Filtering of Nonlinear Stochastic Feedback Systems

F. Carravetta[1] A. Germani[1,2] R. S. Liptser[3] C. Manes[1,2]

[1] Istituto di Analisi dei Sistemi ed Informatica del CNR, IASI-CNR
Viale Manzoni 30, 00185 Roma - Italy, {carravetta,germani}@iasi.rm.cnr.it

[2] Dipartimento di Ingegneria Elettrica, Università degli Studi dell'Aquila
Poggio di Roio, 67040 L'Aquila - Italy, {manes,germani}@ing.univaq.it

[3] Department of Electrical Engineering-Systems, Tel Aviv University
69978 Tel Aviv, Israel, liptser@eng.tau.ac.il

Abstract

This paper concerns the filtering problem for the class of stochastic nonlinear systems on which an output feedback can be closed. It is proven that the optimal filter for the open-loop system remains optimal when the feedback is closed.

1. Introduction

This paper considers the class of stochastic systems described by the equations:

$$\begin{aligned} dX_t &= f(t, X_t, u(t, Y_{[0,t]}))dt + dW'_t, \\ dY_t &= h(t, X_t, u(t, Y_{[0,t]}))dt + dW''_t, \end{aligned} \tag{1.1}$$

subject to random initial condition X_0 and $Y_0 = 0$, where $X_t \in I\!R^d$ is the system state, $Y_t \in I\!R^m$ is the observation process, $u(t, Y_{[0,t]}) \in I\!R^p$ is the input function depending for every t on $Y_{[0,t]} = \{Y_s, 0 \leq s \leq t\}$, $W'_t \in I\!R^{\nu_1}$, $W''_t \in I\!R^{\nu_2}$ are independent Wiener processes with independent components, and f, h are vector functions of suitable dimensions. All vectors are vector-columns.

Let S_t denote the *signal* to be estimated, that is a function of the state defined as

$$S_t = F(X_t), \tag{1.2}$$

where $F : I\!R^d \to I\!R^{d'}$ is a bounded measurable vector function.

Throughout the paper we will use the following *subscript notation*: for a given function $\xi(t), t \geq 0$, we shall denote with $\xi_{[0,t]}$ the whole trajectory of $\{\xi_s, 0 \leq s \leq t\}$. So, the term $u(t, Y_{[0,t]})$ in the system (1.1) represents a causal map of the observation process into the input, describing a behaviour of some control device called a *controller*, and therefore the system (1.1) is under a *feedback law*, and we will refer to such a system as the *closed loop* system.

If in system (1.1) the term $u(t, Y_{[0,t]})$ is replaced by $u(t, \phi_{[0,t]})$, where $\phi(t)$, $t \geq 0$, is a continuous function, we have an *open-loop* system. We shall denote by X_t^ϕ, Y_t^ϕ the solution of the open-loop system, corresponding to the given input function $\phi(t)$:

$$\begin{aligned} dX_t^\phi &= f(t, X_t^\phi, u(t, \phi_{[0,t]}))dt + dW'_t, \\ dY_t^\phi &= h(t, X_t^\phi, u(t, \phi_{[0,t]}))dt + dW''_t, \end{aligned} \tag{1.3}$$

and by $S_t^\phi = F(X_t^\phi)$ the corresponding signal to be estimated. Assume for every fixed t there is a function $\Psi_t(y_{[0,t]}; \phi_{[0,t]})$, $(y_t, \phi(t),\ t \geq 0$, are continuous vector functions valued in $I\!R^p)$ such that

$$\Psi_t(Y^\phi_{[0,t]}; \phi_{[0,t]}) = E\big(S_t^\phi / Y^\phi_{[0,t]}\big).$$

This is the *open-loop filter*, i.e. the optimal filter for the open-loop system (1.3), that is forced by the system output and by the forcing term ϕ. For every t assume also there exists a function $\Phi_t(y_{[0,t]})$ $(Y_t, t \geq 0$, is continuous vector function valued in $I\!R^p)$ such that

$$\Phi_t(Y_{[0,t]}) = E\big(S_t / Y_{[0,t]}\big),\ P\text{-a.s.}$$

This is the *closed-loop filter*, i.e. the optimal filter for the closed-loop system (1.1), that is forced by the system output only.

The following question arises:

$$\Psi_t(Y_{[0,t]}; Y_{[0,t]}) \stackrel{?}{=} \Phi_t(Y_{[0,t]}),\ P\text{-a.s.}, \tag{1.6}$$

stated in other words: if we apply the *open-loop filter to the closed-loop system*, then does the estimate agree with the optimal state-estimate for the closed-loop system?

The question if (1.6) holds or not is not only interesting by itself, but is important in many applications. For instance, in all cases in which a finite-dimensional filter exists for the open loop system (see [8]) identity (1.6) proves that the filter remains optimal and finite-dimensional also when the feedback is closed. Another interesting application is when $\Phi_t(Y_{[0,t]})$ is computed by the Monte-Carlo method via $\Psi_t(Y^\phi_{[0,t]}; \phi_{[0,t]})$.

We point out that an affirmative answer to the question (1.6) is known for a class so called *conditionally linear systems*, (see [6], [7] and references therein included) and also for linear-Gaussian systems under nonlinear feedback [2]) important from an application point of view.

2. Bayes' formulas for $\Psi_t(Y^\phi_{[0,t]};\phi_{[0,t]})$ and $\Phi_t(Y_{[0,t]})$

2.1: Assumptions and notations

Through the paper we will assume that

1. $h(t,0,u(t,0))$ is a bounded function;

2. there are an increasing function $L(t)$ and a measure $\mu(ds)$ on $I\!R_+$, $\int_0^t \mu(ds)<\infty$, $t>0$ such that, denoting with $\alpha(t,x,u(t,y_{[0,t]}))$ either $f(t,x,u(t,y_{[0,t]}))$ or $h(t,x,u(t,y_{[0,t]}))$

$$\begin{aligned}&\big|\alpha(t,x',u(t,y'_{[0,t]}))-\alpha(t,x'',u(t,y''_{[0,t]}))\big|\\ &\le L(t)\Big(\parallel x'-x''\parallel+\int_0^t\parallel y'_s-y''_s\parallel\mu(ds)\Big),\end{aligned}\tag{2.1}$$

where $\parallel . \parallel$ is the Euclidean norm in $I\!R^p$.

Under **1.** and **2.** both systems of the Itô equations (1.1) and (1.3) obey the unique strong solution defined by the initial condition X_0 and Wiener process (W'_t,W''_t), $t\ge 0$. Moreover the first equation from (1.3) obeys the unique strong solution, so there is a measurable functional $Q(\ldots)$ such that

$$\begin{aligned}X^\phi_t&=Q\big(t,X_0,W'_{[0,t]},u_{[0,t]}(\phi)\big),\\ X_t&=Q\big(t,X_0,W'_{[0,t]},u_{[0,t]}(Y)\big),\end{aligned}\tag{2.2}$$

where $u_{[0,t]}(\phi)=\{u(s,\phi_{[0,s]}),s\le t\}$ and $u_{[0,t]}(Y)=\{u(s,Y_{[0,s]}),s\le t\}$.

2.2: The Bayes formula for $\Psi_t(Y^\phi_{[0,t]},\phi_{[0,t]})$

For a detailed description of $\Psi_t(Y^\phi_{[0,t]},\phi_{[0,t]})$ we introduce the following objects:

(i) $G=G(x)$ the distribution function of the vector X_0;

(ii) $\lambda_{W'}$ the Wiener measure on $\mathcal{C}_{[0,\infty)}(I\!R^p)$ being the distribution of $(W'_t)_{t\ge0}$;

(iii)

$$\begin{aligned}&\mathcal{U}_t\big(X_0,W'_{[0,t]},W''_{[0,t]},u_{[0,t]}(\phi)\big)=\\ &\exp\Big\{\int_0^t h^*\big(s,Q(s,X_0,W'_{[0,s]},u_{[0,s]}(\phi)),u(s,\phi_{[0,s]})\big)dW''_s\\ &-\frac12\int_0^t\parallel h\big(s,Q(s,X_0,W'_{[0,s]},u_{[0,s]}(\phi)),u(s,\phi_{[0,s]})\big)\parallel^2ds\Big\},\end{aligned}\tag{2.3}$$

where * is the transposition symbol and $\parallel h\parallel^2=h^*h$;

(iv)

$$\mathcal{Z}_t\big(\phi,W''\big)=\int_{\mathcal{C}_{[0,t]}(I\!R^d\times I\!R^d)}\mathcal{U}_t\big(x,w_{[0,t]},W''_{[0,t]},u_{[0,t]}(\phi)\big)d\lambda_{W'}(w)dG(x)\tag{2.4}$$

(v) $$\rho^{\phi,t}(X_0,W',Y^\phi)=\frac{\mathcal{U}_t\big(X_0,W'_{[0,t]},Y^\phi_{[0,t]},u_{[0,t]}(\phi)\big)}{\mathcal{Z}_t\big(\phi,Y^\phi\big)}.$$

Lemma 2.1. *Under* **1.** *and* **2.**

$$\begin{aligned}&\Psi_t(Y^\phi_{[0,t]};\phi_{[0,t]})\\ &=\int_{\mathcal{C}_{[0,t]}(I\!R^d\times I\!R^d)}F\big(Q(t,x,w_{[0,t]},u_{[0,t]}(\phi)\big)\\ &\qquad\cdot\rho^{\phi,t}(x,w,Y^\phi)d\lambda_{W'}(w)dG(x).\end{aligned}\tag{2.5}$$

Proof. By virtue of (2.2) the second equation from (1.3) is transformed into

$$dY^\phi_t=h\big(t,Q(t,X_0,W'_{[0,t]},u_{[0,t]}(\phi),u(t,\phi_{[0,t]})\big)dt+dW''_t.$$

Further, we deal with two triples of random processes

$$\begin{aligned}&\big(\alpha_t\equiv X_0,W'_t,Y^\phi_t\big)_{t\ge0}\\ &\big(\beta_t\equiv X_0,W'_t,W''_t\big)_{t\ge0}.\end{aligned}$$

Denote by λ_{X_0,W',Y^ϕ} and $\lambda_{X_0,W',W''}$ the distributions of these processes respectively and by $\lambda^t_{X_0,W',Y^\phi}$ and $\lambda^t_{X_0,W',W''}$ their restrictions on $\mathcal{C}_{[0,t]}(I\!R^d\times I\!R^d\times I\!R^p)$. Under the assumptions **1.** and **2.** for any $\phi\in\mathcal{C}_{[0,\infty)}(I\!R^p)$ we have

$$\begin{aligned}&\int_0^t\parallel h\big(s,Q(s,X_0,W'_{[0,s]},u_{[0,s]}(\phi)),u(s,\phi_{[0,s]})\big)\parallel^2ds\\ &\quad<\infty,\quad P\text{-a.s.},\ t>0.\end{aligned}$$

Hence, by Theorem 7.20 and comments from Subsection **7.6.4** after this theorem in [1] it holds $\lambda^t_{X_0,W',Y^\phi}\sim\lambda^t_{X_0,W',W''}$ with the density

$$\begin{aligned}&\frac{d\lambda^t_{X_0,W',Y^\phi}}{d\lambda^t_{X_0,W',W''}}(X_0,W',W'')\\ &=\mathcal{U}_t\big(X_0,W'_{[0,t]},W''_{[0,t]},u_{[0,t]}(\phi)\big)\end{aligned}\tag{2.6}$$

and $\mathcal{U}_t\big(X_0,W'_{[0,t]},W''_{[0,t]},u_{[0,t]}(\phi)\big)$ defined in **(iii)**.

Denote by λ_{Y^ϕ} and $\lambda_{W''}$ the distributions of $(Y^\phi_t)_{t\ge0}$ and $(W''_t)_{t\ge0}$ respectively and by $\lambda^t_{Y^\phi}$ and $\lambda^t_{W''}$ their restrictions on $\mathcal{C}_{[0,t]}(I\!R^p)$. Since $\lambda^t_{Y^\phi}$, $\lambda^t_{W''}$ are marginal distributions of $\lambda^t_{X_0,W',Y^\phi}$, $\lambda^t_{X_0,W',W''}$, it holds $\lambda^t_{Y^\phi}\sim\lambda^t_{W''}$ with the density

$$\frac{d\lambda^t_{Y^\phi}}{d\lambda^t_{W''}}(W'')=\mathcal{Z}_t\big(\phi,W''\big)$$

and $\mathcal{Z}_t(\phi, W'')$ defined in **(iv)**.

Now, following the proof of Lemma 11.5 in [2], we get the required the Bayes formula (2.5). ■

2.3: The Bayes formula for $\Phi_t(y_{[0,t]})$

For a detailed description of $\Phi_t(y_{[0,t]})$ we introduce the following objects:

(i) $G = G(x)$ the distribution function of the vector X_0;

(ii) $\lambda_{W'}$ the Wiener measure on $\mathcal{C}_{[0,\infty)}(I\!R^p)$ being the distribution of $(W'_t)_{t\ge 0}$;

(iii')

$$\mathcal{U}_t\big(X_0, W'_{[0,t]}, W''_{[0,t]}, u_{[0,t]}(W'')\big) = \exp\Big\{\int_0^t h^*\big(s, Q(s, X_0, W'_{[0,s]}, u_{[0,s]}(W'')), u(s, W''_{[0,s]})\big) dW''_s - \frac{1}{2}\int_0^t \| h\big(s, Q(s, X_0, W'_{[0,s]}, u_{[0,s]}(W'')), u(s, W''_{[0,s]})\big) \|^2 ds\Big\} \tag{2.7}$$

(iv')

$$\mathcal{Z}_t(W'', W'') = \int_{\mathcal{C}_{[0,t]}(I\!R^d\times I\!R^d)} \mathcal{U}_t\big(x, w_{[0,t]}, W''_{[0,t]}, u_{[0,t]}(W'')\big)\big) d\lambda_{W'}(w) dG(x);$$

(v')

$$\rho^{Y,t}(X_0, W', Y) = \frac{\mathcal{U}_t\big(X_0, W'_{[0,t]}, Y_{[0,t]}, u_{[0,t]}(Y)\big)}{\mathcal{Z}_t(Y,Y)}.$$

Lemma 2.2. *Under* **1.** *and* **2.**

$$\Phi_t(Y_{[0,t]}) = \int_{\mathcal{C}_{[0,t]}(I\!R^d\times I\!R^d)} F\big(Q(t, x, w_{[0,t]}, u_{[0,t]}(Y)\big) \rho^{Y,t}(x, w, Y) d\lambda_{W'}(w) dG(x). \tag{2.8}$$

Proof. Practically we repeat here the proof of Lemma 2.5. By virtue of (2.2) the second equation from (1.1) is transformed into

$$dY_t = h\big(t, Q(t, X_0, W'_{[0,t]}, u_{[0,t]}(Y), u(t, Y_{[0,t]})\big) dt + dW''_t.$$

Let λ_{X_0,W',Y^ϕ} and $\lambda_{X_0,W',W''}$ be distributions of random processes $\big(\beta_t \equiv X_0, W'_t, Y_t\big)_{t\ge 0}$ and $\big(\alpha_t \equiv X_0, W'_t, W''_t\big)_{t\ge 0}$ and $\lambda^t_{X_0,W',Y}$, $\lambda^t_{X_0,W',W''}$ be the corresponding restrictions on $\mathcal{C}_{[0,t]}(I\!R^d\times I\!R^d\times I\!R^p)$. Let

$$\begin{aligned} &h_1(s, X_0, W', W'') \\ &= h\big(s, Q(s, X_0, W'_{[0,s]}, u_{[0,s]}(W'')), u(s, W''_{[0,s]})\big), \\ &h_2(s, X_0, W', Y) \\ &= h\big(s, Q(s, X_0, W'_{[0,s]}, u_{[0,s]}(Y)), u(s, Y_{[0,s]})\big) \end{aligned}$$

Under the assumptions **1.** and **2.** we have (P-a.s., $t > 0$)

$$\int_0^t \| h_1(s, X_0, W', W'') \|^2 ds < \infty,$$

$$\int_0^t \| h_2(s, X_0, W', Y) \|^2 ds < \infty.$$

Hence, by Theorem 7.20 and comments from Subsection 7.6.4 after this theorem in [1] it holds $\lambda^t_{X_0,W',Y} \sim \lambda^t_{X_0,W',W''}$ with the density

$$\begin{aligned} &\frac{d\lambda^t_{X_0,W',Y^\phi}}{d\lambda^t_{X_0,W',W''}}(X_0, W', W'') \\ &= \mathcal{U}_t\big(X_0, W'_{[0,t]}, W''_{[0,t]}, u_{[0,t]}(W'')\big) \end{aligned} \tag{2.9}$$

where $\mathcal{U}_t\big(X_0, W'_{[0,t]}, W''_{[0,t]}, u_{[0,t]}(W'')\big)$ defined in **(iii')**. Denote by λ_Y and $\lambda_{W''}$ the distributions of $(Y)_{t\ge 0}$ and $(W''_t)_{t\ge 0}$ respectively and by λ^t_Y and $\lambda^t_{W''}$ their restrictions on $\mathcal{C}_{[0,t]}(I\!R^p)$. Since λ^t_Y, $\lambda^t_{W''}$ are marginal distributions of $\lambda^t_{X_0,W',Y}$, $\lambda^t_{X_0,W',W''}$, it holds $\lambda^t_Y \sim \lambda^t_{W''}$ with the density

$$\frac{d\lambda^t_{Y'}}{d\lambda^t_{W''}}(W'') = \mathcal{Z}_t(W'', W'')$$

and $\mathcal{Z}_t(W'', W'')$ defined in **(iv')**.

Now, following the proof of Lemma 11.5 in [2], equation (2.8) is obtained. ■

3. Verification of (1.6)

Theorem 3.1. *Under* **1.** *and* **2.** (1.6) *holds.*

The proof of this theorem uses an auxiliary result given below as lemma.

For notation convenience set

$$H(s, \phi) = h\big(s, Q(s, X_0, W'_{[0,s]}, u_{[0,s]}(\phi)), u(s, \phi_{[0,s]})\big) \tag{3.1}$$

Lemma 3.2. *For every* $t > 0$, *it holds* P*-a.s.*

$$\Big(\int_0^t h^*\big(s, Q(s, X_0, W'_{[0,s]}, u_{[0,s]}(\phi)), u(s, \phi_{[0,s]})\big) dY^\phi_s\Big)\Big|_{\phi\equiv Y} = \int_0^t h^*\big(s, Q(s, X_0, W'_{[0,s]}, u_{[0,s]}(Y)), u(s, Y_{[0,s]})\big) dY_s. \tag{3.2}$$

Proof. Since it is

$$\int_0^t H^*(s,\phi) dY^\phi_s = \int_0^t H^*(s,\phi) dW''_s + \int_0^t \|H(s,\phi)\|^2 ds$$

$$\int_0^t H^*(s,Y) dY_s = \int_0^t H^*(s,Y) dW''_s + \int_0^t \|H(s,Y)\|^2 ds,$$

it suffices to show that P-a.s.

$$\begin{aligned} &\Big(\int_0^t \|H(s,\phi)\|^2 ds\Big)\Big|_{\phi\equiv Y} = \int_0^t \|H(s,Y)\|^2 ds \\ &\Big(\int_0^t H^*(s,\phi) dW''_s\Big)\Big|_{\phi\equiv Y} = \int_0^t H^*(s,Y) dW''_s. \end{aligned} \tag{3.3}$$

The first equality in (3.3) is obvious. To verify the validity of the second one, let us introduce time moments $s_i^n = \frac{t}{n} i$, $i = 1, \dots, n$ and set $H_n(s,\phi) = H(s_i^n, \phi)$; $s \in [s_i^n, s_{i+1}^n)$, $i = 0, \dots, n$. It is clear that for any $n \geq 1$ we have

$$\Big(\int_0^t H_n^*(s,\phi) dW_s''\Big)\Big|_{\phi \equiv Y} = \int_0^t H_n^*(s,Y) dW_s''.$$

Using the latter equality, write

$$\begin{aligned}
&\Big|\Big(\int_0^t H^*(s,\phi) dW_s''\Big)\Big|_{\phi \equiv Y} - \int_0^t H^*(s,Y) dW_s''\Big| \\
&\quad \leq \Big|\Big(\int_0^t \Big\{H^*(s,\phi) - H_n^*(s,\phi)\Big\} dW_s''\Big)\Big|_{\phi \equiv Y}\Big| \\
&\qquad + \Big|\int_0^t \Big\{H^*(s,Y) - H_n^*(s,Y)\Big\} dW_s''\Big|.
\end{aligned}$$

Since $\int_0^t \|H(s,Y) - H_n(s,Y)\|^2 ds \to 0$, $n \to \infty$, the second term in the right side of the above inequality converges to zero in probability. Thus, it remains to establish only that for every $\varepsilon > 0$

$$\lim_n P\Big(\Big|\int_0^t \Big\{H^*(s,\phi) - H_n^*(s,\phi)\Big\} dW_s''\Big| \geq \varepsilon, \phi \equiv Y\Big) = 0. \tag{3.4}$$

Write

$$\begin{aligned}
&\Big\{\Big|\int_0^t \Big\{H^*(s,\phi) - H_n^*(s,\phi)\Big\} dW_s''\Big| \geq \varepsilon, \phi \equiv Y\Big\} \\
&= \Big\{I(\phi_{[0,t]} \equiv Y_{[0,t]})\Big| \\
&\qquad \int_0^t \Big\{H^*(s,\phi) - H_n^*(s,\phi)\Big\} dW_s''\Big| \geq \varepsilon\Big\} \\
&= \Big\{I(\phi_{[0,t]} \equiv Y_{[0,t]})\Big| \\
&\int_0^t I(\phi_{[0,s]} \equiv Y_{[0,s]})\Big\{H^*(s,\phi) - H_n^*(s,\phi)\Big\} dW_s''\Big| \geq \varepsilon\Big\} \\
&\subseteq \Big\{\Big|\int_0^t I(\phi_{[0,s]} \equiv Y_{[0,s]}) \\
&\qquad \cdot \Big\{H^*(s,\phi) - H_n^*(s,\phi)\Big\} dW_s''\Big| \geq \varepsilon\Big\} \\
&= \Big\{\Big|\int_0^t I(\phi_{[0,s]} \equiv Y_{[0,s]}) \\
&\qquad \cdot \Big\{H^*(s,Y) - H_n^*(s,Y)\Big\} dW_s''\Big| \geq \varepsilon\Big\}.
\end{aligned}$$

Further, since

$$\begin{aligned}
&\int_0^t I(\phi_{[0,s]} \equiv Y_{[0,s]}) \|H(s,Y) - H_n(s,Y)\|^2 ds \\
&\quad \leq \int_0^t \|H(s,Y) - H_n(s,Y)\|^2 ds \to 0, \ n \to \infty,
\end{aligned}$$

we conclude that (3.4) is valid. ∎

Proof of Theorem 3.1. Comparing (2.5) and (2.8), one can conclude that (1.6) holds, if

$$\rho^{\phi,t}(x,w,Y^\phi)\Big|_{\phi \equiv Y} = \rho^{Y,t}(x,w,Y), \quad P\text{-a.s.}$$

Now, it remains to note only that the required property is implied by Lemma 3.2. ∎

References

[1] R.S. Liptser, A.N. Shiryayev. *Statistics of Random Processes*, Vol 1, Springer Verlag, New York, 1977.

[2] R.S. Liptser, A.N. Shiryayev. *Statistics of Random Processes*, Vol 2, Springer Verlag, New York, 1978.

[3] R.E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *J. Basic. Eng.*, vol. 1, pp. 35-45, 1960.

[4] A.V. Balakrishnan, *Kalman Filtering Theory*, Optimization Software,Inc., Publication Division, New York, 1984.

[5] B.D.O. Anderson, J.B. Moore, *Optimal Filtering*, Prentice Hall, Inc., Englewood Cliffs, N.J., 1979.

[6] W.J. Kolodziej, R.R. Mohler, "State Estimation and Control of Conditionally Linear Systems," *SIAM Journal on Control and Optimization*, Vol. 24, No. 3, pp. 497-508, May 1986.

[7] W.J. Kolodziej, R.R. Mohler, "Conditionally Linear and Non-Gaussian Processes," in *Non-Gaussian Signal Processing*, E. Wegman and S. Schwartz, eds., North-Holland, Amsterdam, 1987.

[8] M. Cohen De Lara, "Finite-Dimensional Filters. Part I: The Wei Norman Technique", "Finite-Dimensional Filters. Part II: Invariance Group Techniques" , *SIAM J. Contr. & Opt.*, Vol. 35, No. 3, pp. 980–1029.