Communications on Stochastic Analysis

Volume 4 | Number 1 Article 4

3-1-2010

Some solvable classes of filtering problem with Ornstein-Uhlenbeck noise

Zhicheng Liu

Jie Xiong

Communications on Stochastic Analysis
Vol. 4, No. 1 (2010) 35-49

Serials Publications
www.serialspublications.com

SOME SOLVABLE CLASSES OF FILTERING PROBLEM WITH ORNSTEIN-UHLENBECK NOISE

ZHICHENG LIU AND JIE XIONG*

ABSTRACT. This is a companion paper of Crisan *et el* [4]. In this article, we study a few classes of solvable models of the stochastic filtering problems with Ornstein-Uhlenbeck noise: Firstly, we study the singular linear filter with OU noise. Secondly, for nonsingular linear filtering with OU noise, we consider the limit to the classical Kalman-Bucy filter as the OU process converges to the Brownian motion. Finally, we investigate the same filtering problem when the signal is governed by a nonlinear stochastic differential equation.

1. Introduction

The objective of the stochastic filtering is to estimate the stochastic signal X_t based on the observation which is a function $h(X_t)$ of the signal plus a stochastic noise n_t, i.e.,

$$y_t = h(X_t) + n_t. \tag{1.1}$$

The ideal model of n_t is the white noise, i.e., $\{n_t : \ t \in [0,T]\}$ is a family of independent random variables with identical distribution and mean 0. However, the paths of the white noise as functions of t do not exist in the ordinary sense, i.e., they are generalized functions. To overcome this difficulty, many authors considered the accumulated form of the observation model

$$dY_t = h(X_t)dt + dW_t, \tag{1.2}$$

where W_t is the integral of the white noise, which is actually a Brownian motion. We refer the reader to the works of Bucy and Kalman [3], Kushner [9], Fujisaki *et al* [5] and Zakai [12], and the book of Kallianpur [7].

Kunita [8], Mandal and Mandrekar [10] and Gawarecki and Mandrekar [6] studied the model (1.1) when n_t is a general Gaussian process. The most important example is the case when n_t is an Ornstein-Uhlenbeck process (OUP). However, the conditions imposed in these papers are very restrictive. Most notably, the authors assume that the map $t \mapsto h(X_t)$ is differentiable. To remove this restrictive condition, Bhatt and Karandikar [2] consider a variant of the observation model

$$y_t = \alpha \int_{(t-\alpha^{-1})\vee 0}^{t} h(X_s)ds + O_t$$

Received 2009-10-23; Communicated by the editors.
2000 *Mathematics Subject Classification.* Primary 60G35; Secondary 60G30.
Key words and phrases. Stochastic filtering, Ornstein-Uhlenbeck noise, singular filtering.
* This research is supported by the NSF DMS-0906907.

for $\alpha > 0$ and obtain the same results for this modified model.

The smoothing of the observation mentioned above is not needed in the paper of Crisan *et al* [4]. Actually, without this smoothness, a new feature of the optimal filter is found, namely, the optimal filter takes values in the space of singular probability measures. The basic idea is to convert the problem into a classical one with signal dependent noise, which itself is a long-standing open problem in the study of stochastic filtering.

Although some results are obtained in Crisan *et al* [4], the discussion provided in that paper is not very conclusive. In fact, it only provides a procedure on how to approach the problem. More specifically, the conditions imposed by [4] are sequential in nature, i.e., it indicates that, under certain conditions, the procedure can stop and the filtering equations are then derived; and, when those conditions are not satisfied, the filtering problem is then transformed to a new one. The same discussion then continues for the new model. However, it is not clear how many steps will be needed (or, whether the procedure can be stopped at all) and what kind of conclusion will be made based on the original model. We shall say that the filtering problem is *completely solvable* if the conclusions can be made based on the coefficients of the original signal-observation system.

This article serves as a companion paper of [4]. In this paper, we consider three solvable classes of filtering models with OUP as their observation noises. Firstly, we consider a linear filtering model with the d-dimensional signal X_t and the m-dimensional observation y_t governed by the following equations:

$$dX_t = (b_0 + b_1 X_t)dt + b_2 dB_t, \tag{1.3}$$

and

$$y_t = hX_t + O_t, \tag{1.4}$$

where O_t is an m-dimensional OUP governed by the following stochastic differential equation (SDE):

$$dO_t = -a_1 O_t dt + a_2 dB_t, \tag{1.5}$$

the coefficients $b_0,\ b_1, b_2,\ h,\ a_1,\ a_2$ are matrices of the dimensions $d\times 1,\ d\times d,\ d\times k,\ m\times d,\ m\times m,\ m\times k$, respectively, and B_t is a k-dimensional Brownian motion. For this and the second parts, we shall assume the following initial condition:

(IC1): *X_0 is normal with mean $\hat{X}_0 \in \mathbb{R}^d$ and covariance matrix $\gamma_0 \in \mathbb{R}^{d\times d}$.*

Denote by π_t the optimal filter, i.e., the conditional distribution of X_t in $\mathbb{R}^d$ given the σ-field $\mathcal{F}_t^y = \sigma(y_s :\ s \le t)$. We will characterize when the optimal filter is singular according to the rank of a matrix calculated from the coefficients. We will also derive the filtering equation satisfied by π_t.

The matrix $hb_2 + a_2$ will play a major role in this paper. We denote its rank by r, i.e., $r = \text{rank}(hb_2 + a_2)$. Let T_r be the product of elementary matrices of row operations reducing $hb_2 + a_2$ to its echelon form and let T_c be the product of elementary matrices of interchanging columns so that

$$T_r(hb_2 + a_2)T_c = \begin{pmatrix} I & E \\ 0 & 0 \end{pmatrix}, \tag{1.6}$$

where I is the $r \times r$ identity matrix and E is a $r \times (k - r)$ matrix. Note that T_c is a $k \times k$ orthogonal matrix.

Let b_{21} (resp. b_{22}) be the first r (resp. last $k-r$) columns of the matrix b_2T_c. Let

$$F_1 = b_{21}(I+EE^*)^{-\frac{1}{2}} + b_{22}(I+E^*E)^{-1}E^*(I+EE^*)^{\frac{1}{2}}, \tag{1.7}$$

and

$$F_2 = b_{21}(I+EE^*)^{-1}E(I+E^*E)^{\frac{1}{2}} - b_{22}(I+E^*E)^{-\frac{1}{2}}, \tag{1.8}$$

where E^* denotes the transpose of the matrix E.

Let F_3 be the last $m-r$ rows of the matrices $T_r(hb_1+a_1h)$. Let

$$\tilde{h} = (I+EE^*)^{-\frac{1}{2}}T_r(hb_1+a_1h). \tag{1.9}$$

The following condition (CS) will be needed for the optimal filter being singular of continuous type.

Condition (CS): There exists a matrix C such that $F_3(F_1\tilde{h}-b_1) = CF_3$. Further, $F_3F_2 = 0$ and rank$(F_3F_1) = m-r$.

Theorem 1.1. *i) If $F_2 = 0$, then the optimal filter π_t is singular of discrete type. In fact, X_t is $\mathcal{F}_t^y$-measurable and $\pi_t = \delta_{X_t}$, $t > 0$.*

ii) If $F_2 \neq 0$ and $r = m$, then the optimal filter π_t is not singular. The filtering problem is converted to a classical Kalman-Bucy one.

iii) If $F_2 \neq 0$, $r < m$ and the Condition (CS) is satisfied, then the optimal filter π_t is singular of continuous type. In fact, π_t is supported on a random hyperplane M_{Z_t} and is absolutely continuous with respect to the Lebesgue measure on M_{Z_t}.

Remark 1.2. If the conditions of Theorem 1.1 iii) are not satisfied, then, as we will see in Section 2, the filtering problem can be reduced to either case i) or case iii) in last theorem. Note that, we cannot specify in advance which case we will have based on the original coefficients. However, it can be determined in finite many steps. In this sense, we still say that our general linear filtering problem with OUP as noise is solvable.

Secondly, we will consider a special case of (1.5): the OUP is given by

$$dO_t^\beta = -\beta O_t^\beta dt + \beta dW_t, \tag{1.10}$$

where β is a constant and W_t is an m-dimensional Brownian motion. We assume that the signal X_t is given by (1.3) with $k = d$. Namely, B_t is a d-dimensional Brownian motion independent of W_t, and $b_0,\ b_1,\ b_2$ are matrices of sizes $d\times 1$, $d\times d$, $d\times d$, respectively. In this case, the optimal filter is not singular. The filtering equation for this model is studied by Bhatt [1]. The focus of this article is the convergence of the current optimal filter to the Kalman-Bucy filter when the parameter β tends to ∞.

Denote the optimal filter by π_t^β. Note that O_t^β converges to a white noise. More specifically, the integrated noise $\int_0^t O_s^\beta ds$ converges to the Brownian motion W_t as $\beta\to\infty$. Consider the filtering problem with signal X_t given by (1.3) and the observation

$$dY_t = hX_t dt + dW_t. \tag{1.11}$$

Denote the optimal filter by π_t.

Theorem 1.3. *As $\beta \to \infty$, the optimal filter π_t^β converges to π_t in the following sense:*

$$\lim_{\beta\to\infty} \mathbb{E}\left(\rho(\pi_t^\beta, \pi_t)^2\right) = 0,$$

where ρ is the Wasserstein distance (cf. p.122 of Xiong [11] *for its definition).*

Finally, we will consider the filtering problem with signal given by the following non-linear SDEs:

$$dX_t^i = \mu_i X_t^i dt + \sigma_i \sqrt{X_t^i} dB_t^i, \qquad i = 1, 2, \cdots, n, \tag{1.12}$$

where μ_i, σ_i are constant. Suppose that the observation model is given by

$$y_t = \sum_{i=1}^n X_t^i + O_t, \tag{1.13}$$

where O_t is the OUP of the form (1.10) with $\beta = 1$.

Suppose that the initial distribution π_0 of X_t satisfies the following condition

(IC2): π_0 *is absolutely continuous with respect to the Lebesgue measure, and the Radon-Nickodym derivative is a continuous function.*

Here is the main result about the nonlinear filtering with signal (1.12) and the observation (1.13).

Theorem 1.4. *Suppose that Condition (IC2) holds. i) If σ_i, $i = 1, 2, \cdots, n$ are all different, then π_t is singular of discrete type. More specifically, X_t is $\mathcal{F}_t^y$-measurable, and hence, $\pi_t = \delta_{X_t}$, $t > 0$.*

ii) Suppose that there exist $k \ge 1$ and $0 = \ell_0 < \ell_1 < \cdots < \ell_k = n$ such that for each $j = 1, 2, \cdots, k$, $\sigma_{\ell_{j-1}+1} = \cdots = \sigma_{\ell_j}$. For any $z \in \mathbb{R}_+^k$, let

$$M_z = \left\{x \in \mathbb{R}^n : \sum_{i=\ell_{j-1}+1}^{\ell_j} x_i = z_j, \ \ j = 1, \cdots, k\right\}.$$

Then π_t is supported on M_{Z_t}, where $Z_t = (Z_t^1, \cdots, Z_t^k)$ is the k-dimensional observable process defined by

$$Z_t^j = \sum_{i=\ell_{j-1}+1}^{\ell_j} X_i, \qquad j = 1, \cdots, k.$$

The proofs of Theorems 1.1, 1.3 and 1.4 will be presented in Sections 2, 3 and 4, respectively. Throughout this paper, we shall use K to denote a constant whose value can be changed from place to place.

2. Linear Filtering Model with O-U Process as a Noise

In this section, we consider the filtering model (1.3)-(1.4) with OUP noise (1.5). The main idea is to transform the filtering problem with OU noise to one with Brownian noise.

For a square matrix Q, we use e^Q to denotes its exponential, i.e.,

$$e^Q = \sum_{n=1}^\infty \frac{1}{n!} Q^n.$$

Define a new observation process Y_t by

$$dY_t = e^{-a_1 t} d(e^{a_1 t} y_t).$$

It is easy to show that $\mathcal{F}_t^Y = \mathcal{F}_t^y$. By Itô's formula,

$$\begin{aligned} dY_t &= dy_t + a_1 y_t dt \\ &= (hb_0 + (hb_1 + a_1 h)X_t)dt + (hb_2 + a_2)dB_t. \end{aligned} \tag{2.1}$$

It follows from (1.6) that the observation model can be rewritten as

$$T_r dY_t = (T_r(hb_1 + a_1 h)X_t + T_r hb_0)dt + \begin{pmatrix} I & E \\ 0 & 0 \end{pmatrix} dT_c^* B_t,$$

with $\mathcal{F}_T^Y = \mathcal{F}_t^{T_r Y}$, where $T_c^* B$ is a Brownian motion of dimension k, which will be denoted by $\begin{pmatrix} \tilde{B}^1 \\ \tilde{B}^2 \end{pmatrix}$ with $\tilde{B}^1$ and $\tilde{B}^2$ taking values in $\mathbb{R}^r$ and $\mathbb{R}^{k-r}$, respectively.

Define the stochastic processes

$$V_t^1 = (I + EE^*)^{-\frac{1}{2}}(\tilde{B}_t^1 + E\tilde{B}_t^2), \tag{2.2}$$

and

$$V_t^2 = (I + E^*E)^{-\frac{1}{2}}(E^*\tilde{B}_t^1 - \tilde{B}_t^2). \tag{2.3}$$

The following lemma follows from Lévy's characterization of Brownian motion easily. We omit its proof.

Lemma 2.1. *The processes V_t^1 and V_t^2 are two independent Brownian motions taking values in $\mathbb{R}^r$ and $\mathbb{R}^{k-r}$, respectively.*

Note that the observation process Y_t satisfies

$$T_r dY_t = (T_r(hb_1 + a_1 h)X_t + T_r hb_0)dt + \begin{pmatrix} (I + EE^*)^{\frac{1}{2}} dV_t^1 \\ 0 \end{pmatrix}. \tag{2.4}$$

Solve (2.2) and (2.3) for $\tilde{B}_t^i$, $i = 1, 2$, we have

$$\tilde{B}_t^1 = (I + EE^*)^{-1}\left((I + EE^*)^{\frac{1}{2}} V_t^1 + E(I + E^*E)^{\frac{1}{2}} V_t^2\right)$$

and

$$\tilde{B}_t^2 = (I + E^*E)^{-1}\left(E^*(I + EE^*)^{\frac{1}{2}} V_t^1 - (I + E^*E)^{\frac{1}{2}} V_t^2\right).$$

As

$$\begin{aligned} b_2 B_t &= b_2 T_c T_c^* B_t \\ &= b_{21}\tilde{B}_t^1 + b_{22}\tilde{B}_t^2, \end{aligned}$$

the signal process can be rewritten as

$$dX_t = (b_1 X_t + b_0)dt + F_1 dV_t^1 + F_2 dV_t^2, \tag{2.5}$$

where F_1 and F_2 are given by (1.7) and (1.8), respectively.

Now, we are ready to consider the filtering problem with the signal (2.5) and the observation (2.4). First, we consider the case of $F_2 = 0$.

Proof of Theorem 1.1 i). For simplicity of notations, we assume that $r = m$. Then the observation model (2.4) implies that

$$dV_t^1 = (I + EE^*)^{-\frac{1}{2}} T_r dY_t - (I + EE^*)^{-\frac{1}{2}}(T_r(hb_1 + a_1 h)X_t + T_r hb_0)dt.$$

Plugging back to the signal equation (2.5), we have

$$\begin{aligned} dX_t &= \Big(b_1 - F_1(I+EE^*)^{-\frac{1}{2}}T_r(hb_1+a_1h)\Big)X_t dt \\ &\quad + \Big(b_0 - F_1(I+EE^*)^{-\frac{1}{2}}T_r hb_0\Big)dt \\ &\quad + F_1(I+EE^*)^{-\frac{1}{2}}T_r dY_t. \end{aligned} \tag{2.6}$$

It is clear that the SDE (2.6) driven by Y_t has a unique strong solution X_t. Hence, X_t is $\mathcal{F}_t^Y$-measurable. Therefore, $\pi_t = \delta_{X_t}$. □

Next, we consider the case of $F_2 \neq 0$ and $r = m$. In this case, we define the observation process

$$\tilde{Y}_t = (I+EE^*)^{-\frac{1}{2}}T_r Y_t.$$

It is clear that $\mathcal{F}_t^{\tilde{Y}} = \mathcal{F}_t^Y$ and

$$d\tilde{Y}_t = \Big(\tilde{h}X_t + \tilde{h}_0\Big)dt + dV_t^1, \tag{2.7}$$

where $\tilde{h}$ is defined in (1.9) and

$$\tilde{h}_0 = (I+EE^*)^{-\frac{1}{2}}T_r hb_0.$$

The next proposition is the more precise re-statement of the second part of Theorem 1.1. The proof follows from the classical Kalman-Bucy theory (cf. Xiong [11]).

Proposition 2.2. *When $r = m$ and $F_2 \neq 0$, the optimal filter π_t is conditionally Guassian with conditional mean $\hat{X}_t$ and conditional covariance matrix γ_t satisfying*

$$\hat{X}_t = \hat{X}_0 + \int_0^t (b_0 + b_1\hat{X}_s)ds + \int_0^t (F_1 + \gamma_s\tilde{h}^*)d\nu_s, \tag{2.8}$$

and

$$\frac{d}{dt}\gamma_t = \gamma_t b_1^* + b_1\gamma_t + F_1F_1^* + F_2F_2^* - (F_1+\gamma_t\tilde{h}^*)(F_1+\gamma_t\tilde{h}^*)^*, \tag{2.9}$$

where

$$\nu_t = \tilde{Y}_t - \int_0^t (\tilde{h}_0 + \tilde{h}\hat{X}_s)ds$$

is an m-dimensional Brownian motion.

Finally, we consider the case of $r < m$ and $F_2 \neq 0$. In this case, the observation consists of two parts: $\tilde{Y}_t$ given by (2.7) and $\hat{Y}_t$ given by

$$d\hat{Y}_t = (F_3X_t + F_4)dt,$$

where F_3 is given in Section 1 and F_4 consists of the last $m-r$ rows of the matrix $T_r hb_0$. Let $Z_t = F_3X_t$. Then $\mathcal{F}_t^y = \mathcal{F}_t^{\tilde{Y},Z}$ for $t > 0$.

The following result should be available in the literature. However, we cannot find the exact form to suit our purpose. We state here for the convenience of the reader.

Lemma 2.3. *Let X be a $\mathbb{R}^n$-valued normal random variable with mean μ and covariance matrix Σ. Let F be a $m \times n$ matrix and $Z = FX$. Then, given $Z = z$, the conditional distribution of X is normal with mean μ_z and covariance matrix Σ_z given by*

$$\mu_z = G^{-1} \begin{pmatrix} 0 \\ (G_2 - CG_1)\mu \end{pmatrix} + G^{-1} \begin{pmatrix} \tilde{G}_1 \\ C\tilde{G}_1 \end{pmatrix} z,$$

and

$$\Sigma_z = G^{-1} \begin{pmatrix} 0 & 0 \\ 0 & G_2 \Sigma G_2^* \end{pmatrix} (G^{-1})^*,$$

where

$$C = G_2 \Sigma G_1^* (G_1 \Sigma G_1^*)^{-1},$$

G_1 consists of the linearly independent rows which form a basis for the row space of the matrix F, G_2's rows form a basis of the orthogonal complement of the row space of F, $\tilde{G}_1$ is $n \times n$ invertible matrix such that $G_1 = \tilde{G}_1 F$, and $G = \begin{pmatrix} G_1 \\ G_2 \end{pmatrix}$.

The following proposition is a more precise statement of Theorem 1.1 iii).

Proposition 2.4. *Suppose that $F_2 \neq 0$, $r < m$ and the Condition (CS) is satisfied. For any $z \in \mathbb{R}^{m-r}$, let*

$$M_z = \{x \in \mathbb{R}^d : \quad F_3 x = z\}.$$

Then the optimal filter π_t is a conditionally normal probability measure supported on the stochastic hyperplane M_{Z_t}.

Proof. Since Z_t is observable, it is clear that π_t is supported on M_{Z_t}. Note that the row space N of F_3 coincides with the normal space of M_z for any z. Let

$$p = F_3^* (F_3 F_3^*)^{-1} F_3.$$

Then p is the orthogonal projection matrix from $\mathbb{R}^d$ to the subspace N. Note that

$$\begin{aligned} dZ_t &= F_3(b_0 + b_1 X_t)dt + F_3 F_1 dV_t^1 + F_3 F_2 dV_t^2 \\ &= (F_3 b_0 + F_3 b_1 X_t)dt + F_3 F_1 dV_t^1. \end{aligned} \tag{2.10}$$

Let $\rho = F_3^* (F_3 F_3^*)^{-1}$. Then

$$\rho dZ_t = p\left((b_0 + b_1 X_t)dt + F_1 dV_t^1\right).$$

Hence

$$\begin{aligned} dX_t &= (I - p)\left((b_0 + b_1 X_t)dt + F_1 dV_t^1 + F_2 dV_t^2\right) \\ &\quad + p\left((b_0 + b_1 X_t)dt + F_1 dV_t^1 + F_2 dV_t^2\right) \\ &= (I - p)\left((b_0 + b_1 X_t)dt + F_1 dV_t^1 + F_2 dV_t^2\right) + \rho dZ_t. \end{aligned}$$

Let $\kappa_t = X_t - \rho Z_t$. Then κ_t takes values on M_0. Further, κ_t, as the new signal process, satisfies the following equation:

$$d\kappa_t = \left(\tilde{b}_0^t + \tilde{b}_1 \kappa_t\right) dt + \tilde{F}_1 dV_t^1 + \tilde{F}_2 dV_t^2, \tag{2.11}$$

where $\tilde{b}_0^t = (I-p)(b_0 + b_1\rho Z_t)$, $\tilde{b}_1 = (I-p)b_1$ and $\tilde{F}_i = (I-p)F_i$, $i = 1,2$. The observation model can be rewritten as

$$d\tilde{Y}_t = \left(\tilde{h}_0^t + \tilde{h}\kappa_t\right) dt + dV_t^1, \tag{2.12}$$

where $\tilde{h}_0^t = \tilde{h}_0 + \tilde{h}b_1\rho Z_t$. Similar to the Kalman-Bucy filter, we can derive equations for the conditional mean $\hat{\kappa}_t$ and conditional covariance matrix η_t for the filter of π_t. Note that the initials $\hat{\kappa}_0$ and η_0 are given by Lemma 2.3 with F, m and z replaced by F_3, $m-r$ and z_0, respectively.

Finally, the conditional mean and covariance of the optimal filter are given respectively by

$$\hat{X}_t = \hat{\kappa}_t + \rho Z_t \text{ and } \gamma_t = \eta_t.$$

and the proof is finished. □

Finally, we demonstrate Remark 1.2 in more detail. If $\text{rank}(F_3) < m-r$, we may remove the redundant rows in F_3 and, write the remaining matrix as $\tilde{F}_3$ and denote $\tilde{Z}_t = \tilde{F}_3 X_t$. The discussion in Proposition 2.4 remains valid with F_3 and Z_t replaced by $\tilde{F}_3$ and $\tilde{Z}_t$, respectively.

If $F_3F_2 \neq 0$, then there exist independent Brownian motions V_t^{21} and V_t^{22} of appropriate dimensions such that

$$F_3F_2V_t^2 = C_1V_t^{21} \text{ and } F_2V_t^2 = C_2V_t^{21} + C_3V_t^{22}.$$

The discussion in Proposition 2.4 remains valid with V_t^1 and V_t^2 replaced by (V_t^1, V_t^{21}) and V_t^{22}, respectively.

If $F_3(F_1\tilde{h} - b_1)$ cannot be written as CF_3, then the equality

$$d(F_3F_1\tilde{Y}_t - Z_t) = \left(F_3F_1\tilde{h}_0 - F_3b_0 + F_3(F_1\tilde{h} - b_1)X_t\right) dt$$

obtained from (2.10) and (2.12) implies that $Z_t^1 \equiv F_3(F_1\tilde{h} - b_1)X_t$ is also observable. The discussion in Proposition 2.4 remains valid with Z_t replaced by (Z_t, Z_t^1).

3. From OU Noise to Brownian Motion

In this section, we consider the filtering problem with signal X_t given by

$$dX_t = (b_0 + b_1X_t)dt + b_2dB_t,$$

and the observation

$$y_t^\beta = hX_t + O_t^\beta,$$

where O_t^β is an OUP given by

$$dO_t^\beta = -\beta O_t^\beta dt + \beta dW_t,$$

while $\begin{pmatrix} B_t \\ W_t \end{pmatrix}$ is a $(d+m)$-dimensional Brownian motion, and $b_0,\ b_1,\ b_2,\ h$ are matrices of sizes $d\times 1$, $d\times d$, $d\times d$, $m\times d$, respectively.

Comparing with the general model studied in last section, here $k = d + m$ and B_t there is replaced by $\begin{pmatrix} B_t \\ W_t \end{pmatrix}$; $b_2,\ a_1,\ a_2$ are replaced by $(b_2, 0),\ \beta I,\ (0, \beta I)$, respectively. In this case, $r = m$, $E = \beta^{-1}hb_2$, $b_{21} = 0$ and $b_{22} = b_2$. Thus,

$$F_2 = -b_2 \left(I + \beta^{-2}hb_2(hb_2)^*\right)^{-\frac{1}{2}} \neq 0.$$

By Theorem 1.6, the optimal filter π_t^β is not singular.

Denote

$$c_1^\beta = b_2((hb_2)^*hb_2 + \beta^2 I)^{-1}(hb_2)^*(\beta^2 I + hb_2(hb_2)^*)^{1/2},$$

$$c_2^\beta = -b_2((hb_2)^*hb_2 + \beta^2 I)^{-1/2}\beta,$$

$$\tilde{h}_0^\beta = (\beta^2 I + hb_2(hb_2)^*)^{-1/2}hb_0,$$

$$\tilde{h}^\beta = (\beta^2 I + hb_2(hb_2)^*)^{-1/2}(hb_1 + \beta h).$$

and

$$\tilde{Y}_t^\beta = (\beta^2 I + hb_2(hb_2)^*)^{-1/2}Y_t^\beta.$$

By Proposition 2.2, the optimal filter π_t^β is conditionally normal with conditional mean $\hat{X}_t^\beta$ and the conditional covariance γ_t^β satisfying the following (stochastic) differential equations:

$$\hat{X}_t^\beta = \hat{X}_0^\beta + \int_0^t (b_0 + b_1\hat{X}_s^\beta)ds + \int_0^t (c_1^\beta + \gamma_s^\beta \tilde{h}^{\beta *})d\nu_s^\beta, \tag{3.1}$$

and

$$\frac{d}{dt}\gamma_t^\beta = \gamma_t^\beta b_1^* + b_1\gamma_t^\beta + a_t^\beta - (c_1^\beta + \gamma_s^\beta \tilde{h}^{\beta *})(c_1^\beta + \gamma_s^\beta \tilde{h}^{\beta *})^*, \tag{3.2}$$

where

$$a_t^\beta = c_1^\beta c_1^{\beta *} + c_2^\beta c_2^{\beta *}$$

and the process ν_t^β defined by

$$d\nu_t^\beta = d\tilde{Y}_t^\beta - (\tilde{h}_0^\beta + \tilde{h}^\beta \hat{X}_t^\beta)dt,$$

is an m-dimensional Brownian motion.

Denote by $(\hat{X}_t, \gamma_t)$ the Kalman-Bucy filtering of the signal X_t given by (1.3) and observation Y_t given by (1.11). We have

$$\hat{X}_t = \hat{X}_0 + \int_0^t (b_0 + b_1\hat{X}_s)ds + \int_0^t \gamma_s h^* d\nu_s, \tag{3.3}$$

and

$$\frac{d}{dt}\gamma_t = \gamma_t b_1^* + b_1\gamma_t + b_2 b_2^* - \gamma_t h^*(\gamma_t h^*)^*, \tag{3.4}$$

where $\nu_t = Y_t - \int_0^t h\hat{X}_s ds$ is an m-dimensional Brownian motion.

Now we prove that $(\hat{X}_t^{\beta}, \gamma_t^\beta)$ converges to $(\hat{X}_t, \gamma_t)$, as $\beta \to \infty$.

Proposition 3.1. *As $\beta \to \infty$, we have*

$$\lim_{\beta\to\infty} \sup_{0\le t\le T} |\gamma_t^\beta - \gamma_t| = 0,$$

and

$$\lim_{\beta\to\infty} \mathbb{E} \sup_{0\le t\le T} |\hat{X}_t^\beta - \hat{X}_t|^2 = 0.$$

Proof. It is clear that there exists a constant K such that

$$|c_1^\beta| + |c_2^\beta + b_2| + |\tilde{h}_0^\beta| + |\tilde{h}^\beta - h| \le K\beta^{-1}.$$

For two matrices P and Q, we say that $P \le Q$ if $Q - P$ is a positive definite matrix. By comparison, it is easy to show that for any $\beta > 0$ and $t \in [0, T]$, we have $\gamma_t^\beta \le \gamma_t^0$ and $\gamma_t \le \gamma_t^0$, where γ_t^0 is the solution of

$$\frac{d}{dt}\gamma_t^0 = \gamma_t^0 b_1^* + b_1\gamma_t^0 + a^0,$$

here a^0 is a positive definite matrix such that $b_2 b_2^* \le a^0$ and for any $\beta > 0$, $a^\beta \le a^0$. Denote

$$K = \sum_{i,j=1}^d \sup_{0\le t\le T} |(\gamma_t^0)^{ij}|.$$

Then $K < \infty$ and

$$|(\gamma_t^\beta)^{ij}| \le K \text{ and } |\gamma_t^{ij}| \le K, \qquad \forall \beta > 0 \text{ and } t \in [0, T].$$

Take the difference between equations (3.2) and (3.4), we get

$$\begin{aligned}
|\gamma_t^\beta - \gamma_t| &\le 2\int_0^t |b_1||\gamma_s^\beta - \gamma_s| ds + |a^\beta - b_2 b_2^*| \\
&\quad + \int_0^t \left(|c_1^\beta c_1^{\beta *}| + 2|c_1^\beta \tilde{h}^\beta \gamma_s^\beta| + |\gamma_s(\tilde{h}^{\beta *}\tilde{h}^\beta - h^* h)\gamma_s^\beta| \right) ds \\
&\quad + \int_0^t \left(|\tilde{h}^{\beta *}\tilde{h}^\beta \gamma_s^\beta| + |\gamma_s h^* h| \right) |\gamma_s^\beta - \gamma_s| ds \\
&\le K_1(\epsilon) + K_2 \int_0^t |\gamma_s^\beta - \gamma_s| ds,
\end{aligned}$$

where $K_1(\beta) \to 0$, as $\beta \to \infty$, and K_2 is a constant. It follows from Gronwall's inequality that

$$\sup_{t\le T} |\gamma_t^\beta - \gamma_t| \le K_1(\beta) e^{K_2 T} \to 0.$$

Now, we prove the second conclusion. Let

$$z_t^\epsilon = \hat{X}_t^\beta - \hat{X}_t.$$

Then

$$\begin{aligned}
z_t^\beta &= z_0^\beta + \int_0^t b_1 z_s^\beta ds + \int_0^t \left(c_1^\beta + \gamma_s^\beta \tilde{h}^{\beta *} - \gamma_s h^* \right) d\nu_s^\beta \\
&\quad + \int_0^t \gamma_s h^* d(\tilde{Y}_s^\beta - Y_s) - \int_0^t \gamma_s h^* \left(\tilde{h}_0^\beta + (\tilde{h}^\beta - h)\hat{X}_s^\beta + h z_s^\beta \right) ds.
\end{aligned}$$

Note that

$$\tilde{Y}_t^\beta - Y_t = \left(\beta(\beta^2 I + hb_2(hb_2)^*)^{-1/2} - I\right)\int_0^t y_s ds + (\beta^2 I + hb_2(hb_2)^*)^{-1/2} y_t.$$

It is then easy to show that

$$\mathbb{E}\sup_{t\le T}|\tilde{Y}_t^\beta - Y_t|^2 \le K_1\beta^{-2}.$$

Therefore,

$$\begin{aligned} f(t) &\equiv \mathbb{E}\sup_{s\le t}|z_s^\beta|^2 \\ &\le K_2\beta^{-2} + K_3\int_0^t \mathbb{E}|z_s^\beta|^2 ds \\ &\le K_2\beta^{-2} + K_3\int_0^t f(s)ds. \end{aligned}$$

It follows from Gronwall's inequality that

$$\mathbb{E}\sup_{s\le T}|z_s^\beta|^2 \le K_2\beta^{-2}e^{-K_3 T}.$$

□

Now we are ready to finish the

Proof of Theorem 1.3. As in the proof of Corollary 9.38 of Xiong [11], we have

$$\rho(\pi_t^\beta, \pi_t) \le |\hat{X}_t^\beta - \hat{X}_t| + \sqrt{d}\left|\sqrt{\gamma_t^\beta} - \sqrt{\gamma_t}\right|,$$

where $\sqrt{Q}$ stands for the square root of the positive definite matrix Q. The conclusion of Theorem 1.3 then follows from Proposition 3.1. □

4. Nonlinear Filtering with OUP Noise

In this section, we proceed to proving Theorem 1.4. Let

$$Y_t = \int_0^t e^{-s} d\left(e^s y_s\right).$$

Then $\mathcal{F}_t^y = \mathcal{F}_t^Y$. Applying Itô's formula to (1.13), we have

$$dY_t = \sum_{i=1}^n (\mu_i + 1)X_t^i dt + \sum_{i=1}^n \sigma_i\sqrt{X_t^i}dB_t^i + dW_t.$$

It is easy to show that the quadratic variation process of Y_t is

$$\langle Y\rangle_t = \int_0^t \left(\sum_{i=1}^n \sigma_i^2 X_s^i + 1\right) ds,$$

and hence, the process

$$Z_t \equiv \sum_{i=1}^n \sigma_i^2 X_t^i$$

is $\mathcal{F}_t^y$-measurable.

Proof of Theorem 1.4 i). Applying Itô's formula to Z_t, we have

$$dZ_t = \sum_{i=1}^n \mu_i \sigma_i^2 X_t^i dt + \sum_{i=1}^n \sigma_i^3 \sqrt{X_t^i} dB_t^i.$$

As the quadratic covariation process between Y_t and Z_t is given by

$$\langle Y, Z\rangle_t = \int_0^t \sum_{i=1}^n \sigma_i^4 X_s^i ds,$$

the process

$$Z_t^1 \equiv \sum_{i=1}^n \sigma_i^4 X_t^i$$

is $\mathcal{F}_t^y$-measurable.

Replacing Z by Z^1 in the above argument, and continuing in this fashion, we then get that

$$Z_t^k \equiv \sum_{i=1}^n \sigma_i^{2k} X_t^i, \qquad k = 1, 2, \cdots, n$$

are $\mathcal{F}_t^y$-measurable. It is elementary to show that $X_t = (X_t^1, X_t^2, \cdots, X_t^n)$ is a linear transformation of the random vector $(Z_t, Z_t^1, \cdots, Z_t^{n-1})$ when the σ_i's are all different, and hence, X_t is $\mathcal{F}_t^y$-measurable. □

Finally, we proceed to prove the second part of Theorem 1.4.

Proof of Theorem 1.4 ii). For simplicity of notations, we assume $k = 1$. Namely, all σ_i's are the same. Without loss of generality, we assume that $\sigma_i = 1, \quad i = 1, 2, \cdots, n$.

Let

$$\kappa_t \equiv X_t - \frac{1}{\sqrt{n}} \bar{e}(Z_t - z_0),$$

where $\bar{e} = \frac{1}{\sqrt{n}}(1, 1, \cdots, 1)^*$. Then κ_t is an M_{z_0}-valued process. Denote its optimal filter by U_t, i.e., U_t is a $\mathcal{P}(M_{z_0})$-valued process such that

$$\langle U_t, f\rangle = \mathbb{E}\left(f(\kappa_t)|\mathcal{F}_t^Y\right), \qquad \forall f \in C_b(M_{z_0}).$$

The optimal filter π_t of X_t can be represented as

$$\pi_t(\cdot) = U_t\left(\cdot - n^{-1/2}\bar{e}(Z_t - z_0)\right), \qquad t > 0.$$

To finish this section, we derive the SDE satisfied by U_t. We introduce the notations: For $x \in \mathbb{R}^n$, we denote

$$\mu(x) = (\mu_1 x_1, \mu_2 x_2, \cdots, \mu_n x_n)^*,$$

and

$$\bar{\mu}(x) = \sum_{i=1}^n \mu_i x_i.$$

By Itô's formula, we get

$$dZ_t = \bar{\mu}(X_t) dt + \sqrt{Z_t} dW_t^0,$$

where W_t^0 is a real-valued Brownian motion given by

$$dW_t^0 = \sum_{j=1}^n \sqrt{\frac{X_t^j}{Z_t}} dB_t^j.$$

Let

$$\mu(x) = \langle \mu(x), \bar{e} \rangle \bar{e} + \tilde{\mu}(x) \text{ and } e_j = \langle e_j, \bar{e} \rangle \bar{e} + \tilde{e}_j$$

be the orthogonal decompositions on $M_z^\perp \oplus M_z$, where $\{e_j,\ j = 1, 2, \cdots, n\}$ is the standard basis of $\mathbb{R}^n$, and $M_z^\perp$ denotes the orthogonal complement of the subspace M_z.

Note that

$$\begin{aligned} d\kappa_t &= \mu(X_t)dt + \sum_{i=1}^n \sqrt{X_t^i} e_i dB_t^i - \frac{\bar{e}}{\sqrt{n}} \left(m\bar{u}(X_t)dt + \sqrt{Z_t} dW_t^0 \right) \\ &= \tilde{\mu}(X_t)dt + \sum_{i=1}^n \sqrt{X_t^i} \left(e_i - \frac{1}{\sqrt{n}} \bar{e} \right) dB_t^i. \end{aligned} \tag{4.1}$$

It is clear that the $n \times n$-matrix on the right hand side of the equation below has rank less than or equal to $n-1$, there exists $n \times (n-1)$-matrix $M(x)$ such that

$$M(x)M(x)^* = \sum_{i=1}^n x_i \left(\tilde{e}_i - \frac{1}{\sqrt{n}} \bar{e} \right) \left(\tilde{e}_i - \frac{1}{\sqrt{n}} \bar{e} \right)^*.$$

It follows from (4.1) that

$$d\kappa_t = \tilde{\mu}(X_t)dt + M(X_t) d\tilde{W}_t, \tag{4.2}$$

where $\tilde{W}_t$ is a $(n-1)$-dimensional Brownian motion independent of the Brownian motion $\tilde{W}_t^1 \equiv (W_t^0, W_t)^*$ in $\mathbb{R}^2$.

For $\kappa \in M_{z_0}$ and $z \in \mathbb{R}_+$, we define

$$b(\kappa, z) = \tilde{\mu} \left(\kappa + n^{-1/2}(z - z_0)\bar{e} \right),$$

$$c(\kappa, z) = M \left(\kappa + n^{-1/2}(z - z_0)\bar{e} \right),$$

$$\hat{h}_1(\kappa, z) = \sum_{i=1}^n (\mu_i + 1) \left(\kappa^i + n^{-1}(z - z_0) \right),$$

and

$$\hat{h}_2(\kappa, z) = \bar{\mu} \left(\kappa + n^{-1/2}(z - z_0)\bar{e} \right).$$

Define the new observation process $\tilde{Y}_t$ by

$$d\tilde{Y}_t = Q(Z_t)^{-1} d \begin{pmatrix} Y_t \\ Z_t \end{pmatrix},$$

where $Q(z)$ is the 2×2-matrix

$$Q(z) = \begin{pmatrix} \sqrt{z} & 1 \\ \sqrt{z} & 0 \end{pmatrix}.$$

Denote

$$h(\kappa, z) = Q(z)^{-1} \begin{pmatrix} \hat{h}_1(\kappa, z) \\ \hat{h}_2(\kappa, z) \end{pmatrix}.$$

Then the filtering model of κ_t can be written as

$$d\kappa_t = b(\kappa_t, Z_t)dt + c(\kappa_t, Z_t)d\tilde{W}_t \tag{4.3}$$

with observation

$$d\tilde{Y}_t = h(\kappa_t, Z_t)dt + d\tilde{W}_t^1. \tag{4.4}$$

It is clear that the filtering problem is of classical form with state space M_{z_0}, and hence, its optimal filter U_t satisfies the following SDE: For any $f \in C_b^2(M_{z_0})$

$$\langle U_t, f\rangle = \langle U_0, f\rangle + \int_0^t \langle U_s, L_{Z_s} f\rangle\, ds + \int_0^t \left(\left\langle U_s, f h^*_{Z_s}\right\rangle - \langle U_s, f\rangle \left\langle U_s, h^*_{Z_s}\right\rangle \right) d\nu_s,$$

where $h_z(\kappa) = h(\kappa, z)$, L_z is the generator of the process κ_t given by

$$L_z f(\kappa) = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(\kappa, z) \frac{\partial^2 f(\kappa)}{\partial \kappa_i \partial \kappa_j} + \sum_{i=1}^n b_i(\kappa, z) \frac{\partial f(\kappa)}{\partial \kappa_i},$$

with $a(\kappa, z) = c(\kappa, z)c(\kappa, z)^*$ is a $n \times n$-matrix, and

$$\nu_t = \tilde{Y}_t - \int_0^t \langle U_s, h_{Z_s}\rangle\, ds$$

is a two-dimensional Brownian motion. Note that, by Theorem 4.3 in Crisan *et al* [4], the initial measure U_0 is absolutely continuous with respect to the Lebesgue measure on M_{z_0} and

$$\frac{U_0(dx)}{dx} = \frac{\pi_0(x)}{\int_{M_{z_0}} \pi_0(y)dy},$$

where π_0, by abusing the notation a bit, is the Radon-Nickodym derivative of the initial distribution of the signal. □

References

1. Bhatt, A. G.: Linear filtering with Ornstein-Uhlenbeck process as noise. *Sādhanā,* **31**, Part 2, (2006) 141-153.
2. Bhatt, A. G. and Karandikar, R. L.: On filtering with Ornstein-Uhlenbeck process as noise. *J. Indian Statist. Asso.* **41**, no. 2, (2001) 205-220.
3. Bucy, R. S. and Kalman, R. E.: New results in linear filtering and prediction theory. *J. Basic Engineering, Tans. ASME* **83**, (1961) 95-108.
4. Crisan, D., Kouritzin, M., and Xiong, J.: Nonlinear filtering with signal depending noise. *Electronic Journal of Probability* **14**, (2009) 1863-1883.
5. Fujisaki, M., Kallianpur, G., and Kunita, H.: Stochastic differential equations for the non-linear filtering problem. *Osaka J. Math.* **9**, (1972) 19-40.
6. Gawarecki, L. and Mandrekar, V.: On the Zakai equation of filtering with Gaussian noise. *Stochastics in finite and infinite dimensions, Volume in honor of Gopinath Kallianpur, eds. T. Hida et al*, (2000) 145–151.
7. Kallianpur, G.: *Stochastic Filtering Theory.* Springer-Verlag, 1980.
8. Kunita, H.: Representation and stability of nonlinear filters associated with Gaussian noises. *Stochastic Processes: A festschrift in honor of Gopinath Kallianpur, eds. Cambanis et al, Springer-Verlag, New York.* (1993) 201–210.
9. Kushner, H. J.: Dynamic equations for nonlinear filtering. *J. Differ. Equations,* **3**, (1967) 179-190.

10. Mandrekar, V. and Mandal, P. K.: A Bayes formula for Gaussian processes and its applications. *SIAM J. Control Optim.* ***39***, (2000) 852–871.
11. Xiong, J.: *An Introduction to Stochastic Filtering Theory.* Oxford University Press, 2008.
12. Zakai, M.: On the optimal filtering of diffusion processes. *Z. Wahr. Verw. Gebiete* **11**, (1969) 230-243.

Zhicheng Liu: Department of Mathematics, Hebei Normal University, Shijiazhuang 050016, PRC

E-mail address: liuzhicheng0302@163.com

Jie Xiong: Department of Mathematics, University of Tennessee, Knoxville, TN 37996-1300, USA

E-mail address: jxiong@math.utk.edu

URL: http://www.math.utk.edu/∼jxiong