Publ. RIMS. Kyoto Univ.
29 (1993), 857-868

Filtering Problems for Conditionally Linear Systems with Non-Gaussian Initial Conditions

By

Shigeomi HARA*

Abstract

We solve the filtering problems for conditionally linear systems, that is, stochastic, partially observable systems which are linear in unobservable processes and nonlinear in observable processes, with non-Gaussian initial conditions. We assume that all coefficients of given SDEs are Lipshitz continuous in the observable processes, and that certain quantities concerning the diffusion coefficients of given SDEs are uniformly nonsingular. But we assume nothing about the initial distribution and allow that a part of the coefficients of the given SDEs are of linear growth order in the observable processes.

Introduction

The filtering problems for stochastic, partially observable *linear* systems with *Gaussian* initial conditions are solved by Kalman and Bucy [4]. Here "solve" means obtaining the formula to compute the optimal mean square estimate of the present value of the unobservable processes, using the finite dimensional statistics, when we have the data of the past and present observation. In Liptser and Shiryayev [7], they solve the filtering problems for stochastic, partially observable systems that are linear in unobservable processes, and *nonlinear* in observable processes, that is, *conditionally linear systems,* with Gaussian initial conditions. On the other hand, in Makowski [8], they solve the filtering problems for stochastic, partially observable linear systems with *non-Gaussian* initial conditions.

In this paper, we join above two methods in [7] and [8], and solve the filtering problems for conditionally linear systems with non-Gaussian initial conditions. Namely, we consider the following system of stochastic differential equations:

$$dX_t=\{a_0(t, Y)+a_1(t, Y)X_t\}\,dt+b_1(t, Y)dW_1(t)+b_2(t, Y)dW_2(t),$$

$$dY_t=\{A_0(t, Y)+A_1(t, Y)X_t\}\,dt+B_1(t, Y)dW_1(t)+B_2(t, Y)dW_2(t),$$

Communicated by H. Araki, December 25, 1992.
1991 Mathematics Subject Classification: 93E11.
* Department of Applied Mathematics and Physics, Faculty of Engineering, Kyoto University, Kyoto 606, Japan.

where X_t is an unobservable process, Y_t is an observable process, and W_1, W_2 are mutually independent standard Brownian motions. X_0 is an arbitrary $\mathcal{F}_0$-measurable random variable, and $Y_0=0$. We assume that all coefficients of above SDEs are Lipshitz continuous in the observable processes, and that certain quantities concerning the diffusion coefficients of above SDEs are uniformly nonsingular. In [5], Kolodziej and Mohler study the same problem. They assume the only certain integrabilities for the coefficients of given SDEs. But they also assume a certain integrability of the initial distribution, while we assume nothing about the initial distribution. In [2], Haussman and Pardoux study more general problem that contains also "Beneš problem", which is nonlinear in unobservable processes. There they assume only boundedness of the coefficients of given SDEs, while we allow that a part of coefficients are of linear growth order in the observable processes.

In section 1, we give the precise formulation of the problem and the assumptions, and consider the reduction of the problem to the case $b_2\equiv 0$. In section 2, we prove the main theorem under the assumptions in section 1. In section 3, we prove the proposition used in section 1.

§1. Problem Formulation and Assumptions

Let $(\Omega, \mathcal{F}, P)$ be a probability space. In order to guarantee the existence of the regular conditional probability used below, we assume that $(\Omega, \mathcal{F})$ is a standard measurable space (see e.g. [3]). Let $(\mathcal{F}_t)_{t\in[0,T]}$ $(T>0)$ be a right continuous increasing family of sub σ-fields of $\mathcal{F}$.

The unobservable process X_t and the observable process Y_t are M and N dimensional $\mathcal{F}_t$-adapted continuous processes, respectively, which satisfy the following stochastic differential equations:

$$X_t=X_0+\int_0^t\{a_0(s, Y)+a_1(s, Y)X_s\}\,ds+\int_0^t b_1(s, Y)dW_1(s)+\int_0^t b_2(s, Y)dW_2(s), \tag{1.1}$$

$$Y_t=\int_0^t\{A_0(s, Y)+A_1(s, Y)X_s\}\,ds+\int_0^t B_1(s, Y)dW_1(s)+\int_0^t B_2(s, Y)dW_2(s), \qquad t\in[0, T], \tag{1.2}$$

where $W_1(t)$ and $W_2(t)$ are mutually independent M and N dimensional standard $\mathcal{F}_t$-Brownian motions, respectively. a_0, a_1, b_1, b_2, A_0, A_1, B_1, and B_2 are progressively measurable mappings from $[0, T]\times C([0, T]; \boldsymbol{R}^N)$ to $\boldsymbol{R}^M$, $\boldsymbol{R}^{M\times M}$, $\boldsymbol{R}^{M\times M}$, $\boldsymbol{R}^{M\times N}$, $\boldsymbol{R}^N$, $\boldsymbol{R}^{N\times M}$, $\boldsymbol{R}^{N\times M}$, and $\boldsymbol{R}^{N\times N}$, respectively. X_0 is an $\boldsymbol{R}^M$ valued $\mathcal{F}_0$-measurable random variable which is generally non-Gaussian distributed.

We denote by $\mathcal{Y}_t$ the σ-algebra generated by $\{Y_s: s\leqq t\}$. It is well known that if $E[|X_t|^2]<+\infty$, then the optimal mean square estimation of X_t from the observation data $\mathcal{Y}_t$ is

$$E[X_t|\mathcal{Y}_t].$$

We aim at computing this using the solutions of finite dimensional stochastic differential equations.

We assume either following Assumptions A or Assumptions B for the coefficients of the SDEs (1.1)-(1.2). We define the $M\times(M+N)$ matrix b, and the $N\times(M+N)$ matrix B by

$$b=(b_1 \quad b_2), \qquad B=(B_1 \quad B_2).$$

Assumptions A. (1) Let $g(t, y)$ be any element of matrices $a_0(t, y)$, $a_1(t, y)$, $A_0(t, y)$, $A_1(t, y)$, $b_1(t, y)$, $b_2(t, y)$, $B_1(t, y)$, and $B_2(t, y)$. Then, for any $x, y\in C([0, T]; \boldsymbol{R}^N)$ and $t\in[0, T]$, we have

$$|g(t, y)|^2\leqq C\left\{\int_0^t(1+|y_s|^2)dK(s)+(1+|y_t|^2)\right\}, \tag{1.3}$$

$$|g(t, x)-g(t, y)|^2\leqq C\left\{\int_0^t|x_s-y_s|^2dK(s)+|x_t-y_t|^2\right\}, \tag{1.4}$$

where C is a constant and $K(s)$ is a nondecreasing right continuous function, $0\leqq K(s)\leqq 1$.

(2) a_1, A_1 are bounded.

(3) $\exists\delta>0$ $BB^*(t, y)\geqq\delta I$ for any $t\in[0, T]$, $y\in C([0, T]; \boldsymbol{R}^N)$, where $*$ denotes the transpose of a matrix.

(4) B_1, B_2 are bounded.

(5) $\exists\delta>0$ s.t.

(i) $bb^*(t, y)\geqq\delta I$,

(ii) $B(I-b^*(bb^*)^{-1}b)B^*(t, y)\geqq\delta I$, for any $t\in[0, T]$, $y\in C([0, T]; \boldsymbol{R}^N)$. □

Assumptions B. (1)-(3) are the same as (1)-(3) of Assumptions A.

(4) $b_2(t, y)=0$ for any $t\in[0, T]$, $y\in C([0, T]; \boldsymbol{R}^N)$.

(5) There exists an $\boldsymbol{R}^{N\times M}$ valued bounded progressively measurable function $\tilde{A}_1(t, y)$ which satisfies

$$A_1(t, y)=B_2(t, y)\tilde{A}_1(t, y), \qquad t\in[0, T],\ y\in C([0, T]; \boldsymbol{R}^N).$$ □

Remark. When $M=N=1$, then (ii) of Assumptions A (5) becomes

$$(b_1^2+b_2^2)^{-1}(B_1b_2-b_1B_2)^2\geqq\delta.$$

This means that two vectors $(b_1\ b_2)$ and $(B_1\ B_2)$ are not parallel. □

At first, we shall consider the reduction of problem to the case $b_2(t, y)\equiv 0$. Under Assumption B, we have nothing to do for this reduction.

Lemma 1.1 (Lemma 10.4, [6]). *Let $W(t)$ $(0\leqq t\leqq T)$ be an N dimensional standard $\mathcal{F}_t$-Brownian motion, and B_t be an $\boldsymbol{R}^{n\times N}$ valued $\mathcal{F}_t$-adapted process such that*

$$\int_0^T \mathrm{tr}\,(B_t B_t^*)dt<+\infty,$$

where $*$ denotes transpose and tr *denotes trace of square matrices. Let D_t be an $\boldsymbol{R}^{n\times k}$ valued $\mathcal{F}_t$-adapted process such that*

$$D_t D_t^*=B_t B_t^* \quad \text{for a.e. } t\in[0, T] \quad \text{a.s.}.$$

Then, there exists a k dimensional standard $\mathcal{F}_t$-Brownian motion $U(t)$ such that for any $t\in[0, T]$,

$$\int_0^t B_s dW(s)=\int_0^t D_s dU(s) \qquad \text{a.s.}. \qquad \square$$

In view of this lemma, we consider the equation:

$$\begin{pmatrix} b_1 & b_2 \\ B_1 & B_2 \end{pmatrix}\begin{pmatrix} b_1^* & B_1^* \\ b_2^* & B_2^* \end{pmatrix}=\begin{pmatrix} \tilde{b} & 0 \\ \tilde{B}_1 & \tilde{B}_2 \end{pmatrix}\begin{pmatrix} \tilde{b}^* & \tilde{B}_1^* \\ 0 & \tilde{B}_2^* \end{pmatrix},$$

that is,

$$bb^*=b_1b_1^*+b_2b_2^*=\tilde{b}\tilde{b}^*, \tag{1.5}$$

$$Bb^*=B_1b_1^*+B_2b_2^*=\tilde{B}_1\tilde{b}^*, \tag{1.6}$$

$$BB^*=B_1B_1^*+B_2B_2^*=\tilde{B}_1\tilde{B}_1^*+\tilde{B}_2\tilde{B}_2^*=\tilde{B}\tilde{B}^*. \tag{1.7}$$

Proposition 1.2. *Let either Assumptions A or Assumptions B be fulfilled. Then there exist $\boldsymbol{R}^{M\times M}$, $\boldsymbol{R}^{N\times M}$, and $\boldsymbol{R}^{N\times N}$ valued progressively measurable functions $\tilde{b}(t, y)$, $\tilde{B}_1(t, y)$, and $\tilde{B}_2(t, y)$, respectively, which satisfy* (1.5)-(1.7). *Moreover we have the following.*

(1) *Let $g(t, y)$ be any element of matrices $\tilde{b}(t, y)$, $\tilde{B}_1(t, y)$, and $\tilde{B}_2(t, y)$. Then $g(t, y)$ satisfies* (1.3) *and* (1.4).

(2) $\exists\delta>0$ $\tilde{B}\tilde{B}^*(t, y)\geqq\delta I$ *for any $t\in[0, T]$, $y\in C([0, T]; \boldsymbol{R}^N)$.*

(3) *There exists an $\boldsymbol{R}^{N\times M}$ valued bounded, progressively measurable function $\tilde{A}_1(t, y)$ which satisfies*

$$A_1(t, y)=\tilde{B}_2(t, y)\tilde{A}_1(t, y), \qquad t\in[0, T],\ y\in C([0, T]; \boldsymbol{R}^N). \qquad \square$$

We shall prove this proposition in section 3.

By Lemma 1.1 and Proposition 1.2, there exist M and N dimensional, mutually independent, standard $\mathcal{F}_t$-Brownian motions $U(t)$ and $\widetilde{W}_2(t)$, respectively, and the stochastic differential equations (1.1) and (1.2) are transformed to:

$$X_t=X_0+\int_0^t\{a_0(s, Y)+a_1(s, Y)X_s\}ds+\int_0^t\tilde{b}(s, Y)dU(s), \tag{1.8}$$

(1.9) $$Y_t=\int_0^t\{A_0(s, Y)+A_1(s, Y)X_s\}ds+\int_0^t\tilde{B}_1(s, Y)dU(s)$$
$$+\int_0^t\tilde{B}_2(s, Y)d\widetilde{W}_2(s), \qquad t\in[0, T].$$

§ 2. Main Theorem

Following the idea of Makowski [8], we shall introduce the process x_t which satisfies an ordinary differential equation:

(2.1) $$\frac{dx_t}{dt}=a_1(t, Y)x_t, \qquad x_0=X_0,$$

that is, letting $\Phi(t, Y)$ be the fundamental matrix with $\Phi(0, Y)=I$,

$$x_t=\Phi(t, Y)X_0.$$

We denote

$$\bar{X}_t=X_t-x_t,$$

then we can rewrite (1.8)-(1.9) to:

(2.2) $$\bar{X}_t=\int_0^t\{a_0(s, Y)+a_1(s, Y)\bar{X}_s\}ds+\int_0^t\tilde{b}(s, Y)dU(s)$$

(2.3) $$Y_t=\int_0^t\{A_0(s, Y)+A_1(s, Y)\bar{X}_s\}ds+\int_0^t\tilde{B}_1(s, Y)dU(s)$$
$$+\int_0^t A_1(s, Y)x_s ds+\int_0^t\tilde{B}_2(s, Y)d\widetilde{W}_2(s), \qquad t\in[0, T].$$

Noting (3) of Proposition 1.2, we set

(2.4) $$V_t=\int_0^t\tilde{A}_1(s, Y)x_s ds+\widetilde{W}_2(t)$$
$$=\int_0^t F(s, Y)X_0 ds+\widetilde{W}_2(t),$$

where,

(2.5) $$F(t, y)=\tilde{A}_1(t, y)\Phi(t, y).$$

In order to apply Girsanov's theorem to (2.4), if we set

(2.6) $$\Lambda_t=\exp\left[-\int_0^t\{F(s, Y)X_0\}^*d\widetilde{W}_2(s)-\frac{1}{2}\int_0^t|F(s, Y)X_0|^2ds\right], \quad t\in[0, T],$$

then we get the

Lemma 2.1. $$E[\Lambda_T]=1.$$ □

Proof. Here we follow essentially the argument of Example 2.2 in [2]. Let $p(x, d\omega)$ be the regular conditional probability given $X_0=x$. Since X_0 is $\mathcal{F}_0$-measurable, $\widetilde{W}_2(t)$ is still an $\mathcal{F}_t$-Brownian motion under the probability measure $p(x, \cdot)$ for P^{X_0}-a.e. $x\in\boldsymbol{R}^M$. From (2) of Assumptions A or B and (3) of Proposition 1.2, F is bounded, so for any $x\in\boldsymbol{R}^M$,

$$\int_\Omega \exp\left\{\frac{1}{2}\int_0^T |F(t, Y(\omega))x|^2 dt\right\} p(x, d\omega)<+\infty .$$

Hence, by Novikov's theorem (see e.g. Theorem III-5.3 in [3]), we have

$$\int_\Omega \Lambda_T(\omega)p(x, d\omega)=1 , \qquad \text{for } P^{X_0}\text{-a.e. } x\in\boldsymbol{R}^M .$$

Integrating this equation over $\boldsymbol{R}^M$ by $P^{X_0}(dx)$, we obtain Lemma 2.1. □

By Lemma 2.1, using Girsanov's theorem, we see that V_t is an $\mathcal{F}_t$-Brownian motion under the probability measure $\tilde{P}$, where

$$\tilde{P}(A)=\int_A \Lambda_T(\omega)P(d\omega), \qquad A\in\mathcal{F} . \tag{2.7}$$

Lemma 2.2. *On the probability space* $(\Omega, \mathcal{F}, \tilde{P})$,
(i) X_0, U, *and* V *are mutually independent;*
(ii) $\tilde{P}^{X_0}=P^{X_0}$ *on* $(\boldsymbol{R}^M, \mathcal{B}^M)$.

Proof. (i) By Girsanov's theorem, $(U(t), V(t))$ is an $M+N$ dimensional $\mathcal{F}_t$-Brownian motion under the probability $\tilde{P}$. Thus (i) is clear.
(ii) Since Λ_t is a martingale with respect to $(\mathcal{F}_t)$ under the probability P, for any $A\in\mathcal{B}^M$ we have

$$\tilde{E}[1_A(X_0)]=E[1_A(X_0)\Lambda_T]=E[1_A(X_0)\Lambda_0]=E[1_A(X_0)] . \qquad \square$$

Now, we have the stochastic differential equations:

$$\bar{X}_t=\int_0^t \{a_0(s, Y)+a_1(s, Y)\bar{X}_s\} ds+\int_0^t \tilde{b}(s, Y)dU(s) \tag{2.8}$$

$$Y_t=\int_0^t \{A_0(s, Y)+A_1(s, Y)\bar{X}_s\} ds+\int_0^t \tilde{B}_1(s, Y)dU(s) + \int_0^t \tilde{B}_2(s, Y)dV(s), \qquad t\in[0, T] , \tag{2.9}$$

on the probability space $(\Omega, \mathcal{F}, \tilde{P})$. By (1), (2) of Assumptions A or B and (1) of Proposition 1.2, the SDEs (2.8)-(2.9) have a unique strong solution, so we have

$$\sigma(\bar{X}_s, Y_s ; 0\leqq s\leqq t)\subset\sigma(U(s), V(s) ; 0\leqq s\leqq t) \tag{2.10}$$

For any $\phi : \boldsymbol{R}^M\to\boldsymbol{R}$, bounded Borel measurable mapping, we have the Kallianpur-Striebel formula:

$$E[\phi(X_t)|\mathcal{Y}_t]=\frac{\tilde{E}[\phi(X_t)Z_t|\mathcal{Y}_t]}{\tilde{E}[Z_t|\mathcal{Y}_t]}, \qquad t\in[0, T], \tag{2.11}$$

where $\tilde{E}$ denotes the expectation under the probability $\tilde{P}$, and $Z_t=\Lambda_t^{-1}$, i.e.

$$Z_t=\exp\Big[X_0^*\int_0^t F(s, Y)^*dV(s)-\frac{1}{2}X_0^*\int_0^t(F^*F)(s, Y)X_0ds\Big].$$

We set

$$h_t=\int_0^t F(s, Y)^*dV(s), \tag{2.12}$$

and

$$m_t=\tilde{E}\Big[\begin{pmatrix}\bar{X}_t\\ h_t\end{pmatrix}\Big|\mathcal{Y}_t\Big],\quad m_t'=\tilde{E}[\bar{X}_t|\mathcal{Y}_t],\quad \gamma_t=\tilde{E}\Big[\Big(\begin{pmatrix}\bar{X}_t\\ h_t\end{pmatrix}-m_t\Big)\Big(\begin{pmatrix}\bar{X}_t\\ h_t\end{pmatrix}-m_t\Big)^*|\mathcal{Y}_t\Big].$$

Proposition 2.3. *Let either Assumptions A or Assumptions B be fulfilled, and* $(\Omega, \mathcal{F}, \tilde{P})$ *be a probability space. Then, for any* $t\in[0, T]$, $(\bar{X}_t, h_t)$ *is Gaussian with mean vector* $m_t(\omega)$ *and covariance matrix* $\gamma_t(\omega)$ *under the regular conditional probability given* $\mathcal{Y}_t$, $\tilde{P}_{\mathcal{Y}_t}(\omega, A)$ $(A\in\mathcal{F})$, *for* $\tilde{P}$*-a.e.* ω. *Moreover,*

$$dm_t=\begin{pmatrix}a_0+a_1m_t'\\ 0\end{pmatrix}dt+\Big\{\begin{pmatrix}\tilde{b}\tilde{B}_1^*\\ F\tilde{B}_2^*\end{pmatrix}+\gamma_t\begin{pmatrix}A_1^*\\ 0\end{pmatrix}\Big\}(\tilde{B}\tilde{B}^*)^{-1} \tag{2.13}$$

$$\times\{dY_t-(A_0+A_1m_t')dt\}, \qquad m_0=0,$$

$$\dot{\gamma}_t=\begin{pmatrix}a_1 & 0\\ 0 & 0\end{pmatrix}\gamma_t+\gamma_t\begin{pmatrix}a_1^* & 0\\ 0 & 0\end{pmatrix}+\begin{pmatrix}\tilde{b}\tilde{b}^* & 0\\ 0 & FF^*\end{pmatrix}-\Big\{\begin{pmatrix}\tilde{b}\tilde{B}_1^*\\ F\tilde{B}_2^*\end{pmatrix} \tag{2.14}$$

$$+\gamma_t\begin{pmatrix}A_1^*\\ 0\end{pmatrix}\Big\}(\tilde{B}\tilde{B}^*)^{-1}\Big\{\begin{pmatrix}\tilde{b}\tilde{B}_1^*\\ F\tilde{B}_2^*\end{pmatrix}+\gamma_t\begin{pmatrix}A_1^*\\ 0\end{pmatrix}\Big\}^*, \quad \gamma_0=0. \qquad \square$$

Proof. By assumptions, we can apply Theorem 12.7 of [7] to equations (2.8), (2.9) and (2.12). □

We set

$$g_t=\frac{1}{2}X_0^*\int_0^t(FF^*)(s, Y)X_0ds.$$

Clearly we have

$$\tilde{E}[\phi(X_t)Z_t|\mathcal{Y}_t]=\tilde{E}[\tilde{E}[\phi(\bar{X}_t+x_t)\exp\{X_0^*h_t-g_t\}|\mathcal{Y}_t\vee\sigma(X_0)]|\mathcal{Y}_t]. \tag{2.15}$$

We denote by $\tilde{P}_0$ the regular conditional probability given $\mathcal{Y}_t\vee\sigma(X_0)$ on the probability space $(\Omega, \mathcal{F}, \tilde{P})$. Since x_t, X_0, and g_t are $\mathcal{Y}_t\vee\sigma(X_0)$-measurable, by the property of the conditional expectation we have

(2.16) $$\check{E}[\phi(\bar{X}_t+x_t)\exp\{X_0^*h_t-g_t\}\,|\,\mathcal{Y}_t\vee\sigma(X_0)](\omega)$$
$$=\int_\Omega\phi(\bar{X}_t(\tilde{\omega})+x_t(\omega))\exp\{X_0^*(\omega)h_t(\tilde{\omega})-g_t(\omega)\}\tilde{P}_0(\omega,\,d\tilde{\omega}),\quad\text{for }\tilde{P}\text{-a.e. }\omega.$$

By (2.10) and (2.12), we have

$$\sigma(\bar{X}_s,\,Y_s,\,h_s\,;\,0\leqq s\leqq t)\subset\sigma(U(s),\,V(s)\,;\,0\leqq s\leqq t).$$

Hence, from Lemma 2.2 (i),

(2.17) $(\bar{X},\,Y,\,h)$ and X_0 are mutually independent under $\tilde{P}$.

Therefore we can see that, under the conditional probability $\tilde{P}_0(\omega,\,\cdot)$, $(\bar{X}_t,\,h_t)$ is Gaussian with mean vector $m_t(\omega)$ and covariance matrix $\gamma_t(\omega)$ for $\tilde{P}$-a.e. ω. We denote by $\mu(dx\,;\,m,\,\gamma)$ the Gaussian distribution over $\boldsymbol{R}^{2M}$ with mean vector m and covariance matrix γ. Thus, from (2.16) we have

(2.18) $$\tilde{E}[\phi(\bar{X}_t+x_t)\exp\{X_0^*h_t-g_t\}\,|\,\mathcal{Y}_t\vee\sigma(X_0)](\omega)$$
$$=\int_{\boldsymbol{R}^{2M}}\phi(x'+x_t(\omega))\exp\{X_0^*(\omega)x''-g_t(\omega)\}\mu(dx\,;\,m_t(\omega),\,\gamma_t(\omega)),$$

for $\tilde{P}$-a. e. ω, where $x=\begin{pmatrix}x'\\x''\end{pmatrix}$, $x',\,x''\in\boldsymbol{R}^M$. Hence, by (2.15) and (2.18) we have

(2.19) $$\tilde{E}[\phi(X_t)Z_t\,|\,\mathcal{Y}_t]$$
$$=\tilde{E}\,[\int_{\boldsymbol{R}^{2M}}\phi(x'+x_t)\exp\{X_0^*x''-g_t\}\mu(dx\,;\,m_t,\,\gamma_t)\,|\,\mathcal{Y}_t],\quad\tilde{P}\text{-a. s.}$$

Since X_0 and $\mathcal{Y}_t$ are mutually independent and m_t, γ_t are $\mathcal{Y}_t$-measurable, we have

(2.20) $$\tilde{E}[\phi(X_t)Z_t\,|\,\mathcal{Y}_t](\omega)$$
$$=\tilde{E}\left[\int_{\boldsymbol{R}^{2M}}\phi(x'+\Phi(t,\,y)X_0)\exp\left\{X_0^*x''-\frac{1}{2}X_0^*\int_0^t(FF^*)(s,\,y)X_0ds\right\}\right.$$
$$\left.\mu(dx\,;\,m,\,\gamma)\right]_{y=Y(\omega),\,m=m_t(\omega),\,\gamma=\gamma_t(\omega)},$$
$$=\int_{\boldsymbol{R}^M}\int_{\boldsymbol{R}^{2M}}\phi(x'+\Phi(t,\,y)z)\exp\left\{z^*x''-\frac{1}{2}z^*\int_0^t(FF^*)(s,\,y)zds\right\}$$
$$\mu(dx\,;\,m,\,\gamma)P^{X_0}(dz)\,|_{\,y=Y(\omega),\,m=m_t(\omega),\,\gamma=\gamma_t(\omega)},$$

for $\tilde{P}$-a. e. ω. Here we have used Lemma 2.2 (ii).

Theorem 2.4. *Let either Assumptions A or Assumptions B be fulfilled. Then for any $\phi:\boldsymbol{R}^M\to\boldsymbol{R}$, bounded Borel measurable, the optimal mean square estimation $E[\phi(X_t)\,|\,\mathcal{Y}_t]$ satisfies Kallianpur-Striebel formula* (2.11), *and we have*

$$\tilde{E}[\phi(X_t)Z_t|\mathcal{Y}_t](\omega)$$
$$=\int_{R^M}\int_{R^{2M}}\phi(x'+\Phi(t, y)z)\exp\left\{z^*x''-\frac{1}{2}z^*\int_0^t(FF^*)(s, y)zds\right\}$$
$$\mu(dx\,;m, \gamma)P^{X_0}(dz)|_{y=Y(\omega),\, m=m_t(\omega),\, \gamma=\gamma_t(\omega)},\quad \text{for } P\text{-a.e. } \omega,$$

where $x=\binom{x'}{x''}$, $x', x''\in \boldsymbol{R}^M$. *Further,* $m_t(\omega)$ *and* $\gamma_t(\omega)$ *satisfy the stochastic differential equations* (2.13) *and* (2.14). □

§ 3. Proof of Proposition 1.2

Under Assumptions B, if we set

$$\tilde{b}=b_1,\qquad \tilde{B}_1=B_1,\qquad \tilde{B}_2=B_2,$$

then the conclusion of Proposition 1.2 holds clearly.

Next, we prove the conclusion of Proposition 1.2 under Assumptions A. If we set

$$\tilde{b}(t, y)=(bb^*)^{1/2}(t, y)\,; \tag{3.1}$$

$$\tilde{B}_1(t, y)=Bb^*(bb^*)^{-1/2}(t, y)\,; \tag{3.2}$$

$$\tilde{B}_2(t, y)=\{B(I-b^*(bb^*)^{-1}b)B^*\}^{1/2}(t, y),\quad t\in[0, T],\ y\in C([0, T]\,;\boldsymbol{R}^N), \tag{3.3}$$

then (1.5)-(1.7) are satisfied. By (i) of Assumptions A (5), $(bb^*)^{-1}$ is well defined. (2) of Proposition 1.2 clearly holds by (3) of Assumptions A and (1.7). For (3) of Proposition 1.2, we set

$$\tilde{A}_1(t, y)=\tilde{B}_2^{-1}A_1(t, y)\,. \tag{3.4}$$

By (ii) of Assumptions A (5), $\tilde{B}_2^{-1}$ is well defined, and

$$|\tilde{B}_2^{-1}(t, y)|\leqq C\delta^{-1/2}\,.$$

Hence, noting A_1 is also bounded, we can see that (3) of Proposition 1.2 holds. At last we shall prove (1) of Proposition 1.2. It is clear that

$$|\tilde{b}|\leqq C(|b_1|+|b_2|)\,, \tag{3.5}$$

so $\tilde{b}$ satisfies (1.3).

Lemma 3.1. *Let* $b: \boldsymbol{R}\to\boldsymbol{R}^{n\times m}$ $(n\leqq m)$ *satisfy the following*:

$$|b(t)-b(s)|\leqq K|t-s|,\qquad t, s\in\boldsymbol{R}\,;$$

$$bb^*>0\,.$$

Then if we set $a=bb^*$, *we have*

$$|a^{1/2}(t)-a^{1/2}(s)|\leqq CK|t-s|,\qquad t, s\in\boldsymbol{R}\,,$$

where the constant C depends only on n and m. □

Proof. We follow essentially the proof of Theorem 5.2.2 in [9], i.e., Lemma 3.2 in this paper.

At first we consider the case $b(t)\in C^1(\boldsymbol{R})$. Then clearly $a(t)\in C^1(\boldsymbol{R})$. By Lemma 6.1.1 in [1], we have also $a^{1/2}(t)\in C^1(\boldsymbol{R})$. It is sufficient to prove that

$$\sup_{t\in\boldsymbol{R}}|(a^{1/2})'(t)|\leqq CK\,. \tag{3.6}$$

Here $'$ denotes the differential in t. Let any $t\in\boldsymbol{R}$ be fixed. We may assume that $a(t)$ is a diagonal matrix. Indeed, there exists an orthogonal matrix Q such that $\tilde{a}(t)\equiv Qa(t)Q^*$ is diagonal. Clearly $\tilde{a}^{1/2}(t)=Qa^{1/2}(t)Q^*$. Therefore

$$|(\tilde{a}^{1/2})'(t)|=|(a^{1/2})'(t)|\,.$$

So let $a(t)$ be diagonal. Since $a^{1/2}(t)$ is also diagonal, we have

$$a'(t)_{ij}=(a^{1/2})'(t)_{ij}a^{1/2}(t)_{jj}+a^{1/2}(t)_{ii}(a^{1/2})'(t)_{ij}\,.$$

Here the suffix ij denotes the (i, j) component of matrices. Hence

$$(a^{1/2})'(t)_{ij}=\frac{a'(t)_{ij}}{a^{1/2}(t)_{ii}+a^{1/2}(t)_{jj}}\,.$$

On the other hand,

$$a'(t)_{ij}=\sum_{k=1}^{m}\{b'(t)_{ik}b(t)_{jk}+b(t)_{ik}b'(t)_{jk}\}\,.$$

So,

$$|a'(t)_{ij}|\leqq K\sum_{k=1}^{m}\{|b(t)_{jk}|+|b(t)_{ik}|\}\,.$$

Since $a(t)$ is diagonal, we have

$$a^{1/2}(t)_{ii}=\sqrt{a(t)_{ii}}=\sqrt{\sum_{k=1}^{m}b(t)_{ik}^2}\,.$$

Hence,

$$|(a^{1/2})'(t)_{ij}|\leqq\frac{K\sum_{k=1}^{m}\{|b(t)_{jk}|+|b(t)_{ik}|\}}{\sqrt{\sum_{k=1}^{m}b(t)_{ik}^2}+\sqrt{\sum_{k=1}^{m}b(t)_{jk}^2}}\leqq mK\,,$$

and we obtain (3.6).

When $b(t)\notin C^1(\boldsymbol{R})$, it is sufficient that we approximate $b(t)$ by $b_\varepsilon(t)$ with the molifier. □

By this lemma, we can directly prove that $\tilde{b}$ satisfies (1.4). So we see that $\tilde{b}$ satisfies (1) of Proposition 1.2.

In view of the definition of $\tilde{B}_1$, say (3.2), we set

$$A=(bb^*)^{-1/2}b\,.$$

Since

$$AA^*=(bb^*)^{-1/2}bb^*(bb^*)^{-1/2}=I\,,$$

we have

$$|A|^2\leqq C_1|AA^*|\leqq C_1\,.$$

Hence

$$|\tilde{B}_1|=|BA^*|\leqq C_2|B|\cdot|A^*|\leqq C_3|B|\leqq C_4(|B_1|+|B_2|)\,.$$

Therefore $\tilde{B}_1$ is bounded, and satisfies (1.3). From (1.6), we have

$$\tilde{B}_1(x)-\tilde{B}_1(y)=-\hat{B}_1(y)(\tilde{b}(x)-\tilde{b}(y))\tilde{b}(x)^{-1} \tag{3.7}$$

$$+\sum_{i=1}^{2}\{(B_i(x)-B_i(y))b_i^*(x)\tilde{b}^{-1}(x)+B_i(y)(b_i^*(x)-b_i^*(y))\tilde{b}^{-1}(x)\}\,,$$

for $x,\ y\in C([0,\ T]\,;\boldsymbol{R}^N)$. Here we abbreviated the argument t. B_1, B_2 and $\tilde{B}_1$ are bounded. Clearly $b_i^*\tilde{b}^{-1}=b_i^*(bb^*)^{-1/2}$ $(i=1,\ 2)$ is bounded. By (i) of Assumptions A (5), $\tilde{b}^{-1}$ is bounded. Hence from (3.7) we can see that $\tilde{B}_1$ satisfies (1.4). So $\tilde{B}_1$ satisfies (1) of Proposition 1.2.

From (1.7), using the boundedness of B_1, B_2 and $\tilde{B}_1$, we can see that $\tilde{B}_2$ is also bounded. So $\tilde{B}_2$ satisfies (1.3). By the definition of A, we have

$$A(x)-A(y)=\tilde{b}(y)^{-1}\{(b(x)-b(y))-(\tilde{b}(x)-\tilde{b}(y))A(x)\}$$

By the boundedness of b^{-1} and A, A satisfies (1.4). Since B and A is bounded and satisfy (1.4), so is $B(I-A^*A)B^*=\tilde{B}_2^2$.

Lemma 3.2 (**Theorem 5.2.2,** [**9**]). *Let $a(t)$ be a symmetric matrix valued function of $t\in\boldsymbol{R}$, which satisfies*

$$|a(t)-a(s)|\leqq K|t-s| \qquad t,\ s\in\boldsymbol{R}\,;$$

$$\exists\delta>0 \quad a(t)\geqq\delta I \qquad \text{for any } t\in\boldsymbol{R}\,.$$

Then we have

$$|a^{1/2}(t)-a^{1/2}(s)|\leqq CK|t-s|\,,$$

where C is a constant. □

By (ii) of Assumptions A (5), using this lemma, we can see that $\hat{B}_2$ satisfies (1.4).

This completes the proof of Proposition 1.2. □

References

[1] Friedman, A., *Stochastic differential equations and applications*, **1**, Academic Press, New York, 1975.

[2] Haussmann, U. G. and Pardoux, E., A conditionally almost linear filtering problem with non-Gaussian initial condition, *Stochastics*, **23** (1988), 241–275.

[3] Ikeda, N. and Watanabe, S., *Stochastic differential equations and diffusion proc-*

esses, North Holland/Kodansha, Amsterdam/Tokyo, 1981.

[4] Kalman, R. E. and Bucy, R. S., New results in linear filtering and prediction theory, *Trans. ASME J. Basic Engrg.*, **83D** (1961), 95–108.

[5] Kolodziej, W. J. and Mohler, R. R., State estimation and control of conditionally linear systems, *SIAM J. Control and Optimization*, **24** (1986), 497–508.

[6] Liptser, R. S. and Shiryayev, A. N., *Statistics of Random Processes I–General Theory*, Springer-Verlag, New York, 1977.

[7] Liptser, R. S. and Shiryayev, A. N., *Statistics of Random Processes II–Applications*, Springer-Verlag, New York, 1978.

[8] Makowski, A. M., Filtering formulae for partially observed linear systems with non-Gaussian initial conditions, *Stochastics*, **16** (1986), 1–24.

[9] Strook, D. W. and Varadhan, S. R. S., *Multidimensional diffusion processes*, Springer-Verlag, New York, 1979.