta,m-1} - Y^{\beta,m-2})(u)|^2 \right] du. \end{aligned}$$

Similarly, we get

$$\begin{aligned} & |(Y^{\beta,m} - Y^{\beta,m-1})(t)|^2 \\ &\leq 2C^2T \int_0^t \left[ |(X^{\beta,m-1} - X^{\beta,m-2})(u)|^2 + |(Y^{\beta,m-1} - Y^{\beta,m-2})(u)|^2 \right] du. \end{aligned}$$

Define  $\eta^m(t) = \sup_{0 \leq u \leq t} \{ E [ |(X^{\beta,m} - X^{\beta,m-1})(t)|^2 + |(Y^{\beta,m} - Y^{\beta,m-1})(t)|^2 ] \}$ . Then using Fubini's theorem it follows that

$$\eta^m(t) \leq 4C^2T \int_0^t \eta^{m-1}(u) du \leq \dots \leq \frac{(4C^2T)^{m-1}}{(m-1)!} \eta^1(t), \quad (3.9)$$

where

$$\begin{aligned} \eta^1(t) &\leq \sup_{0 \leq u \leq t} \left\{ E \left[ |X^{\beta,1}(u) - X_0|^2 + |Y^{\beta,1}(u)|^2 \right] \right\} \\ &\leq 2 \sup_{0 \leq u \leq t} \left\{ E \left| \int_0^t (a_0(u) + a_1(u)X_0 - X_0) du \right|^2 + E \left| \int_0^t b(u) dW^S(u) \right|^2 \right. \\ &\quad \left. + E \left| \int_0^t (h_0(u) + h_1(u)X_0) du \right|^2 + E(O^\beta(t))^2 \right\} \\ &< \infty. \end{aligned} \quad (3.10)$$

It now follows from (3.9) and (3.10) that  $(X^{\beta,m}, Y^{\beta,m})$  converge to a process  $(X^\beta, Y^\beta)$ , which is necessarily Gaussian. It is also clear from the construction that the limiting process is a solution of (3.7)–(3.8). This completes the proof.  $\square$

The optimal filter  $\pi^\beta(t)$  for the model (3.7)–(3.8) is the conditional distribution of  $X^\beta(t)$  given the observation  $\sigma$ -field  $\mathcal{F}_t^{Y^\beta} := \sigma\{Y^\beta(s); 0 \leq s \leq t\}$ . Theorem 3.1 implies that  $\pi^\beta(t)$  is Gaussian and hence, as in the classical case, it is completely determined by the first two conditional moments  $\hat{X}^\beta(t)$  and  $P^\beta(t)$  where

$$\hat{X}^\beta(t) = E[X^\beta(t) | \mathcal{F}_t^{Y^\beta}] \quad (3.11)$$

and

$$P^\beta(t) = E[(X^\beta(t) - \hat{X}^\beta(t))^2 | \mathcal{F}_t^{Y^\beta}] = E[(X^\beta(t))^2 | \mathcal{F}_t^{Y^\beta}] - (\hat{X}^\beta(t))^2. \quad (3.12)$$

Further it is well-known that the conditional variance  $P^\beta(t)$  is in fact deterministic and is equal to  $E[(X^\beta(t) - \hat{X}^\beta(t))^2]$ . (See e.g. page 522 of Rao (1985)).



**Remark 3.1.** In Bhatt and Karandikar (2002) the Ornstein-Uhlenbeck noise  $V^\beta$  was assumed to be stationary which implies that  $V^\beta(0) \sim N(0, \beta/2)$ . However this is not a necessary assumption and the analysis there will work for a general  $V^\beta(0)$ . In this article, as seen from the above comments, we are going to use the Gaussian nature of the process  $(X^\beta, Y^\beta)$ . To ensure this we would need  $V^\beta(0)$  to be normally distributed. Thus the assumption that  $V^\beta(0) = 0$  is only for convenience and is not a restriction.

Our aim now is to characterize  $(\hat{X}^\beta(t), P^\beta(t))$  via filtering equations. Towards this end we proceed by making some *transformations* of the model (3.7)–(3.8). First let us define

$$y^\beta(t) = \frac{d}{dt}Y^\beta(t); \quad \xi^\beta(t) = h_0(t) + h_1(t)X^\beta(t) + h_2(t)Y^\beta(t). \quad (3.13)$$

Then in view of (3.5) the observation model (3.8) can be rewritten in an equivalent form as

$$y^\beta(t) = \xi^\beta(t) + V^\beta(t), \quad 0 \leq t \leq T. \quad (3.14)$$

**Remark 3.2.** Model (3.14) can be thought of as the instantaneous observation model as opposed to (3.8) which is an accumulative observation model.

Now let

$$\check{y}^\beta(t) = e^{\beta t}y^\beta(t); \quad \check{\xi}^\beta(t) = e^{\beta t}\xi^\beta(t); \quad \check{V}^\beta(t) = e^{\beta t}V^\beta(t) \quad (3.15)$$

Then it follows that

$$\check{y}^\beta(t) = \check{\xi}^\beta(t) + \check{V}^\beta(t), \quad 0 \leq t \leq T. \quad (3.16)$$

Using (3.4) we have

$$\check{V}^\beta(t) = \beta \int_0^t e^{\beta u} dW^N(u).$$

Thus we get that  $\check{V}^\beta$  is a semimartingale and

$$\int_0^t \frac{e^{-\beta u}}{\beta} d\check{V}^\beta(u) = W^N(t). \quad (3.17)$$

The relation (3.16) implies that  $\check{y}^\beta(t)$  is also a semimartingale. Now define

$$\tilde{Y}^\beta(t) = \int_0^t \frac{e^{-\beta u}}{\beta} d\check{y}^\beta(u); \quad \tilde{\xi}^\beta(t) = \int_0^t \frac{e^{-\beta u}}{\beta} d\check{\xi}^\beta(u). \quad (3.18)$$

It follows from (3.16)–(3.18) that

$$\tilde{Y}^\beta(t) = \tilde{\xi}^\beta(t) + W^N(t), \quad 0 \leq t \leq T. \quad (3.19)$$

Now, using (3.7), (3.8), (3.13), (3.15), (3.18) and applying integration by parts, we get

$$\begin{aligned}
\tilde{\xi}^\beta(t) &= \int_0^t \frac{e^{-\beta u}}{\beta} d\check{\xi}^\beta(u) \\
&= \int_0^t \frac{e^{-\beta u}}{\beta} d[e^{\beta u} \xi^\beta(u)] \\
&= \int_0^t \xi^\beta(u) du + \int_0^t \frac{1}{\beta} d\xi^\beta(u) \\
&= \int_0^t \xi^\beta(u) du + \frac{1}{\beta} \int_0^t [h'_0(u) + h'_1(u)X^\beta(u) + h'_2(u)Y^\beta(u)] du \\
&\quad + \frac{1}{\beta} \int_0^t h_1(u) dX^\beta(u) + \frac{1}{\beta} \int_0^t h_2(u) dY^\beta(u) \\
&= \int_0^t \left[ H_0^\beta(u) + H_1^\beta(u)X^\beta(u) + H_2^\beta(u)Y^\beta(u) \right] du \\
&\quad + \int_0^t B_1^\beta(u) dW^S(u) + \int_0^t B_2^\beta(u) dW^N(u)
\end{aligned} \tag{3.20}$$

where

$$H_0^\beta(u) = h_0(u) + \frac{1}{\beta} [h'_0(u) + a_0(u)h_1(u) + h_0(u)h_2(u)], \tag{3.21a}$$

$$H_1^\beta(u) = h_1(u) + \frac{1}{\beta} [h'_1(u) + a_1(u)h_1(u) + h_1(u)h_2(u)], \tag{3.21b}$$

$$H_2^\beta(u) = h_2(u) + \frac{1}{\beta} [h'_2(u) + a_2(u)h_1(u) + (h_2(u))^2], \tag{3.21c}$$

$$B_1^\beta(u) = \frac{1}{\beta} h_1(u)b(u) \tag{3.21d}$$

and

$$B_2^\beta(u) = \frac{1}{\beta} h_2(u). \tag{3.21e}$$

Let a deterministic function  $\alpha^\beta$  be defined by

$$(\alpha^\beta(t))^2 = (B_1^\beta(t))^2 + (B_2^\beta(t) + 1)^2 + 2\rho(t)B_1^\beta(t)(B_2^\beta(t) + 1), \quad 0 \leq t \leq T \tag{3.22}$$

where  $\rho$  is as in (3.1). It follows from (3.2) that  $(\alpha^\beta(t))^2 > 0$  for all  $t$ . Let

$$W(t) = \int_0^t \frac{B_1^\beta(u)}{\alpha^\beta(u)} dW^S(u) + \int_0^t \frac{B_2^\beta(u) + 1}{\alpha^\beta(u)} dW^N(u) \tag{3.23}$$

Then  $W$  is a Standard Brownian motion. Moreover, using (3.20)–(3.23) we can now write the observation model (3.19) as

$$\begin{aligned}
\tilde{Y}^\beta(t) &= \int_0^t \left[ H_0^\beta(u) + H_1^\beta(u)X^\beta(u) + H_2^\beta(u)Y^\beta(u) \right] du + \int_0^t \alpha^\beta(u) dW(u), \\
&\quad 0 \leq t \leq T.
\end{aligned} \tag{3.24}$$

Also, it follows from (3.13), (3.15), and (3.18) that

$$\begin{aligned}\sigma\left(Y^\beta(u):u\leq t\right) &= \sigma\left(y^\beta(u):u\leq t\right) = \sigma\left(\tilde{y}^\beta(u):u\leq t\right) \\ &= \sigma\left(\tilde{Y}^\beta(u):u\leq t\right).\end{aligned}\tag{3.25}$$

Let us denote the  $\sigma$ -field in (3.25) by  $\mathcal{F}_t^\beta$ . Thus the optimal filter  $(\hat{X}^\beta(t), P^\beta(t))$  is the conditional mean and variance of  $X^\beta(t)$  given  $\mathcal{F}_t^\beta$ .

**Remark 3.3.** The new signal-observation model is now given by (3.7) and (3.24). The advantage in making the transformations is that both these equations are now driven by (correlated) Brownian motions  $W^S$  and  $W$ . We also note here, using (3.1) and (3.23), that

$$\begin{aligned}\langle W^S, W \rangle(t) &= \int_0^t \left[ (\alpha^\beta(u))^{-1} B_1^\beta(u) + (\alpha^\beta(u))^{-1} \rho(u) (B_2^\beta(u) + 1) \right] du \\ &:= \int_0^t \lambda^\beta(u) du.\end{aligned}\tag{3.26}$$

**Remark 3.4.** It follows from (3.24) that the *transformed* observation process  $\tilde{Y}^\beta$  is still a Gaussian process. However the SDE's (3.7) and (3.24) are no longer linear in  $(X^\beta, \tilde{Y}^\beta)$ . In fact, it is easy to see that

$$Y^\beta(t) = \int_0^t e^{-\beta u} \left( \int_0^u \beta e^{\beta r} d\tilde{Y}^\beta(r) \right) du.$$

Thus the coefficients in (3.7) and (3.24) depend on the entire past of  $\tilde{Y}^\beta$  (in a non-anticipative way). Note that the dependence on  $X^\beta$  is still *Markovian* and linear.

Let us denote the non-anticipative functional by  $H^\beta$ . *i.e.*

$$H^\beta(t, X^\beta(t), \tilde{Y}^\beta) = H_0^\beta(t) + H_1^\beta(t) X^\beta(t) + H_2^\beta(t) Y^\beta(t) \quad 0 \leq t \leq T.\tag{3.27}$$

Let  $\hat{H}^\beta(t)$  denote its conditional expectation  $E[H^\beta(t, X^\beta(t), \tilde{Y}^\beta) | \mathcal{F}_t^\beta]$ . Define the *innovations process*  $I^\beta$  as follows.

$$I^\beta(t) = \int_0^t \left( \alpha^\beta(u) \right)^{-1} d\tilde{Y}^\beta(u) - \int_0^t \left( \alpha^\beta(u) \right)^{-1} \hat{H}^\beta(u) du \quad 0 \leq t \leq T.\tag{3.28}$$

Then it is well-known (and also easy to see) that  $I^\beta$  is a standard Brownian motion adapted to the filtration  $\mathcal{F}_t^\beta$ . Now we are ready to prove the main theorem of this section.

**Theorem 3.2.** *Let the processes  $W^S, W^N, O^\beta$ , the random variable  $X_0$  and the coefficients  $a_0, a_1, a_2, h_0, h_1, h_2$  and  $b$  be as above. Let the signal-observation model be given by (3.7)-(3.8). Let  $H_1^\beta$  be defined by (3.21b). Then the optimal filter  $(\hat{X}^\beta, P^\beta)$  (given by (3.11), (3.12)) is the*

unique solution of the equations

$$\begin{aligned}\widehat{X}^\beta(t) &= E[X_0] + \int_0^t \left( a_0(s) + a_1(s)\widehat{X}^\beta(s) + a_2(s)Y^\beta(s) \right) ds \\ &\quad + \int_0^t \left( b(s)\lambda^\beta(s) + \left( \alpha^\beta(s) \right)^{-1} H_1^\beta(s)P^\beta(s) \right) dI^\beta(s)\end{aligned}\quad (3.29)$$

$$\begin{aligned}P^\beta(t) &= V[X_0] + \int_0^t \left\{ 2a_1(s)P^\beta(s) + (b(s))^2 \right. \\ &\quad \left. - \left( b(s)\lambda^\beta(s) + \left( \alpha^\beta(s) \right)^{-1} H_1^\beta(s)P^\beta(s) \right)^2 \right\} ds.\end{aligned}\quad (3.30)$$

*Proof.* We have already noted that the observation model (3.24) is equivalent to the original model (3.8). Also, as remarked after equation (3.25), to find the optimal filter for the original model it suffices to work only with the observation model (3.24). However, Remark 3.4 tells us that we cannot directly apply the classical Kalman-Bucy filter formula. Instead, we will use Theorem 2.2, which is a more general result from non-linear filtering theory.

Our assumption of boundedness of all the coefficients, together with the fact that  $(X^\beta, Y^\beta)$  is Gaussian, implies that the observation function  $H^\beta$  (as in (3.27)) satisfies the energy condition (2.8). Thus Theorem 2.2 is applicable. In fact, we need to apply the Theorem for only two functions, *viz.*,  $f_1(x) = x$  and  $f_2(x) = x^2$ . This is easily done by identifying the various components appearing in (2.9).

It is immediate from (3.7) and (3.26) that

$$M^1(t) = X^\beta(t) - \int_0^t \left[ a_0(s) + a_1(s)X^\beta(s) + a_2(s)Y^\beta(s) \right] ds$$

is a martingale with  $\langle M^1, W \rangle(t) = \int_0^t b(s)\lambda^\beta(s)ds$ . Also, recalling (3.12) and using the fact that  $Y^\beta$  is  $(\mathcal{F}_t^\beta)$  adapted, the term  $\pi(u, Hf_1) - \pi(u, H)\pi(u, f_1)$  appearing in the stochastic integral in (2.9) reduces to  $H_1^\beta(u)P^\beta(u)$ . Thus we get that  $\widehat{X}^\beta(t)$  satisfies (3.29).

Similarly, to evaluate  $\pi^\beta(t, f_2)$ , we apply Ito's formula to (3.7) and use (3.26) to get that

$$\begin{aligned}M^2(t) &= (X^\beta(t))^2 - \int_0^t \left[ 2a_0(s)X^\beta(s) + 2a_1(s)(X^\beta(s))^2 \right. \\ &\quad \left. + 2a_2(s)X^\beta(s)Y^\beta(s) + (b(s))^2 \right] ds\end{aligned}$$

is a martingale with  $\langle M^2, W \rangle(t) = \int_0^t 2b(s)\lambda^\beta(s)X^\beta(s)ds$ . The cross quadratic variation gives the first term  $(\pi^\beta(\cdot, \widehat{D}f_2))$  in the integrand of the stochastic integral in (2.9). Let  $f_3(x) = x^3$ . Then the second term can be easily seen to equal

$$H_1^\beta(u)\pi^\beta(u, f_3) - H_1^\beta(u)\widehat{X}^\beta(u)\pi^\beta(u, f_2).$$

Hence we get that  $\pi^\beta(t, f_2)$  satisfies the SDE

$$\begin{aligned} \pi^\beta(t, f_2) &= E[X_0^2] + \int_0^t \left[ 2a_0(s)\widehat{X}^\beta(s) + 2a_1(s)\pi^\beta(s, f_2) \right. \\ &\quad \left. + 2a_2(s)\widehat{X}^\beta(s)Y^\beta(s) + (b(s))^2 \right] ds \\ &+ \int_0^t \left[ 2b(s)\lambda^\beta(s)\widehat{X}^\beta(s) + \left(\alpha^\beta(s)\right)^{-1} H_1^\beta(s) \left( \pi^\beta(s, f_3) - \widehat{X}^\beta(s)\pi^\beta(s, f_2) \right) \right] dI^\beta(s). \end{aligned} \quad (3.31)$$

Now, another application of Ito's formula to equation (3.29) gives

$$\begin{aligned} \left(\widehat{X}^\beta(t)\right)^2 &= [E(X_0)]^2 + \int_0^t \left[ 2a_0(s)\widehat{X}^\beta(s) + 2a_1(s) \left(\widehat{X}^\beta(s)\right)^2 + 2a_2(s)\widehat{X}^\beta(s)Y^\beta(s) \right. \\ &\quad \left. + \left( b(s)\lambda^\beta(s) + \left(\alpha^\beta(s)\right)^{-1} H_1^\beta(s)P^\beta(s) \right)^2 \right] ds \\ &+ \int_0^t \left( 2b(s)\lambda^\beta(s)\widehat{X}^\beta(s) + 2 \left(\alpha^\beta(s)\right)^{-1} H_1^\beta(s)P^\beta(s)\widehat{X}^\beta(s) \right) dI^\beta(s). \end{aligned} \quad (3.32)$$

Finally subtracting (3.32) from (3.31) we get

$$\begin{aligned} P^\beta(t) &= \pi^\beta(t, f_2) - \left(\widehat{X}^\beta(t)\right)^2 \\ &= V(X_0) + \int_0^t \left[ 2a_1(s)P^\beta(s) + (b(s))^2 \right. \\ &\quad \left. - \left( b(s)\lambda^\beta(s) + \left(\alpha^\beta(s)\right)^{-1} H_1^\beta(s)P^\beta(s) \right)^2 \right] ds \\ &+ \int_0^t \left(\alpha^\beta(s)\right)^{-1} H_1^\beta(s) \left( \pi^\beta(s, f_3) - \pi^\beta(s, f_2)\widehat{X}^\beta(s) - 2P^\beta(s)\widehat{X}^\beta(s) \right) dI^\beta(s). \end{aligned} \quad (3.33)$$

The integrand in the stochastic integral above is zero as expected, since

$$\begin{aligned} &\pi^\beta(s, f_3) - \pi^\beta(s, f_2)\widehat{X}^\beta(s) - 2P^\beta(s)\widehat{X}^\beta(s) \\ &= \pi^\beta(s, f_3) + 2 \left(\widehat{X}^\beta(s)\right)^3 - 3\widehat{X}^\beta(s)\pi^\beta(s, f_2) \\ &= E \left[ \left( X^\beta(s) - \widehat{X}^\beta(s) \right)^3 \mid \mathcal{F}_s^\beta \right] = 0. \end{aligned}$$

Thus (3.33) is same as (3.30).

Equations (3.29) and (3.30) have the same form as the filtering equations in the classical linear filtering case. (*i.e.* (2.6) and (2.7)). Uniqueness of solution is also similarly proved. (See Proof of Theorem 10.5.1 in Kallianpur (1980)).

□

## References

- [1] Bhatt, A.G., Kallianpur, G. and Karandikar, R.L. (1995) *Uniqueness and robustness of solution of measure valued equations of nonlinear filtering* Ann. Probab. **23** 1895-1938.

- [2] Bhatt, A.G. and Karandikar, R.L. (1999) *Characterization of the Optimal Filter: The Non- Markov Case* Stochastics and Stoch. Rep. **66** 177–204.
- [3] Bhatt, A.G., and Karandikar, R.L. (2002) *On Filtering with Ornstein-Uhlenbeck Process as Noise* Submitted for publication.
- [4] Bhatt, A.G., Rajput, B. and Xiong, J. (2002) *Zakai equation of nonlinear filtering with Ornstein-Uhlenbeck noise: Existence and Uniqueness* submitted for publication.
- [5] Elliott, R.J. (1982) *Stochastic Calculus and Applications*. Springer, New York.
- [6] Fujisaki, M., Kallianpur, G. and Kunita, H. (1972) *Stochastic differential equations for the nonlinear filtering problem*. Osaka J. Math. **9** 19-40.
- [7] Gawarecki, L. and Mandrekar, V. (2000) *On the Zakai equation of filtering with Gaussian noise* Stochastics in finite and infinite dimensions, Volume in honour of Gopinath Kallianpur, eds. T. Hida et al, **145–151**.
- [8] Kallianpur, G. (1980) *Stochastic Filtering theory*. Springer-Verlag, New York, Heidelberg, Berlin.
- [9] Kallianpur, G. and Karandikar, R.L. (1988) *White noise theory of Prediction, Filtering and Smoothing*. Gordon and Breach, New York.
- [10] Kalman, R.E. (1960) *New results in linear filtering and prediction theory*. J. Basic Eng. ASME **82** 33-45.
- [11] Kalman, R.E. and Bucy R.S. (1961) *A new approach to linear filtering and prediction problems*. J. Basic Eng. ASME **83** 95-108.
- [12] Karandikar, R.L. (1989) *On Métivier-Pellaumail inequality, Emery topology and Pathwise formulae in stochastic calculus* **Sankhya A 51** 121–143.
- [13] Kouritzin, M. and Xiong, J. (2002) *Nonlinear filtering: from Ornstein-Uhlenbeck to white noise* preprint.
- [14] Kunita, H. (1993) *Representation and stability of nonlinear filters associated with Gaussian noises* **Stochastic Processes: A Festschrift in Honour of Gopinath Kallianpur**. 201–210 *Ed. S. Cambanis, J.K. Ghosh, R.L. Karandikar, P.K. Sen*.
- [15] Mandrekar, V and Mandal, P.K. (2000) *A Bayes Formula For Gaussian Processes and its Applications* **SIAM J. Control Optim.** **39(3)** 852-871.
- [16] Métivier, M. (1982) *Semimartingales: A Course on Stochastic Processes*. Walter de Gruyter, Berlin.
- [17] Nelson, E. (1967) *Dynamical Theories of Brownian Motion*. Princeton University press.

[18] Rao, C.R. (1985) *Linear Statistical Inference and its Applications*. Wiley Eastern.