

# The Separation Principle in Stochastic Control, Redux

Tryphon T. Georgiou and Anders Lindquist

## Abstract

Over the last 50 years a steady stream of accounts have been written on the separation principle of stochastic control. Even in the context of the linear-quadratic regulator in continuous time with Gaussian white noise, subtle difficulties arise, unexpected by many, that are often overlooked. In this paper we propose a new framework for establishing the separation principle. This approach takes the viewpoint that stochastic systems are well-defined maps between sample paths rather than stochastic processes *per se* and allows us to extend the separation principle to systems driven by martingales with possible jumps. While the approach is more in line with “real-life” engineering thinking where signals travel around the feedback loop, it is unconventional from a probabilistic point of view in that control laws for which the feedback equations are satisfied almost surely, and not deterministically for every sample path, are excluded.

## I. INTRODUCTION

ONE of the most fundamental principles of feedback theory is that the problems of optimal control and state estimation can be decoupled in certain cases. This is known as the *separation principle*. The concept was coined early on in [17], [32] and is closely connected to the idea of *certainty equivalence*; see, e.g., [38]. In studying the literature on the separation principle of stochastic control, one is struck by the level of sophistication and technical complexity. The source of the difficulties can be traced to the circular dependence between control and observations. The goal of this paper is to present a rigorous approach to the separation principle in continuous time which is rooted in the engineering view of systems as maps between signal spaces.

The most basic setting begins with a linear system

$$\begin{cases} dx = A(t)x(t)dt + B_1(t)u(t)dt + B_2(t)dw \\ dy = C(t)x(t)dt + D(t)dw \end{cases} \quad (1)$$

with a state process  $x$ , an output process  $y$  and a control  $u$ , where  $w(t)$  is a vector-valued Wiener process,  $x(0)$  is a zero-mean Gaussian random vector independent of  $w(t)$ ,  $y(0) = 0$ , and  $A$ ,  $B_1$ ,  $B_2$ ,  $C$ ,  $D$  are matrix-valued functions of compatible dimensions, which we take to be continuous of bounded variation. Moreover,  $DD'$  is nonsingular on the interval  $[0, T]$ , and if we want the noise processes in the state and output equations to be independent, as often is assumed but not required here, we take  $B_2D' \equiv 0$ . All random variables and processes are defined over a common complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ .

The control problem is to design an output feedback law

$$\pi : y \mapsto u \quad (2)$$

over the window  $[0, T]$  which maps the observation process  $y$  to the control input  $u$ , in a nonanticipatory manner, so that the value of the functional

$$J(u) = E \left\{ \int_0^T x(t)'Q(t)x(t)dt + \int_0^T u(t)'R(t)u(t)dt + x(T)'Sx(T) \right\} \quad (3)$$

This research was supported by grants from AFOSR, NSF, VR, SSF and the Göran Gustafsson Foundation.

T. T. Georgiou is with the Department of Electrical & Computer Engineering, University of Minnesota, Minneapolis, Minnesota 55455, USA, tryphon@umn.edu

A. Lindquist is with the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, and the Center for Industrial and Applied Mathematics (CIAM) and the ACCESS Linnaeus Center, Royal Institute of Technology, 100 44 Stockholm, Sweden, alq@kth.se

is minimized, where  $Q$  and  $R$  are continuous matrix functions of bounded variation,  $Q(t)$  is positive semi-definite and  $R(t)$  is positive definite for all  $t$ . How to choose the admissible class of control laws  $\pi$  has been the subject of much discussion in the literature [27]. The conclusion, under varying conditions, has been that  $\pi$  can be chosen to be linear in the data and, more specifically, in the form

$$u(t) = K(t)\hat{x}(t), \quad (4)$$

where  $\hat{x}(t)$  is the Kalman estimate of the state vector  $x(t)$  obtained from the Kalman filter

$$d\hat{x} = A(t)\hat{x}(t)dt + B_1(t)u(t)dt + L(t)(dy - C(t)\hat{x}(t)dt), \quad \hat{x}(0) = 0, \quad (5)$$

and the gains  $K$  and  $L$  computed by solving to a pair of dual Riccati equations.

A result of this kind is far from obvious, and the early literature was marred by treatments of the separation principle where the non-Gaussian element introduced by an *a priori* nonlinear control law  $\pi$  was overlooked. The subtlety lies in excluding the possibility that a nonlinear controller extracts more information from the data than it is otherwise possible. This point will be explained in detail in Section II, where a brief historical account of the problem will be given. Early expositions of the separation principle often fall in one of two categories: either the subtle issues are overlooked and inadmissible shortcuts are taken; or the treatment is mathematically quite sophisticated and technically very demanding. The short survey in Section II will thus serve the purpose of introducing the theoretical challenges at hand, as well as setting up notation.

In this paper we take the point of view that feedback laws (2) should act on sample paths of the stochastic process  $y$  rather than on the process itself. This is motivated by engineering thinking where systems and feedback loops process signals. Thus, our *key* assumption on admissible control laws (2) is that the resulting feedback loop is *deterministically well-posed* in the sense that the feedback equations admit a unique solution path-wise which causally depends on the input. For this class of control laws we prove that the separation principle stated above holds and moreover that it extends to systems driven by general martingale noise. More precisely, in this non-Gaussian situation the Wiener process  $w$  in (1) is replaced by an arbitrary martingale process with possible jumps such as a Poisson process martingale; see, e.g., [19, p. 87]. Then, we only need to exchange the (linear) Kalman estimate  $\hat{x}$  by the strict sense conditional mean

$$\hat{x}(t) = E\{x(t) \mid \mathcal{Y}_t\}, \quad (6)$$

where

$$\mathcal{Y}_t := \sigma\{y(\tau), \tau \in [0, t]\}, \quad 0 \leq t \leq T, \quad (7)$$

is the *filtration* generated by the output process; i.e., the family of increasing sigma fields representing the data as it is produced. The estimate  $\hat{x}$  needs to be defined with care so that it constitutes a sufficiently regular stochastic process and realized by a map acting on observations [2, page 17], [11]. Unfortunately, the results in the present paper come at a cost since our key assumption of well-posedness excludes control laws for which the feedback system fails to be defined sample-wise. Existence of strong solutions of the feedback equations is not enough to ensure well-posedness in our sense as we will discuss below. In addition, the condition of deterministic well-posedness is often difficult to verify. Yet, besides the fact that we prove the separation principle for general martingale noise, the sample-wise viewpoint provides a simple explanation of why the separation principle may hold in the first place.

Before proceeding we recast the system model (1) in an integrated form which allows similar conclusions for more general linear systems in a unified setting. To this end, let

$$z(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}.$$

System (1) can now be expressed in the form

$$\begin{cases} z(t) = z_0(t) + \int_0^t G(t, \tau)u(\tau)d\tau \\ y(t) = Hz(t), \end{cases} \quad (8)$$

where  $z_0$  is the process  $z$  obtained by setting  $u = 0$  and  $G$  is a Volterra kernel. This integrated form encompasses a considerably wider class of controlled linear systems which includes delay-differential equations, following [26], [27], which will be taken up in Section VI. The corresponding feedback

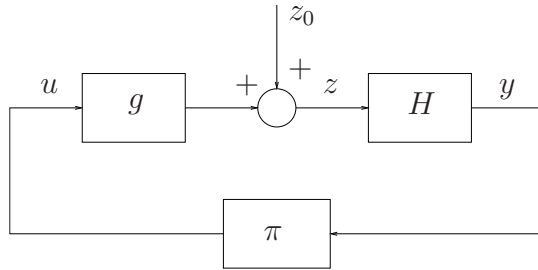


Fig. 1. A feedback interconnection.

configuration is shown in Figure 1 where  $g$  represents the Volterra operator

$$g : (t, u) \mapsto \int_0^t G(t, \tau)u(\tau)d\tau, \quad (9)$$

and  $H$  is a constant matrix. As usual, Figure 1 is a graphical representation of the algebraic relationship

$$z = z_0 + g\pi Hz. \quad (10)$$

For the particular model in (1),  $H = [0, I]$ , but in general  $H$  could be any matrix or linear system. Setting  $z := x$  and  $H = I$  we obtain the special case of complete state information.

In a stochastic setting, the feedback equation (10) is said to have a unique strong solution if there exists a non-anticipating function  $F$  such that  $z = F(z_0)$  satisfies (10) with probability one and all other solutions coincide with  $z$  with probability one. It is important to note that in our sample-wise setting we require more, namely that such a unique solution exists and that (10) holds for *all*  $z_0$ , not only “almost all.” Consequences of this requirement will be further elaborated upon below.

The outline of the paper is as follows. In Section II we begin by reviewing the standard quadratic regulator problem and pointing out the subtleties created by possible nonlinearities in the control law. We then review several strategies in the literature to establish a separation principle, chiefly restricting the class of admissible controls. Section III defines notions of signals and systems used in our framework, and in Section IV we establish necessary conditions for a feedback loop to make sense and deduce a basic fact about propagation of information in the loop through linear components. In Section V we state and prove our main results on the separation principle for linear-quadratic regulator problems, allowing also for more general martingale noise. Finally, in Section VI we prove a separation theorem for delay systems with Gaussian martingale noise.

## II. HISTORICAL REMARKS

A common approach to establishing the basic separation principle stated at the beginning of Section I is a completion-of-squares argument similar to the one used in deterministic linear-quadratic-regulator theory; see e.g. [1]. For ease of reference, we briefly review this construction. Given the system (1) and the solution of the matrix Riccati equation

$$\dot{P} = -A'P - PA + PB_1R^{-1}B_1'P - Q, \quad P(T) = S. \quad (11a)$$

Itô’s differential rule (see, e.g., [19], [31]) yields

$$d(x'Px) = x'\dot{P}xdt + 2x'Pdx + \text{tr}(B_2'PB_2)dt,$$

where  $\text{tr}(M)$  denotes the trace of the matrix  $M$ . Then from (1) and (11a) it readily follows that

$$d(x'Px) = [-x'Qx - u'Ru + (u - Kx)'R(u - Kx)]dt + \text{tr}(B_2'PB_2)dt + 2x'PB_2dw,$$

where

$$K(t) := -R(t)^{-1}B_1(t)'P(t). \quad (11b)$$

Integrating this from 0 to  $T$  and taking mathematical expectation, we obtain the following expression for the cost functional (3):

$$J(u) = E \left\{ x(0)'P(0)x(0) + \int_0^T (u - Kx)'R(u - Kx)dt \right\} + \int_0^T \text{tr}(B_2'PB_2)dt. \quad (12)$$

To ensure that  $\int_0^T x'PB_2dw$  has zero expectation, we need to check that the integrand is square integrable almost surely. It is clear that  $u$  is square integrable for otherwise  $J(u) = \infty$ . Then the state process

$$x(t) = x_0(t) + \int_0^t \Phi(t, s)B_1(s)u(s)ds \quad (13)$$

is square integrable as well. Here  $x_0$  is the (square integrable) state process corresponding to  $u = 0$ , and  $\Phi(t, s)$  is the transition matrix function of the system (1).

Now, if we had complete state information with (1) replaced by

$$\begin{cases} dx = A(t)x(t)dt + B_1(t)u(t)dt + B_2(t)dw \\ y = x \end{cases} \quad (14)$$

we could immediately conclude that the feedback law

$$u(t) = K(t)x(t) \quad (15)$$

is optimal, because the last term in (12) does not depend on the control. However, when we have incomplete state information with the control being a function of the observed process  $\{y(s); 0 \leq s \leq t\}$ , things become more complicated. Mathematically we formalize this by having any control process adapted to the filtration (7); i.e., having  $u(t)$   $\mathcal{Y}_t$ -measurable for each  $t \in [0, T]$ . Then, setting

$$\tilde{x}(t) := x(t) - \hat{x}(t) \quad (16)$$

with  $\hat{x}$  given by (6), we have  $E\{[u(t) - K(t)\hat{x}(t)]\tilde{x}(t)'\} = 0$ , and therefore

$$E \int_0^T (u - Kx)'R(u - Kx)dt = E \int_0^T [(u - K\hat{x})'R(u - K\hat{x}) + \text{tr}(K'RK\Sigma)]dt, \quad (17)$$

where  $\Sigma$  is the covariance matrix

$$\Sigma(t) := E\{\tilde{x}(t)\tilde{x}(t)'\}. \quad (18)$$

A common mistake in the early literature on the separation principle is to assume *without further investigation* that  $\Sigma$  does not depend on the choice of control. Indeed, if this were the case, it would follow directly that (12) is minimized by choosing the control as (4), and the proof of the separation principle would be immediate. (Of course, in the end this will be the case under suitable conditions, but this has to be proven.) This mistake probably originates from the observation that the control term in (13) cancels when forming (16) so that

$$\tilde{x}(t) = \tilde{x}_0(t) := x_0(t) - \hat{x}_0(t), \quad (19)$$

where

$$\hat{x}_0(t) := E\{x_0(t) \mid \mathcal{Y}_t\}. \quad (20)$$

However, in this analysis, we have *not* ruled out that  $\hat{x}_0$  depends on the control or, what would follow from this, that the filtration (7) does. A detailed discussion of this conundrum can be found in [27]. In fact, since the control process  $u$  is in general a *nonlinear* function of the data and thus non-Gaussian, then so is the output process  $y$ .<sup>1</sup> Consequently, the conditional expectation (20) might not in general coincide with the *wide sense* conditional expectation obtained by projections of the components of  $x_0(t)$  onto the closed linear span of the components of  $\{y(\tau), \tau \in [0, t]\}$ , and therefore, *a priori*, it could happen that  $\hat{x}$  is not generated by the Kalman filter (5).

To avoid these problems one might begin by uncoupling the feedback loop as described in Figure 2, and determine an optimal control process in the class of stochastic processes  $u$  that are adapted to the family of sigma fields

$$\mathcal{Y}_t^0 := \sigma\{y_0(\tau), \tau \in [0, t]\}, \quad 0 \leq t \leq T, \quad (21)$$

i.e., such that, for each  $t \in [0, T]$ ,  $u(t)$  is a function of  $y_0(s)$ ;  $0 \leq s \leq T$ . Such a problem, where one optimizes over the class of all control processes adapted to a fixed filtration, was called a *stochastic open loop (SOL) problem* in [27]. In the literature on the separation principle it is not uncommon to assume from the outset that the control is adapted to  $\{\mathcal{Y}_t^0\}$ ; see, e.g., [6, Section 2.3], [16], [40].

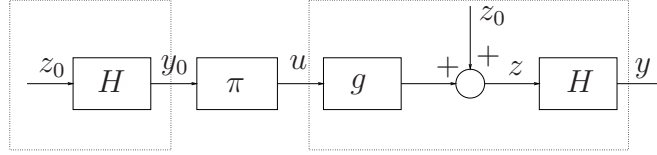


Fig. 2. A stochastic open loop (SOL) configuration.

In [27] it was suggested how to embed the class of admissible controls in various SOL classes in a problem-dependent manner, and then construct the corresponding feedback law. More precisely, in the present context, the class of admissible feedback laws was taken to consist of the nonanticipatory functions  $u := \pi(y)$  such that the feedback loop

$$z = z_0 + g\pi Hz \quad (22)$$

has a unique solution  $z_\pi$  and  $u = \pi(Hz_\pi)$  is adapted to  $\{\mathcal{Y}_t^0\}$ . Next, we shall give a few examples of specific classes of feedback laws that belong to this general class.

*Example 1:* It is common to restrict the admissible class of control laws to contain only linear ones; see, e.g., [12]. In a more general direction, let  $\mathcal{L}$  be the class

$$(\mathcal{L}) \quad u(t) = \bar{u}_0(t) + \int_0^t F(t, \tau) dy, \quad (23)$$

where  $\bar{u}$  is a deterministic function and  $F$  is an  $L_2$  kernel. In this way, the Gaussian property will be preserved, and  $\hat{x}$  will be generated by the Kalman filter (5). Then it follows from (1) and (5) that  $\tilde{x}$  is generated by

$$d\tilde{x} = (A - LC)\tilde{x}dt + (B_2 - LD)dw, \quad \tilde{x}(0) = x(0),$$

which is clearly independent of the choice of control. Then so is the error covariance (18), as desired. Even in the more general setting described by (8), it was shown in [26, pp. 95–96] that

$$\mathcal{Y}_t = \mathcal{Y}_t^0, \quad t \in [0, T], \quad (24)$$

for any  $\pi \in \mathcal{L}$ , where (21) is the filtration generated by the uncontrolled output process  $y_0$  obtained by setting  $u = 0$  in (8).

<sup>1</sup>However, the model is conditionally Gaussian given the filtration  $\{\mathcal{Y}_t\}$ ; see Remark 6.

*Example 2:* In his influential paper [41], Wonham proposed the class of control laws

$$u(t) = \psi(t, \hat{x}(t)) \quad (25)$$

in terms of the state estimate (6), where  $\psi(t, x)$  is Lipschitz continuous in  $x$ . For pedagogical reasons, we first highlight a somewhat more restrictive construction due to Kushner [21]. Let

$$\hat{\xi}_0(t) := E\{x_0(t) \mid \mathcal{Y}_t^0\}$$

be the Kalman state estimate of the uncontrolled system

$$\begin{cases} dx_0 = A(t)x_0(t)dt + B_2(t)dw \\ dy_0 = C(t)x_0(t)dt + D(t)dw \end{cases} \quad (26)$$

Here we use the notation  $\hat{\xi}_0$  to distinguish it from  $\hat{x}_0$ , defined by (20), which *a priori* depends on the control. Then the Kalman filter takes the form

$$d\hat{\xi}_0 = A\hat{\xi}_0(t)dt + L(t)dv_0, \quad \hat{\xi}_0(0) = 0$$

where the innovation process

$$dv_0 = dy_0 - C\hat{\xi}_0(t)dt, \quad v_0(0) = 0$$

generates the same filtration,  $\{\mathcal{V}_t^0\}$ , as  $y_0$ ; i.e.,  $\mathcal{V}_t^0 = \mathcal{Y}_t^0$  for  $t \in [0, T]$ . This is well-known, but a simple proof is given on page 18 in Section VI in a more general setting; see (65). Now, along the lines of (13), define

$$\hat{\xi}(t) = \hat{\xi}_0(t) + \int_0^t \Phi(t, s)B_1(s)u(s)ds,$$

where the control is chosen as

$$u(t) = \psi(t, \hat{\xi}(t)). \quad (27)$$

Since  $\psi$  is Lipschitz,  $\hat{\xi}$  is the unique strong solution of the stochastic differential equation

$$d\hat{\xi} = (A\hat{\xi}(t) + B_1\psi(t, \hat{\xi}(t)))dt + L(t)dv_0, \quad \hat{\xi}(0) = 0, \quad (28)$$

and it is thus adapted to  $\{\mathcal{V}_t^0\}$  and hence to  $\{\mathcal{Y}_t^0\}$ ; see, e.g., [19, p. 120]. Hence the selection (27) of control law forces  $u$  to be adapted to  $\{\mathcal{Y}_t^0\}$ , and hence, due to

$$dy = dy_0 + \int_0^t C(t)\Phi(t, s)B_1(s)u(s)dsdt, \quad (29)$$

obtained from (13),  $\mathcal{Y}_t \subset \mathcal{Y}_t^0$  for  $t \in [0, T]$ . However, since the control-dependent terms cancel,

$$dv_0 = dy_0 - C\hat{\xi}_0(t)dt = dy - C\hat{\xi}(t)dt,$$

which inserted into (28) yields a stochastic differential equation, obeying the appropriate Lipschitz condition, driven by  $dy$  and having  $\hat{\xi}$  as a strong solution. Therefore,  $\hat{\xi}$  is adapted to  $\{\mathcal{Y}_t\}$ , and hence, by (27), so is  $u$ . Consequently, (29) implies that  $\mathcal{Y}_t^0 \subset \mathcal{Y}_t$  for  $t \in [0, T]$  so that actually (24) holds. Finally, this implies that  $\hat{\xi} = \hat{x}$ , and thus  $u$  is given by (25). However, it should be noted that the class of control laws (27) is a subclass of (25) as it has been constructed to make  $u$  *a priori* adapted to  $\{\mathcal{Y}_t^0\}$ . Therefore, the relevance of these results, presented in [21], for the proof in [22, page 348] is unclear. In their popular textbook [20], widely used as a reference source for the validity of the separation principle over a general class of admissible (including nonlinear) controls, Kwakernaak and Sivan prove the separation principle over a class of linear laws but claim with reference to [22], [21] that it holds “without qualification” in general [20, p. 390]. (However, see Remark 6 below.)

In his pioneering paper [41] Wonham proved the separation theorem for controls in the class (25) even with a more general cost functional than (3). However, the proof is far from simple and marred by many

technical assumptions. A case in point is the assumption that  $C(t)$  is square and has a determinant bounded away from zero, which is a serious restriction. A later proof by Fleming and Rishel [15] is considerably simpler. They also prove the separation theorem with quadratic cost functional (3) for a class of Lipschitz continuous feedback laws, namely

$$u(t) = \phi(t, y), \tag{30}$$

where  $\phi : [0, T] \times C^n[0, T] \rightarrow \mathbb{R}^m$  is a nonanticipatory function of  $y$  which is Lipschitz continuous in this argument.

*Example 3:* It is interesting to note that if there is a delay in the processing of the observed data so that, for each  $t$ ,  $u(t)$  is a function of  $y(\tau)$ ;  $0 \leq \tau \leq t - \varepsilon$ , then

$$\mathcal{Y}_t = \mathcal{Y}_t^0, \quad t \in [0, T].$$

To see this, let  $n$  be a positive integer, and suppose that  $\mathcal{Y}_t = \mathcal{Y}_t^0$  for  $t \in [0, n\varepsilon]$ . Since  $u(t)$  is  $\mathcal{Y}_{t-\varepsilon}$ -measurable on  $[0, (n+1)\varepsilon]$ , it is at the same time  $\mathcal{Y}_{t-\varepsilon}$  as well as  $\mathcal{Y}_{t-\varepsilon}^0$ -measurable. Then, since

$$y(t) = y_0(t) + \int_0^t HG(t, s)u(s)ds,$$

it follows that  $\mathcal{Y}_t = \mathcal{Y}_t^0$  for  $t \in [0, (n+1)\varepsilon]$ . Since  $\mathcal{Y}_t = \mathcal{Y}_t^0$  for  $t \in [0, \varepsilon]$ , (24) follows by induction.

*Remark 4:* This is the reason why the problem with possibly control-dependent sigma fields does not occur in the usual discrete-time formulation. Indeed, in this setting, the error covariance (18) will not depend on the control, while, as we have mentioned, some more analysis is needed to rule out that its continuous-time counterpart does. This invalidates a procedure used in several textbooks (see, e.g., [36]) in which the continuous-time  $\Sigma$  is constructed as the limit of finite difference quotients of the discrete-time  $\Sigma$ , which, as we have seen in Example 3, does not depend on the control, and which simply is the solution of a discrete-time matrix Riccati equation. However, we cannot *a priori* conclude that continuous-time  $\Sigma$  satisfies this Riccati equation. For this we need (24), or alternatively arguments such as in Remark 6. Otherwise the argument is circular.

*Remark 5:* Historically, a popular approach was introduced in Duncan and Varaiya, and Davis and Varaiya [14], [13] based on *weak solutions* of the relevant stochastic differential equation. The driving noise is Wiener and the approach utilizes the Girsanov transformation to recast the problem in a way so that the filtration of the observation process is independent of the input process (see [6, Section 2.4]). Very briefly, by an appropriate change of probability measure,

$$d\tilde{w} = B_1 u dt + B_2 dw$$

can be transformed into a new Wiener process, which in the sense of weak solutions [19] is the same as any other Wiener process. In this way, the filtration  $\{\mathcal{Y}_t\}$  can be fixed to be constant with respect to variations in the control. In this paper we do not consider weak solutions since our observation process is not arbitrary from an applications point of view.

*Remark 6:* Yet another approach to the separation principle is based on the fact that, although (1) with a nonlinear control is non-Gaussian, the model is *conditionally Gaussian* given the filtration  $\{\mathcal{Y}_t\}$  [29, Chapters 16.1]. This fact can be used to show that  $\hat{x}$  is actually generated by a Kalman filter [29, Chapters 11 and 12]. This last approach requires quite a sophisticated analysis and is restricted to the case where the driving noise  $w$  is a Wiener process.

A key point for establishing the separation principle is to identify admissible control laws for which (24) holds. For each such control law  $\pi$  we need a solution of the feedback equation (10), i.e., a pair  $(z_0, z)$  of stochastic processes that satisfies

$$z = z_0 + g\pi Hz. \tag{31}$$

Since  $z_0$  is the driving process, it is natural to seek a solution  $z$  which causally depends on  $z_0$  and is unique. If this is the case then  $z$  is a *strong solution*; otherwise it is a *weak solution*. There are well-known examples of stochastic differential equations that have only weak solutions [19, page 137], [37],

[5]. Moreover, as we have mentioned in Remark 5, weak solutions circumvent the need to establish the equivalence (24) between filtrations. Thus, it has been suggested that the framework of weak solutions is the appropriate one for control problems [34, page 149]. Yet, from an applications point of view, where the control needs to be causally dependent on observed data, this is in our view questionable. In fact, in the present paper we take an even more stringent view on the causal dependence. We require that (31) has a unique strong solution which specifies a measurable map  $z_0 \rightarrow z$  between sample-paths (cf. [19, Remark 5.2, p. 128], [34, p. 122]), thus modeling correspondence between signals – we further elaborate upon this in Section IV.

In short, we only allow control laws which are physically realizable in an engineering sense, in that they induce a signal that travels through the feedback loop. This comes at a price since there are stochastic differential equations having strong solutions that do not fall in this category (Remark 12). Moreover, verifying that a control law is admissible in our sense may be difficult to ascertain in general. On the other hand, an advantage of the approach is that the class of control laws includes discontinuous ones and allows for statements about linear systems driven by non-Gaussian noise with possible jumps. We now proceed to develop the approach and the key property of *deterministic well-posedness*.

### III. SIGNALS AND SYSTEMS

Signals are thought of as sample paths of a stochastic process with possible discontinuities. This is quite natural from several points of view. First, it encompasses the response of a typical nonlinear operation that involves thresholding and switching, and second, it includes sample paths of counting processes and other martingales. More specifically we consider signals to belong to the *Skorohod space*  $D$ ; this is defined as the space of functions which are continuous on the right and have a left limit at all points, i.e., the space of *càdlàg* functions.<sup>2</sup> It contains the space  $C$  of continuous functions as a proper subspace. The notation  $D[0, T]$  or  $C[0, T]$  emphasizes the time interval where signals are being considered.

Traditionally, the comparison of two continuous functions in the uniform topology relates to how much their graphs need to be perturbed so as to be carried onto one another by changing only the ordinates, with the time-*abscissa* being kept fixed. However, in order to metrize  $D$  in a natural manner one must recognize the effect of uncertainty in measuring time and allow a respective deformation of the time axis as well. To this end, let  $\mathcal{K}$  denote the class of strictly increasing, continuous mappings of  $[0, T]$  onto itself and let  $I$  denote the identity map. Then, for  $x, y \in D[0, T]$ ,

$$d(x, y) := \inf_{\kappa \in \mathcal{K}} \max\{\|\kappa - I\|, \|x - y\kappa\|\}$$

defines a metric on  $D[0, T]$  which induces the so called *Skorohod topology*. A further refinement so as to ensure bounds on the slopes of the chords of  $\kappa$ , renders  $D[0, T]$  separable and complete, that is,  $D[0, T]$  is a *Polish* space; see [7, Theorem 12.2].

Systems are thought of as general measurable nonanticipatory maps from  $D \rightarrow D$  sending sample paths to sample paths so that their outputs at any given time  $t$  is a measurable function of past values of the input and of time. More precisely, let

$$\Pi_\tau : x \mapsto \Pi_\tau x := \begin{cases} x(t) & \text{for } t < \tau \\ x(\tau) & \text{for } t \geq \tau. \end{cases}$$

Then, a measurable map  $f : D[0, T] \rightarrow D[0, T]$  is said to be a *system* if and only if

$$\Pi_\tau f \Pi_\tau = \Pi_\tau f \quad \text{for all } \tau \in [0, T].$$

An important class of systems is provided by stochastic differential equations with Lipschitz coefficients driven by a Wiener process [34, Theorem 13.1]. These have path-wise unique strong solutions. Strong solutions induce maps between corresponding path spaces [34, page 127], [19, pages 126-128]. Also,

<sup>2</sup>“continu à droite, limite à gauche” in French, alternatively RCLL (“right continuous with left limits”) in English.



under fairly general conditions (see e.g., [33, Chapter V]), stochastic differential equations driven by martingales with sample paths in  $D$  have strong solutions who are semi-martingales.

Besides stochastic differential equations in general, and those in (8) in particular, other nonlinear maps may serve as systems. For instance, discontinuous hystereses nonlinearities as well as non-Lipschitz static maps such as  $u \mapsto y := \sqrt{|u|}$ , are reasonable as systems, from an engineering viewpoint. Indeed, these induce maps from  $D \rightarrow D$  (or from  $C \rightarrow D$ , as in the case of relay hysteresis), are seen to be systems according to our definition,<sup>3</sup> and can be considered as components of nonlinear feedback laws. We note that a nonlinearity such as  $u \mapsto y = \text{sign}(u)$  is not a system in the sense of our definition since the output is not in general in  $D$ . Such nonlinearities, which often appear in bang-bang control, need to be approximated with a physically realizable hysteretic system.

#### IV. WELL-POSEDNESS AND A KEY LEMMA

It is straightforward to construct examples of deterministically well-posed feedback interconnections with elements as above. However, the situation is a bit more delicate when considering feedback loops since it is also perfectly possible that, at least mathematically, they give rise to unrealistic behavior. A standard example is that of a feedback loop with causal components that “implements” a perfect predictor. Indeed, consider a system  $f$  which superimposes its input with a delayed version of it, i.e.,

$$f : z(t) \mapsto z(t) + z(t - t_{\text{delay}}),$$

for  $t \geq 0$ , and assume initial conditions  $z(t) = 0$  for  $t < 0$ . Then the feedback interconnection of Figure 3 is unrealistic as it behaves as a perfect predictor. The feedback equation

$$z(t) = z_0(t) + f(z(t)) = z_0(t) + z(t) + z(t - t_{\text{delay}})$$

gives rise to  $0 = z_0(t) + z(t - t_{\text{delay}})$ , and hence,

$$z(t) = -z_0(t + t_{\text{delay}}).$$

Therefore, the output process  $z$  is not causally dependent on the input. The question of well-posedness of feedback systems has been studied from different angles for over forty years. See for instance the monograph by Jan Willems [39].

In our present setting of stochastic control we need a concept of well-posedness which ensures that signals inside a feedback loop are causally dependent on external inputs. This is a natural assumption from a systems point of view.

*Definition 7:* A feedback system is *deterministically well-posed* if the closed-loop maps are themselves systems; i.e., the feedback equation  $z = z_0 + f(z)$  has a unique solution  $z$  for inputs  $z_0$  and the operator  $(1 - f)^{-1}$  is itself a system.

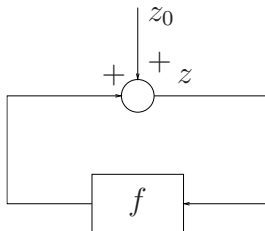


Fig. 3. Basic feedback system.

<sup>3</sup>More precisely, to be seen as a system, relay hysteresis needs to be preceded by a low-pass filter since its domain consists of continuous functions.

Thus, now thinking about  $z_0$  and  $z$  in the feedback system in Figure 3 as stochastic processes, deterministic well-posedness implies that  $\mathcal{Z}_t \subset \mathcal{Z}_t^0$  for  $t \in [0, T]$ , where  $\mathcal{Z}_t$  and  $\mathcal{Z}_t^0$  are the sigma-fields generated by  $z$  and  $z_0$ , respectively. This is a consequence of the fact that  $(1 - f)^{-1}$  is a system. Likewise, since  $(1 - f)$  is also a system,  $\mathcal{Z}_t^0 \subset \mathcal{Z}_t$  so that in fact

$$\mathcal{Z}_t^0 = \mathcal{Z}_t, \quad t \in [0, T]. \quad (32)$$

Next we consider the situation in Figure 1 and the relation between  $\mathcal{Y}_t$  and the filtration  $\mathcal{Y}_t^0$  of the process  $y_0 = Hz_0$ . The latter represents the ‘‘uncontrolled’’ output process where the control law  $\pi$  is taken to be identically zero. A key technical lemma for what follows states that the filtrations  $\mathcal{Y}_t$  and  $\mathcal{Y}_t^0$  are also identical if the feedback system is deterministically well-posed. This is not obvious at first sight, solely on the basis of the linear relationships  $y = Hz$  and  $y_0 = Hz_0$ , as the following simple example demonstrates: the two vector processes  $\begin{pmatrix} 0 \\ w \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ w \end{pmatrix}$  generate the same filtrations while  $(1 \ 0) \begin{pmatrix} 0 \\ w \end{pmatrix}$  and  $(1 \ 0) \begin{pmatrix} 0 \\ w \end{pmatrix}$  do not.

*Lemma 8:* If the feedback interconnection in Figure 1 is deterministically well-posed,  $g\pi$  is a system, and  $H$  is a linear system having a right inverse  $H^{-R}$  that is also a system, then  $(1 - Hg\pi)^{-1}$  is a system and  $\mathcal{Y}_t = \mathcal{Y}_t^0$ ,  $t \in [0, T]$ .

*Remark 9:* Note that, for the prototype problem involving (1), the conditions on  $H$  in Lemma 8 are trivial as  $H = [0 \ I]$  and hence  $H^{-R} := H'$  is a right inverse. The requirement in the lemma that  $g\pi$  is a system allows for a more general situation where  $\pi$  is not itself a system (e.g., generating outputs not in  $\mathcal{D}$ ), but where the cascade connection is still admissible.

*Proof:* By well-posedness  $(1 - g\pi H)^{-1}$  is a system. To show that  $(1 - Hg\pi)^{-1}$  exists and is a system, first note that

$$(1 - Hg\pi)H = H - Hg\pi H = H(1 - g\pi H). \quad (33)$$

The first step is using left distributivity and the second is using the fact that  $H$  is linear. But then

$$(1 - Hg\pi) \underbrace{H(1 - g\pi H)^{-1}H^{-R}}_h = I, \quad (34)$$

where  $HH^{-R} = I$ . Thus,  $h$  is a ‘‘right inverse’’ of  $p := (1 - Hg\pi)$  in that the composition  $p \circ h$  of the two maps is the identity. We claim that  $h$  is in fact the inverse of  $p$  (which is necessarily unique) in that  $y = h(y_0)$  and

$$(1 - Hg\pi)y = y_0 \quad (35)$$

establish a bijective correspondence between  $y$  and  $y_0$ , i.e., that both  $p \circ h$  as well as  $h \circ p$  are identity maps. We need to show the latter. The only potential problem would be if two distinct values  $y$  and  $\hat{y}$  satisfy (35) for the same value for  $y_0$ . We now show that this is not possible.

Since  $H$  is right invertible,  $y_0$  can be written in the form  $y_0 = Hz_0$  for  $z_0 = H^{-R}y_0$ . Let  $z = (1 - g\pi H)^{-1}z_0$  and  $y = Hz$ . Then  $y = h(y_0)$ , so by (34)  $y$  is a particular solution of equation (35). Now let  $\hat{y}$  be another solution, i.e., suppose that

$$(1 - Hg\pi)\hat{y} = y_0 \quad (36)$$

and that  $\hat{y} \neq y$ . We begin by writing  $\hat{y}$  in the form  $\hat{y} = H\hat{z}$ , which can always be done since  $H$  is right invertible. Next we set  $\hat{z}_0 := (1 - g\pi H)\hat{z}$ . Then, by well-posedness,  $\hat{z}$  is the unique solution of

$$\hat{z} = \hat{z}_0 + g\pi H(\hat{z}). \quad (37)$$

Moreover, by (33) and (36),  $H\hat{z}_0 = y_0$ , and consequently  $\hat{z}_0 = z_0 + v$  with  $Hv = 0$ . We now claim that  $\hat{z} = z + v$  which would then contradict the assumption that  $\hat{y} \neq y$ . To show this, note that, since  $z = z_0 + g\pi Hz$  and  $H$  is linear,

$$z + v = z_0 + v + g\pi H(z + v).$$

But the solution to (37) is unique by well-posedness. Hence,  $\hat{z} = z + v$  which proves our claim.

Therefore, finally,  $(1 - Hg\pi)$  is invertible and

$$(1 - Hg\pi)^{-1} = h = H(1 - g\pi H)^{-1}H^{-R}$$

is itself is a system, being a composition of systems. Thus, the configuration in Fig. 4 is deterministically well-posed. Using (33) once again,

$$H(1 - g\pi H)^{-1} = (1 - Hg\pi)^{-1}H. \tag{38}$$

It now follows that

$$y = H(1 - g\pi H)^{-1}z_0 = (1 - Hg\pi)^{-1}Hz_0 = (1 - Hg\pi)^{-1}y_0, \tag{39}$$

while also (35) holds. Equation (39) shows that  $\mathcal{Y}_t \subset \mathcal{Y}_t^0$ , whereas (35) shows that  $\mathcal{Y}_t^0 \subset \mathcal{Y}_t$ . ■

The essence of the lemma<sup>4</sup> is to underscore the equivalence between the configuration in Figure 1 and that in Figure 4. It is this equivalence which accounts for the identity  $\mathcal{Y}_t = \mathcal{Y}_t^0$  between the respective  $\sigma$ -algebras. An analogous notion of well-posedness was considered by Willems in [40] where however, in contrast, the well-posedness of the feedback configuration in Figure 4, and consequently the validity of  $\mathcal{Y}_t = \mathcal{Y}_t^0$ , is assumed at the outset.

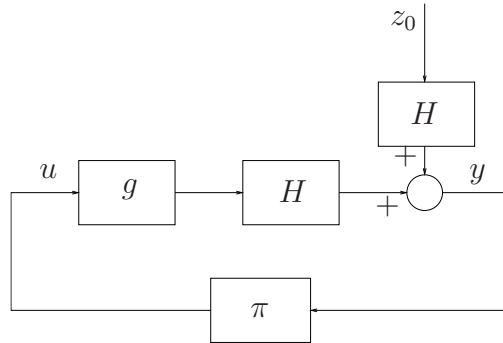


Fig. 4. An equivalent feedback configuration.

In the present paper we consider only feedback laws that render the feedback system deterministically well-posed. Therefore we highlight the conditions in a formal definition.

*Definition 10:* A feedback law  $\pi$  is *deterministically well-posed* for the system (8) if  $g\pi$  is a system and the feedback loop of Figure 1 is deterministically well-posed.

If the feedback law  $\pi$  is deterministically well-posed, then, by Lemma 8, the feedback loop in Figure 4 is also deterministically well-posed. Thus, in essence, given the assumption that  $z = z_0 + g\pi Hz$  admits a pathwise unique strong solution, so does  $y = y_0 + Hg\pi y$ .

*Remark 11:* For pedagogical reasons, we consider the case of complete state information, corresponding to (14). This corresponds to taking  $H = I$  and  $z = x$ , and the basic feedback loop is as depicted in Figure 5. Then the basic condition for well-posedness (32) states that the filtration  $\{\mathcal{X}_t\}$ , where  $\mathcal{X}_t := \sigma\{x(s); s \in [0, T]\}$ , is constant under variations of the control. Consequently, we do not need Lemma 8 to resolve an issue of circular control dependence. This is completely consistent with the analysis leading up to (15) in Section II.

<sup>4</sup>It is interesting to note, as was pointed out by a referee, that the proof of the lemma relies critically on the action of the operator  $(1 - g\pi H)^{-1}$  on a null set, as the probability  $\mathbb{P}(z_0 = H^{-R}y_0) = 0$  for any nontrivial model. This fact may be disturbing from a probabilistic point of view but does not invalidate the lemma.

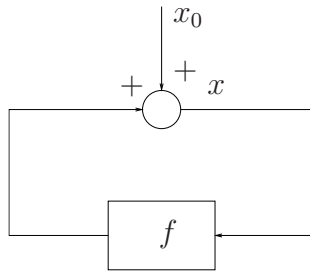


Fig. 5. Feedback loop for complete state information.

*Remark 12:* We now present an example of a feedback system which fails to be deterministically well-posed<sup>5</sup>. Consider the system

$$\begin{cases} dx = udt + dw \\ y = x \end{cases}$$

where  $w$  is a Wiener process. Then, the control law  $u = \pi(y)$  with  $\pi(y) = \max\{|x|^{2/3}, 1\}$  is not deterministically well-posed although the stochastic differential equation

$$dx = \pi(x)dt + dw$$

has a unique strong solution [18, Chapter 5, Proposition 5.17] in the sense that any other solution has same sample paths with probability one (indistinguishable). The failure to be deterministically well-posed can be traced to the fact that this control law allows for multiple consistent responses for  $w \equiv 0$ , a physically questionable situation. Indeed, the ordinary differential equation  $\dot{x} = \pi(x)$  is not Lipschitz and has infinitely many solutions.

## V. THE SEPARATION PRINCIPLE

Our first result is a very general separation theorem for the classical stochastic control problem stated at the beginning of Section I.

*Theorem 13:* Given the system (1), consider the problem of minimizing the functional (3) over the class of all feedback laws  $\pi$  that are deterministically well-posed for (1). Then the unique optimal control law is given by (4), where  $K$  is defined by (11), and  $\hat{x}$  is given by the Kalman filter (5).

*Proof:* By Lemma 8, (18) does not depend on the control. Therefore, given the analysis at the beginning of Section II, (4) is the unique optimal control provided it defines a deterministically well-posed control law. It remains to show this.

Inserting (4) into (5) yields

$$\hat{x}(t) = \int_0^t \Psi(t, s)L(s)dy(s),$$

where the transition matrix  $\Psi(t, s)$  of  $[A(t) + B_1(t)K(t) - L(t)C(t)]$  has partial derivatives in both arguments. Together with (4) this yields

$$u(t) = (\pi_{\text{opt}}y)(t) := \int_0^t M(t, s)dy(s), \quad (40)$$

where  $M(t, s) := K(t)\Psi(t, s)L(s)$ . Clearly  $s \mapsto M(t, s)$  has bounded variation for each  $t \in [0, T]$ , and therefore integration by parts yields

$$(\pi_{\text{opt}}y)(t) = M(t, t)y(t) - \int_0^t d_s M(t, s)y(s)ds, \quad (41)$$

<sup>5</sup>This example was kindly suggested by a referee.

which is defined samplewise. Now inserting  $u = \pi_{\text{opt}}Hz$  into (9) and (10) we obtain

$$z = z_0 + g\pi_{\text{opt}}Hz, \quad (42)$$

where  $g\pi_{\text{opt}}Hz$  takes the form

$$(g\pi_{\text{opt}}Hz)(t) = \int_0^t N(t, s)dz(s) \quad \text{with} \quad N(t, s) = \int_s^t G(t, \tau)M(\tau, s)Hd\tau,$$

where  $G$  is the kernel of the Volterra operator (9). A simple calculation yields

$$\frac{\partial G}{\partial s}(t, s) = \begin{bmatrix} A(t) \\ C(t) \end{bmatrix} \Phi(t, s)B_1(s),$$

where  $\Phi(t, s)$  is the transition matrix of  $A$ , and therefore  $Q(t, s) := \frac{\partial N}{\partial s}(t, s)$  is a continuous Volterra kernel, and so is the unique solution  $R$  of the resolvent equation

$$R(t, s) = \int_s^t R(t, \tau)Q(\tau, s)d\tau + Q(t, s) \quad (43)$$

[35], [42]. From (42) we have

$$dz = dz_0 + \int_0^t Q(t, s)dz(s)dt$$

from which it follows that

$$\int_0^t Q(t, s)dz(s) = \int_0^t R(t, s)dz_0(s).$$

Consequently,  $(1 - g\pi_{\text{opt}}H)$  has a unique preimage given by

$$[(1 - g\pi_{\text{opt}}H)^{-1}z](t) = z_0(t) + \int_0^t \int_\tau^t R(t, s)dsdz_0(\tau),$$

which is clearly a system, as claimed. Hence the feedback loop is deterministically well-posed.  $\blacksquare$

Consequently, for a system driven by a Wiener process with Gaussian initial condition, the linear control law defined by (4) and (5) is optimal in the class of all linear and nonlinear control laws for which the feedback system is deterministically well-posed.

If we forsake the requirement that  $\hat{x}$  is given by the Kalman filter (5), we can now allow  $x_0$  to be non-Gaussian and  $w$  to be an arbitrary martingale, even allowing jumps.

*Theorem 14:* Given the system (1), where  $w$  is a martingale and  $x(0)$  is an arbitrary zero mean random vector independent of  $w$ , consider the problem of minimizing the functional (3) over the class of all feedback laws  $\pi$  that are deterministically well-posed for (1). Then, provided it is deterministically well-posed, the unique optimal control law is given by (4), where  $K$  is defined by (11) and  $\hat{x}$  is the conditional mean (6).

*Proof:* Given Lemma 8, we can use the same completion-of-squares argument as in Section II except that we now need to use Ito's differential rule for martingales (see, e.g., [19], [33]), which, in integrated form, becomes

$$\begin{aligned} x(T)'P(T)x(T) - x(0)'P(0)x(0) &= f_\Delta + \\ &+ \int_0^T \{x(t)'\dot{P}(t)x(t)dt + 2x(t_-)'P(t)dx + \text{tr}(P(t)d[x, x'])\}, \end{aligned} \quad (44)$$

where  $[x, x']$  is the quadratic variation of  $x$  and  $f_\Delta$  is an extra term which is in general nontrivial when  $w$  has a jump component. Now let

$$q(t) := \int_0^t \Phi(t, s)(A(s)x(s) + B_1(s)u(s))ds,$$

where  $\Phi$  is the transition function of (1) which is differentiable in both arguments. Then,  $x = q + v$ , where  $dv = B_2 dw$  and  $q$  is a continuous process with bounded variation. Therefore

$$[x, x'] = [q, q'] + 2[q, v'] + [v, v'] = [v, v'].$$

In fact,  $[q, q'] = [q, v'] = 0$  [19, Corollary 8.5]. Since  $v$  does not depend on the control  $u$ , neither does the last term in the integral in (44). If  $w$  has a jump component, we have a nontrivial extra term in (44), namely

$$f_\Delta = \sum_{s \leq T} [x(s)'P(s)x(s) - x(s_-)'P(s)x(s_-) - 2x(s_-)'P(s)\Delta_s - \Delta_s'P(s)\Delta_s]$$

where the sum is over all jump times  $s$  on the interval  $[0, T]$  and  $\Delta_s := x(s) - x(s_-)$  is the jump, and we need to ensure that this term does not depend on the control either. However, since  $x(s) = x(s_-) + \Delta_s$ , we have  $f_\Delta = 0$ .

Then the rest of the proof that (4) with  $\hat{x}$  given by (6) is the unique minimizer of (3) over all deterministically well-posed control laws follows from an argument as in Section II. More precisely, using (11) and completing the squares we obtain

$$\begin{aligned} & \int_0^T x(t)'Q(t)x(t)dt + \int_0^T u(t)'R(t)u(t)dt + x(T)'Sx(T) \\ &= x(0)'P(0)x(0) + \int_0^T (u - Kx)'R(u - Kx)dt \\ &+ \int_0^T \text{tr}(P(t)d[v, v']) + \int_0^T x(t_-)'P(t)B_2(t)dw. \end{aligned} \quad (45)$$

Next we show that  $E \left\{ \int_0^T x(t_-)'P(t)B_2(t)dw \right\}$  does not depend on the control in this more general case as well. Since only the second term in (13) depends on the control, the problem reduces to showing that

$$E \left\{ \int_0^T \left[ \int_0^t \Phi(t, s)B_1(s)u(s)ds \right]' P(t)B_2(t)dw(t) \right\} = 0.$$

After a change in the order of integration this is equivalent to

$$E \left\{ \int_0^T u(t)' \int_t^T F(t, s)dw(s)dt \right\} = 0, \quad (46)$$

where  $F(t, s) = B_1(t)' \Phi(s, t)' P(s) B_2(s)$ . However, since  $u(t)$  is  $\mathcal{W}_t$ -measurable, where  $\mathcal{W}_t$  is the sigma-field generated by  $\{w(s); 0 \leq s \leq t\}$ , (46) can be written

$$E \left\{ \int_0^T u(t)' E \left\{ \int_t^T F(t, s)dw(s) \mid \mathcal{W}_t \right\} dt \right\},$$

which is zero since  $w$  is a martingale. In view of (17) where (18) does not depend on the control (Lemma 8) the statement of the theorem follows.  $\blacksquare$

We note that in general the optimal control law does not belong to  $\mathcal{L}$  and that  $\hat{x}$  is not given by the Kalman filter (5) but by the conditional mean (6), which then has to be chosen with some care since it is only defined almost surely as projection for each individual time  $t$ . To this end it is standard to select the optional projection of  $x(t)$  on  $\mathcal{Y}_t$  which is a stochastic process with a càdlàg version [2, page 17]. Often  $\hat{x}$  is given by a nonlinear filter as in the following example. However, even in those cases, it is difficult to ascertain well-posedness. At present, we are unable to establish that the control law in the example is deterministically well-posed and hence optimal in our admissible class of controls. We conjecture that Theorem 14 can be strengthened by removing the *a priori* assumption of well-posedness

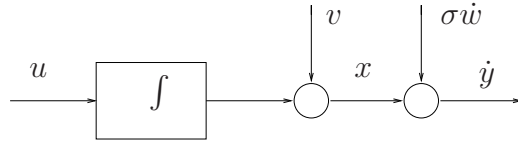


Fig. 6. Model for step change in white noise.

for the case where the optimal filter can be expressed as a stochastic differential equation with locally Lipschitz coefficients by suitable use of stopping times. Such a strengthening would suffice to prove optimality for the following example where we are currently unable to prove well-posedness.

*Example 15:* Consider the system in Figure 6. Here,  $x$  represents a parameter which undergoes a sudden random step change due to a random external forcing  $v$ . The step can be in either direction. Thus, as a stochastic process  $v(t)$  is defined

$$v(t) = \begin{cases} \theta & t \geq \tau \\ 0 & t < \tau \end{cases} \quad (47)$$

where  $\theta = \pm 1$  with equal probability and  $\tau$  is a random variable uniformly distributed on  $[0, T]$ . Clearly  $v$  is a martingale. Our goal is to maintain a value for the state  $x$  close to zero on the interval  $[0, T]$  via integral control action through  $u$ , indirectly, by demanding that

$$E \left\{ \int_0^T (x^2 + Ru^2) dt \right\}$$

be minimal with  $R > 0$ . Here,  $u$  denotes the control. The process  $x$  is observed in additive white noise  $\dot{w}$ . The system is now written in the standard form (1) as follows:

$$\begin{cases} dx = u(t)dt + dv, & x(0) = 0, \\ dy = x(t)dt + \sigma dw \end{cases} \quad (48)$$

where  $w$  is a Wiener process. We first solve the Riccati equation  $\dot{k} = -k^2 + R^{-1}$  with boundary condition  $k(T) = 0$  to obtain  $k(t) = -R^{-1/2} \tanh(R^{-1/2}(T-t))$ . The control law in Theorem 14 is

$$u(t) = k(t)\hat{x}(t), \quad (49a)$$

where the conditional expectation is determined separately using a (nonlinear) Wonham-Shiryayev filter

$$d\hat{x} = k(t)\hat{x}(t)dt + \frac{1}{\sigma^2}(1 - \rho(t)^2 - 2(T-t)\phi(t))(dy - \hat{x}(t)dt) \quad (49b)$$

$$d\rho = \frac{1}{\sigma^2}(1 - \rho(t)^2 - 2(T-t)\phi(t))(dy - \hat{x}(t)dt) \quad (49c)$$

$$d\phi = -\frac{1}{\sigma^2}\phi(t)\rho(t)(dy - \hat{x}(t)dt) \quad (49d)$$

with  $\rho(0) = 0$  and  $\phi(0) = 1$ . Following [16, page 222] we explain the steps for deriving the filter equations in Appendix VIII.

In order to conclude that the control law (49) is *actually* optimal we need to establish that the feedback loop is deterministically well-posed. This requires that (10) has a unique solution for each  $z_0 = (v \ w)'$ . Noting that the innovation  $dy - \hat{x}dt$  can be expressed as

$$dy - \hat{x}(t)dt = (v(t) - \rho(t))dt + dw,$$

this requires that the stochastic differential equations (48)-(49) can be uniquely solved path-wise as a map from  $z_0 = (v \ w)'$  to  $z = (x \ y)'$ . There are conditions in the literature for when path-wise uniqueness

holds (see [34, page 126, Theorem 10.4], [19, page 128], and the references therein). However, we are not able at present to verify that these hold in our case.

In view of Remark 11 we immediately have the following corollary to Theorem 14 for the case of complete state information. A similar statement was given in [27] in a different context.

*Corollary 16:* Given the system (14), where  $w$  is a martingale and  $x(0)$  is an arbitrary random vector independent of  $w$ , consider the problem of minimizing the functional (3) over the class of all feedback laws  $\pi$  that are deterministically well-posed for (14). Then the unique optimal control law is given by (15), where  $K$  is defined by (11).

*Proof:* It just remains to prove that the control law (15) is deterministically well-posed. To this end, we first note that (with  $z = x$ ) the feedback equation (10) becomes

$$x(t) = x_0(t) + \int_0^t Q(t, s)x(s)ds,$$

where  $Q(t, s) = \Phi(t, s)B_1(s)K(s)$  with  $\Phi$  (as before) being the transition matrix function of  $A$ . Then a straight-forward calculation shows that

$$x(t) = x_0(t) + \int_0^t R(t, s)x_0(s)ds,$$

where  $R$  is the unique solution of the resolvent equation (43). This establishes well-posedness.  $\blacksquare$

*Example 17:* Let the driving noise  $w$  in (14) be given by either a Poisson martingale [19, page 87], or a *geometric Brownian motion* [19, page 124]

$$dw = \mu w(t)dt + \sigma w(t)dv,$$

where  $v$  is a Wiener process, or a combination. Then the control law  $u(t) = K(t)x(t)$  is optimal for the problem to minimize (3).

## VI. THE SEPARATION PRINCIPLE FOR DELAY-DIFFERENTIAL SYSTEMS

The formulation (8) covers more general stochastic systems than the ones considered above. An example is a delay-differential system of the type

$$\begin{cases} dx = A_1(t)x(t)dt + A_2(t)x(t-h)dt + \int_{t-h}^t A_0(t, s)x(s)dsdt + B_1(t)u(t)dt + B_2(t)dw \\ dy = C_1(t)x(t)dt + C_2(t)x(t-h)dt + D(t)dw \end{cases}$$

Apparently, stochastic control for various versions of such systems were first studied in [23], [24], [26], [27], and [9], although [9] relies on the strong assumption that the observation  $y$  is “functionally independent” of the control  $u$ , thus avoiding the key question studied in the present paper.

Here, as in [26], we shall consider the wider class of stochastic systems

$$\begin{cases} dx = \left( \int_{t-h}^t d_s A(t, s)x(s) \right) dt + B_1(t)u(t)dt + B_2(t)dw \\ dy = \left( \int_{t-h}^t d_s C(t, s)x(s) \right) dt + D(t)dw \end{cases} \quad (50)$$

where  $A$  and  $C$  are of bounded variation in the first argument and continuous on the right in the second,  $x(t) = \xi(t)$  is deterministic (for simplicity) for  $-h \leq t \leq 0$ , and  $y(0) = 0$ . More precisely,  $A(t, s) = 0$  for  $s \geq t$ ,  $A(t, s) = A(t, t-h)$  for  $t \leq t-h$ , and the total variation of  $s \mapsto A(t, s)$  is bounded by an integrable function in the variable  $t$ , and the same holds for  $C$ . Moreover, to avoid technicalities we assume that  $w$  is now a (square-integrable) Gaussian (vector) martingale. Now, the first of equations (50) can be written in the form

$$\begin{aligned} x(t) = & \Phi(t, 0)\xi(0) + \int_{-h}^0 d_\tau \left\{ \int_0^t \Phi(t, s)A(s, \tau)ds \right\} \xi(\tau) \\ & + \int_0^t \Phi(t, s)B_1(s)u(s)ds + \int_0^t \Phi(t, s)B_2(s)dw \end{aligned} \quad (51)$$



[26, p. 85], where  $\Phi$  is the Green's function corresponding to the deterministic system [3] (also see, e.g., [26, p. 101]). In the same way, we can express the second equation in integrated form. Consequently, (50) can be written in the form (8), where  $K$  and  $H$  are computed as in [26, pp. 101–103]. The problem is to find a feedback law (2) that minimizes

$$J(u) := E\{V_0(x, u)\} \quad (52)$$

subject to the constraint (50), where

$$V_s(x, u) := \left\{ \int_s^T x(t)'Q(t)x(t)d\alpha(t) + \int_s^T u(t)'R(t)u(t)dt \right\} \quad (53)$$

and  $d\alpha$  is a positive Stieltjes measure.

Lemma 8 enables us to strengthen the results in [26]. To this end, to avoid technicalities, we shall appeal to a representation result from [27] rather than using a completion-of-squares argument, although the latter strategy would lead to a stronger result where  $w$  could be an arbitrary martingale. A completion-of-squares argument for a considerably simpler problem was given in [8], but, as pointed out in [28], this paper suffers from a similar mistake as the one pointed out earlier on page 4 in the present paper. In this context, we also mention the recent paper [4], which considers optimal control of a stochastic system with delay in the control. This paper assumes at the outset that the separation principle for delay systems is valid with a reference to [20]. Instead of basing the argument on [20], which is not quite appropriate here, their claim could be justified by noting that the delay in the control also implies a delay in information as in Example 3 above.

Now, it can be shown that the corresponding deterministic control problem obtained by setting  $w = 0$  has an optimal linear feedback control law

$$u(t) = \int_{t-h}^t d_\tau K(t, \tau)x(\tau), \quad (54)$$

where we refer the reader to [26] for the computation of  $K$ . The following theorem is a considerable strengthening of the corresponding result in [26].

*Theorem 18:* Given the system (50), where  $w$  is a Gaussian martingale, consider the problem of minimizing the functional (52) over the class of all feedback laws  $\pi$  that are deterministically well-posed for (1). Then the unique optimal control law is given by

$$u(t) = \int_{t-h}^t d_s K(t, s)\hat{x}(s|t), \quad (55)$$

where  $K$  is the deterministic control gain (54) and

$$\hat{x}(s|t) := E\{x(s) \mid \mathcal{Y}_t\} \quad (56)$$

is given by a linear (distributed) filter

$$d\hat{x}(t|t) = \int_{t-h}^t d_s A(t, s)\hat{x}(s|t)dt + B_1 u dt + X(t, t)dv \quad (57a)$$

$$d_t \hat{x}(s|t) = X(s, t)dv, \quad s \leq t \quad (57b)$$

where  $v$  is the innovation process

$$dv = dy - \int_{t-h}^t d_s C(t, s)\hat{x}(s|t)dt, \quad v(0) = 0, \quad (58)$$

and the gain  $X$  is as defined in [26, p.120].

For the proof of Theorem 18 we shall need two lemmas. The first is a slight reformulation of Lemma 4.1 in [27] and only requires that  $v$  be a martingale.

*Lemma 19* ([27]): Let  $v$  be a square-integrable martingale with natural filtration

$$\mathcal{V}_t = \sigma\{v(s), s \in [0, t]\}, \quad 0 \leq t \leq T \quad (59)$$

and satisfying  $[v_j, v_k] = \beta_j \delta_{jk}$ , where  $\beta_k$ ,  $k = 1, 2, \dots, p$ , are nondecreasing functions, and  $\delta_{jk}$  is the Kronecker delta equal to one for  $j = k$  and zero otherwise. With  $u$  a square-integrable control process adapted to  $\{\mathcal{V}_t\}$ , let

$$u(t) = \bar{u}(t) + \sum_{k=1}^p \int_0^t u_k(t, s) dv_k(s) + \tilde{u}(t) \quad (60)$$

be the unique orthogonal decomposition for which  $\bar{u}$  is deterministic and, for each  $t \in [0, T]$ ,  $\tilde{u}$  is orthogonal to the linear span of the components of  $\{v(s), s \in [0, t]\}$ . Moreover, let  $x_0$  be a square-integrable process adapted to  $\{\mathcal{V}_t\}$  and having a corresponding orthogonal decomposition

$$x_0(t) = \bar{x}_0(t) + \sum_{k=1}^p \int_0^t x_k^0(t, s) dv_k(s) + \tilde{x}_0(t). \quad (61)$$

Then  $x = x_0 + g(u)$ , defined by (8) exchanging  $z$  for  $x$ , has the orthogonal decomposition

$$x(t) = \bar{x}(t) + \sum_{k=1}^p \int_0^t x_k(t, s) dv_k(s) + \tilde{x}(t), \quad (62)$$

where

$$\bar{x}(t) = \bar{x}_0(t) + \int_0^t G(t, \tau) \bar{u}(\tau) d\tau \quad (63a)$$

$$x_k(t, s) = x_k^0(t, s) + \int_s^t G(t, \tau) u_k(\tau, s) d\tau \quad (63b)$$

$$\tilde{x}(t) = \tilde{x}_0(t) + \int_0^t G(t, \tau) \tilde{u}(\tau) d\tau \quad (63c)$$

and

$$E\{V_0(x, u)\} = E\{V_0(\bar{x}, \bar{u}) + \sum_{k=1}^p \int_0^T V_s(x_k(\cdot, s), u_k(\cdot, s)) d\beta_k + E\{V_0(\tilde{x}, \tilde{u})\}. \quad (64)$$

For a proof of this lemma, we refer the reader to [27].

*Lemma 20*: Let  $y$  be the output process of the closed-loop system obtained after applying a deterministically well-posed feedback law  $u = \pi(y)$  to the system (50). Then the innovation process (58) is a Gaussian martingale, and the corresponding filtration (59) satisfies

$$\mathcal{V}_t = \mathcal{Y}_t, \quad 0 \leq t \leq T. \quad (65)$$

*Proof*: As can be seen from the equation (51) and the remark following it, the process  $y_0$  obtained by setting  $u = 0$  in (50) is given by  $dy_0 = q_0(t)dt + D(t)dw$  for a process  $q_0$  adapted to  $\{\mathcal{W}_t\}$ . Define  $dv_0 = dy_0 - \hat{q}_0(t)dt$ , where  $\hat{q}_0(t) := E\{q_0(t) \mid \mathcal{Y}_t^0\}$ . Now,  $q_0$  and  $w$  are jointly Gaussian, and therefore, for each  $t \in [0, T]$ , the components of  $\hat{q}_0(t)$  belong to the closed linear span of the components of the martingale  $\{y_0, t \in [0, T]\}$ , and hence

$$\hat{q}_0(t) = \int_0^t M(t, s) dy_0$$

for some  $L^2$ -kernel  $M$ . Therefore,  $v_0$  is Gaussian, and its natural filtration  $\mathcal{V}_t^0$  satisfies  $\mathcal{V}_t^0 \subset \mathcal{Y}_t^0$ . Now let  $R$  be the resolvent of the Volterra equation with kernel  $M$ ; i.e., the unique solution of the resolvent equation

$$R(t, s) = \int_s^t R(t, \tau) M(\tau, s) d\tau + M(t, s)$$

[35], [42]. Then

$$\int_s^t R(t, s) dv_0(s) = \int_s^t M(t, s) dy_0(s) = \hat{q}_0(t),$$

and hence  $\mathcal{Y}_t^0 \subset \mathcal{V}_t^0$ . Consequently, in view of Lemma 8,  $\mathcal{V}_t^0 = \mathcal{Y}_t^0 = \mathcal{Y}_t$ . Next observe that

$$dy = q(t) + D(t)dw, \quad q(t) := q_0(t) + h(u)(t),$$

where  $h(u)$  is a causal (linear) function of the control  $u$ . Since  $h(u)$  is adapted to  $\{\mathcal{Y}_t\}$ ,

$$\hat{q}(t) := \hat{q}_0(t) + h(u)(t),$$

and therefore the innovation process (58) satisfies  $dv = dy - \hat{q}(t)dt = dy_0 - \hat{q}_0(t)dt = dv_0$ . Equation (65) now follows.

Finally, to prove that the innovation process  $v$  is a martingale we need to show that

$$E\{v(s) - v(t) \mid \mathcal{V}_t\} = 0 \text{ for all } s \geq t.$$

To this end, first note that

$$E\{v(s) - v(t) \mid \mathcal{V}_t\} = E\left\{\int_t^s \tilde{q}(\tau) d\tau \mid \mathcal{V}_t\right\} + E\left\{\int_t^s B(\tau) dw \mid \mathcal{V}_t\right\}, \quad (66)$$

where  $\tilde{q}(t) := q(t) - \hat{q}(t)$ . Since all the processes are jointly Gaussian (the control-dependent terms have been canceled in forming  $\tilde{q}$ ), independence is the same as orthogonality. Since  $\tilde{q}(\tau) \perp \mathcal{V}_\tau \supset \mathcal{V}_t$  for  $\tau \geq t$ , the first term in (66) is zero. The second term can be written

$$E\left\{E\left\{\int_t^s B(\tau) dw \mid \mathcal{W}_t\right\} \mid \mathcal{V}_t\right\},$$

which is zero since  $w$  is a martingale. ■

We are now in a position to prove Theorem 18. Lemma 20 shows that the innovation process (58) is a martingale. It is no restriction to assume that  $E\{v(t)v(t)'\}$  is diagonal; if it is not, we just normalize the innovation process by replacing  $v(t)$  by  $R(t)^{-1/2}v(t)$ , where  $R(t) := E\{v(t)v(t)'\} > 0$ . Then we set  $\beta_k(t) := E\{v_k^2\}$ ,  $k = 1, 2, \dots, p$ . Since  $\mathcal{V}_t = \mathcal{Y}_t$  for  $t \in [0, T]$  (Lemma 20), admissible controls take the form (60). Moreover, the process  $\hat{x}(t) := E\{x(t) \mid \mathcal{Y}_t\}$  is adapted to  $\{\mathcal{V}_t\}$ , and hence, analogously to (60), it has the decomposition

$$\hat{x}(t) = \bar{x}(t) + \sum_{k=1}^p \int_0^t x_k(t, s) dv_k(s) + \tilde{x}(t), \quad (67)$$

which now will take the place of (62) in Lemma 19. As before, let  $\hat{x}_0$  be the process  $\hat{x}$  obtained by setting  $u = 0$ . By Lemma 8,  $\hat{x}_0$  does not depend on the control  $u$ . Moreover, since  $x_0$  and  $v$  are jointly Gaussian,

$$\hat{x}_0(t) = \bar{x}_0(t) + \sum_{k=1}^p \int_0^t x_k^0(t, s) dv_k(s), \quad (68)$$

replacing (61) in Lemma 19. Moreover,

$$E\{V_0(x, u)\} = E\{V_0(\hat{x}, u)\} + E\{V_0(x - \hat{x}, 0)\},$$

where the last term does not depend on the control, since  $x - \hat{x} = x_0 - \hat{x}_0$ . Hence, by Lemma 19, the problem is now reduced to finding a control (60) and a state process (67) minimizing  $E\{V_0(\hat{x}, u)\}$  subject to

$$\bar{x}(t) = \bar{x}_0(t) + \int_0^t G(t, \tau) \bar{u}(\tau) d\tau \quad (69a)$$

$$x_k(t, s) = x_k^0(t, s) + \int_s^t G(t, \tau) u_k(\tau, s) d\tau \quad (69b)$$

$$\tilde{x}(t) = \int_0^t G(t, \tau) \tilde{u}(\tau) d\tau \quad (69c)$$

where the last equation has been modified to account for the fact that  $\tilde{x}_0 = 0$ . Clearly, this problem decomposes into several distinct problems. First  $\bar{u}$  need to chosen so that  $V_0(\bar{x}, \bar{u})$  is minimized subject to (69a). This is a deterministic control problem with the feedback solution

$$\bar{u}(t) = \int_{t-h}^t d_\tau K(t, \tau) \bar{x}(\tau), \quad (70)$$

where  $K$  is as in (54). Secondly, for each  $s \in [0, T]$  and  $k = 1, 2, \dots, p$ ,  $u_k(t, s)$  has to be chosen so as to minimize  $V_s(x_k(\cdot, s), u_k(\cdot, s))$  subject to (69b). This again is a deterministic control problem with the optimal feedback solution

$$u_k(t, s) = \int_{t-h}^t d_\tau K(t, \tau) x_k(\tau, s). \quad (71)$$

Finally,  $\tilde{u}$  should be chosen so as to minimize  $E\{V_0(\tilde{x}, \tilde{u})\}$  subject to (69c). This problem clearly has the solution  $\tilde{u} = 0$ , and hence  $\tilde{x} = 0$  as well. Combining these results inserting them into (60) then yields the optimal feedback control

$$u(t) = \int_{t-h}^t d_\tau K(t, \tau) (\bar{x}(\tau) + \sum_{k=1}^p \int_0^t x_k(t, s) dv_k(s))$$

It remains to show that this is exactly the same as (55); i.e., that

$$\hat{x}(\tau|t) = \bar{x}(\tau) + \sum_{k=1}^p \int_0^t x_k(t, s) dv_k(s). \quad (72)$$

To this end, first note that, since the optimal control is linear in  $dv$ ,  $\hat{x}(\tau|t)$  will take the form

$$\hat{x}(\tau|t) = \bar{x}(\tau) + \int_0^t X_t(\tau, s) dv(s),$$

where  $\bar{x}(\tau) = E\{x(\tau)\}$ , the same as in (72). Clearly  $E\{[x(\tau) - \hat{x}(\tau|t)]v(s)'\} = 0$  for  $s \in [0, t]$ , and therefore

$$E\{x(\tau)v(s)'\} = E\{\hat{x}(\tau|t)v(s)'\} = \int_0^s X_t(\tau, s) d\beta(s),$$

showing that the kernel  $X_t$  does not depend on  $t$ ; hence this index will be dropped. Now, setting  $\tau = t$ , comparing with (67) and noting that  $\tilde{x} = 0$ , we see that  $X(t, s)$  is the matrix with columns  $x_1(t, s), x_2(t, s), \dots, x_p(t, s)$ , establishing (72), which from now we shall write

$$\hat{x}(\tau|t) = \bar{x}(\tau) + \int_0^t X(\tau, s) dv(s). \quad (73)$$

Hence, (55) is the optimal control, as claimed. Moreover,

$$\hat{x}(\tau|t) = \hat{x}(s) + \int_s^t X(\tau, s) dv(s),$$

which yields (57a). To derive (57b), follow the procedure in [26].

It remains to show that the optimal control law (55) is deterministically well-posed. To this end, it is no restriction to assume that  $\bar{x}_0 \equiv 0$  so that all processes have zero mean. Then it follows from (55) and the unsymmetric Fubini Theorem of Cameron and Martin [10] that

$$u(t) = \int_0^t P(t, s) dv(s), \quad \text{where } P(t, s) = \int_{t-h}^t d_\tau K(t, \tau) X(\tau, s) d\tau,$$

and likewise from (58) that

$$dv = dy - \int_0^t S(t, s)dv(s)dt, \quad \text{where } S(t, s) = \int_{t-h}^t d_\tau C(t, \tau)X(\tau, s)d\tau.$$

The function  $S$  is a Volterra kernel and therefore the Volterra resolvent equation

$$V(t, s) = \int_s^t V(t, \tau)S(\tau, s)d\tau + S(t, s)$$

has a unique solution  $V$ , from which it follows that

$$dv = dy - \int_0^t V(t, s)dy(s).$$

Then the optimal control law is given by (40), where now  $M$  is given by

$$M(t, s) = P(t, s) - \int_s^t P(t, \tau)V(\tau, s)d\tau.$$

Now, for the optimal control law,  $s \mapsto X(t, s)$  is of bounded variation for each  $t$  [26], and hence so is  $s \mapsto M(t, s)$ . Hence  $\pi_{\text{opt}}$  can be defined samplewise as in (41). To complete the proof that the optimal feedback loop is deterministically well-posed we proceed exactly as in the proof of Theorem 13, noting that in the present setting

$$\frac{\partial G}{\partial s}(t, s) = \int_s^t d_\tau \begin{bmatrix} A(t, \tau) \\ C(t, \tau) \end{bmatrix} \Phi(\tau, s)B_1(s),$$

where  $\Phi(t, s)$  is the transition matrix of  $A$  [26, p.101].

*Remark 21:* It was shown in [27] that, in the case of complete state information ( $y = x$ ), the control (54) is optimal even when  $w$  is an arbitrary (not necessarily Gaussian) martingale.

## VII. CONCLUSIONS

In studying the literature on the separation principle of stochastic control, one encounters many expositions where subtle difficulties are overlooked and inadmissible shortcuts are taken. On the other hand, for most papers and monographs that provide rigorous derivations, one is struck by the level of mathematical sophistication and technical complexity, which makes the material hard to include in standard textbooks in a self-contained fashion. It is our hope that our use of deterministic well-posedness provides an alternative mechanism for understanding the separation principle that is more palatable and transparent to the engineering community, while still rigorous. The new insights offered by the approach allow us to establish the separation principle also for systems driven by non-Gaussian martingale noise. However, in this more general framework the key issue of establishing well-posedness for particular control systems is challenging and more work needs to be done.

## ACKNOWLEDGEMENT

We are indebted to an anonymous referee for significant input which has helped us improved the paper considerably.

## VIII. APPENDIX

Consider the “uncontrolled” observation process  $dy_0 = v(t)dt + \sigma dw$ . If  $d\mathbb{P}$  denotes the law of  $(\theta, \tau, w)$  and  $\Lambda(t) = e^{\sigma^{-2} \int_0^t v(s)dy_0 - \frac{1}{2}\sigma^{-2} \int_0^t v(s)^2 ds}$ , then, under a new measure  $d\mathbb{Q} := \Lambda(T)^{-1}d\mathbb{P}$ ,  $y_0$  becomes a Wiener process while the law of  $v$  (i.e., of  $\theta$  and  $\tau$ ) is the same as before. Under  $d\mathbb{Q}$ , the two processes  $y_0$  and  $v$  are independent. The conditional expectation is now given by (Bayes’ formula [16, page 174])

$$\begin{aligned} E_{\mathbb{P}}(v(t)|\mathcal{Y}_t) &= \frac{E_{\mathbb{Q}}(v(t)\Lambda(t)|\mathcal{Y}_t)}{E_{\mathbb{Q}}(\Lambda(t)|\mathcal{Y}_t)} = \frac{E_{\mathbb{Q}}(\theta I_{t \geq \tau} e^{\sigma^{-2} \int_0^t \theta I_{s \geq \tau} dy_0 - \frac{1}{2}\sigma^{-2} \int_0^t I_{s \geq \tau} ds} | \mathcal{Y}_t)}{E_{\mathbb{Q}}(e^{\sigma^{-2} \int_0^t \theta I_{s \geq \tau} dy_0 - \frac{1}{2}\sigma^{-2} \int_0^t I_{s \geq \tau} ds} | \mathcal{Y}_t)} \\ &= \frac{E_{\mathbb{Q}}(I_{t \geq \tau} e^{(y_0(t) - y_0(t \wedge \tau) - \frac{1}{2}(t - \tau)^+)/\sigma^2} - I_{t \geq \tau} e^{-(y_0(t) - y_0(t \wedge \tau) - \frac{1}{2}(t - \tau)^+)/\sigma^2} | \mathcal{Y}_t)}{E_{\mathbb{Q}}(e^{(y_0(t) - y_0(t \wedge \tau) - \frac{1}{2}(t - \tau)^+)/\sigma^2} - e^{-(y_0(t) - y_0(t \wedge \tau) - \frac{1}{2}(t - \tau)^+)/\sigma^2} | \mathcal{Y}_t)}. \end{aligned} \quad (74)$$

Here  $t \wedge \tau := \min(t, \tau)$ ,  $I_{t \geq \tau}(t) = 1$  when  $t \geq \tau$  and 0 otherwise, and  $(t - \tau)^+ = (t - \tau)I_{t \geq \tau}$ . Note that  $v(t) = \theta I_{t \geq \tau}(t)$ . For convenience we define  $\rho(t) := E_{\mathbb{P}}(v(t)|\mathcal{Y}_t)$  and

$$\begin{aligned} \Sigma(t) &:= \int_0^t (e^{y_0(t) - y_0(s) - \frac{1}{2}(t-s)/\sigma^2}) ds, \text{ and} \\ \bar{\Sigma}(t) &:= \int_0^t e^{-(y_0(t) - y_0(s) - \frac{1}{2}(t-s)/\sigma^2)} ds. \end{aligned}$$

From (74),  $\rho(t) = N(t)/D(t)$  where

$$\begin{aligned} N(t) &= \Sigma(t) - \bar{\Sigma}(t) \\ D(t) &= \Sigma(t) + \bar{\Sigma}(t) + 2(T - t). \end{aligned}$$

By first noting that  $\Sigma$  and  $\bar{\Sigma}$  satisfy the stochastic differential equations

$$\begin{aligned} d\Sigma &= \Sigma(t)dy_0 + dt \\ d\bar{\Sigma} &= -\bar{\Sigma}(t)dy_0 + dt, \end{aligned}$$

respectively, the Itô rule applied to the expression  $N(t)/D(t)$  for the conditional expectation gives the filter equations (setting  $\phi = D^{-1}$ )

$$d\rho = \sigma^{-2}(1 - \rho(t)^2 - 2(T - t)\phi(t))(dy_0 - \rho dt) \quad (75a)$$

$$d\phi = -\sigma^{-2}\phi(t)\rho(t)(dy_0(t) - \rho(t)dt). \quad (75b)$$

Finally, noting that the innovation  $dy_0 - \rho dt$  is equal to  $dy - \hat{x}dt$  for the controlled system, we obtain the filter equations (49).

## REFERENCES

- [1] K.J. Åström, *Introduction to Stochastic Control Theory*, Academic Press, 1970.
- [2] A. Bain and D. Crisan, *Fundamentals of Stochastic Filtering*, Springer, 2009.
- [3] H. T. Banks, “Representations for solutions of linear functional differential equations”, *J. Diff. Equ* **5** (1969), 399.
- [4] M. V. Basin, J.G. Rodriguez Gonzalez, R.M. Zuniga, “Optimal controller for linear systems with time delays in input and observations,” *Dynamics of Continuous, Discrete and Impulsive Systems, Series B: Appl. & Algorithms* **12** (2005), 1-14.
- [5] V. E. Beneš, “Nonexistence of strong nonanticipating solutions to stochastic DEs: implications for functional DEs, filtering, and control,” *Stoch. Proc. Appl.*, **5** (3): 243263, 1977.
- [6] A. Bensoussan, *Stochastic Control of Partially Observable Systems*, Cambridge University Press, 1992.
- [7] P. Billingsley, *Convergence of Probability Measures*, second edition, Wiley, 1999.
- [8] R. E. Schmootzer and G. L. Blankenship, “A simple proof of the separation theorem for linear stochastic systems with time delays,” *IEEE Trans. Automatic Control* **AC 23**, (1978), pp. 734-735.
- [9] R. Brooks, “Linear Stochastic Control: An extended separation principle,” *J. Math. Anal. Appl.* **38** (1972), 569 - 587.
- [10] R. H. Cameron and W. T. Martin, An unsymmetric Fubini theorem, *Bull. Amer. Math. Soc.* **47** (1941), 121.

- [11] J.M.C. Clark and D. Crisan, "On a robust version of the integral representation formula of nonlinear filtering," *Probab. Theory Relat. Fields*, **133**: 43-56, 2005.
- [12] M.H.A. Davis, *Linear Estimation and Stochastic Control*, Chapman and Hall, 1977.
- [13] M.H.A. Davis and P. Varaiya, "Information states for stochastic systems", *J. Math. Anal. Applications* **37** (1972), 384-402.
- [14] T. Duncan and P. Varaiya, "On the solutions of a stochastic control system," *SIAM J. Control* **9(3)** (1971), 354-371.
- [15] W.H. Fleming and R.W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, 1975.
- [16] R. van Handel, *Stochastic Calculus, Filtering, and Stochastic Control*, unpublished lecture notes, Caltech, Spring 2007.
- [17] P.D. Joseph and J.T. Tou, "On linear control theory", *AIEE Trans. Applications and Industry* **80** (1961), 193-196.
- [18] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, 1988.
- [19] F.C. Klebaner, *Introduction to Stochastic Calculus with Applications*, Imperial College Press, 2005.
- [20] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, John Wiley & Sons, 1972.
- [21] H. J. Kushner, *Stochastic Stability and Control*, Academic Press, 1967.
- [22] H. Kushner, *Introduction to Stochastic Control*, Holt, Rinehart and Winston, 1971.
- [23] A. Lindquist, "On optimal stochastic control with smoothed information," *Information Sciences* **1** (1968), 55-85.
- [24] A. Lindquist, "An innovations approach to optimal control of linear stochastic systems with time delay," *Information Sciences* **1** (1969), 279-295.
- [25] A. Lindquist, "A theorem on duality between estimation and control for linear stochastic systems with time delay," *Journal of Mathematical Analysis and Applications* **37** (1972), 516-536.
- [26] A. Lindquist, "Optimal control of linear stochastic systems with applications to time lag systems", *Information Sciences* **5** (1973), 81-126.
- [27] A. Lindquist, "On feedback control of linear stochastic systems," *SIAM J. Control* **11**: 323-343, 1973.
- [28] A. Lindquist, "Comments on 'A Simple proof of the separation theorem for linear stochastic systems with time delays,'" *IEEE Trans. Automatic Control* **AC 25**, (1980), pp. 274 - 275.
- [29] R. S. Liptser and A. N. Shirayev, *Statistics of Random Processes II, Applications*, Springer-Verlag, 1978.
- [30] S.K. Mitter, "Filtering and stochastic control: A historical perspective", *IEEE Control Systems Mag.* **16(3)** (1996), 67-76.
- [31] B. Øksendal, *Stochastic Differential Equations*, Springer, 2007.
- [32] J. E. Potter, "A guidance-navigation separation theorem", Rep. Re-II, Experim. Astronomy Lab., MIT, Cambridge, 1964.
- [33] P.E. Protter, *Stochastic Integration and Differential Equations*, Springer, 2004.
- [34] L.C.G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales, Volume 2, Itô Calculus*, Cambridge University Press, 2nd edition, 2000.
- [35] F. Smithies, *Integral Equations*, Cambridge University Press, 1958.
- [36] R.F. Stengel, *Optimal Control and Estimation*, unabridged and corrected, Dover (1994) republication of *Stochastic Optimal Control: Theory and Applications*, Wiley, 1986.
- [37] Tsirel'son, B.S., "An example of a stochastic differential equation which has no strong solution," *Teoriya Veroyatnostei i ee Primeneniya*, **20(2)**: 427-430, 1975.
- [38] H. van Warter and J. C. Willems, "The certainty equivalence property in stochastic control theory", *IEEE Trans. Automatic Control* **AC-26(5)** (1981), 1080-1087.
- [39] J. C. Willems, *The Analysis of feedback systems*, MIT Press, 1971.
- [40] J. C. Willems, "Recursive filtering," *Statistica Neerlandica*, **32(1)** (1978), 1-39.
- [41] W.M. Wonham, "On the separation theorem of stochastic control", *SIAM J. Control* **6(2)** (1968), 312-326.
- [42] P. P. Zabreyko et. al., *Integral Equations – A Reference Text*, Noordhoff International Publishing, Leyden, 1975.