

Parameter Identification for Partially Observed Diffusions

T. E. DABBOUS¹ AND N. U. AHMED²

Communicated by T. S. Angell

Abstract. In this paper, we consider the identification problem of drift and dispersion parameters for a class of partially observed systems governed by Ito equations. Using the pathwise description of the Zakai equation, we formulate the original identification problem as a deterministic control problem in which the unnormalized conditional density (solution of the Zakai equation) is treated as the state, the unknown parameters as controls, and the likelihood ratio as the objective functional. The question of existence of elements in the parameter set that maximize the likelihood ratio is discussed. Further, using variational arguments and the Gateaux differentiability of the unnormalized density on the parameter set, we obtain the necessary conditions for optimal identification.

Key Words. Nonlinear filtering, likelihood ratio, parameter identification, optimal control, distributed-parameter systems.

1. Introduction

In the last few years, considerable attention has been focused on the identification problem of systems governed by linear or nonlinear Ito equations (Refs. 1–6). In Ref. 2, the identification problem for partially observed linear time-invariant systems has been considered. Using linear filter theory, the maximum likelihood approach, and the smoothness of solutions of the algebraic Riccati equation, sufficient conditions were obtained for the consistency of the maximum likelihood estimate.

¹Associate Professor, Department of Electrical Engineering, Bilkent University, Ankara, Turkey.

²Professor, Department of Electrical Engineering, University of Ottawa, Ottawa, Ontario, Canada.

In Ref. 3, Lipster and Shirayev have considered the identification problem for a class of completely observed systems governed by a stochastic differential equation of the form

$$dx(t) = ah(t, x(t)) dt + dW(t), \quad t \geq 0,$$

where $x \in R$ and α is some unknown parameter. Using the maximum likelihood approach, an explicit expression for the maximum likelihood estimate $\hat{\alpha}_t$ was obtained. Further, utilizing the law of iterated logarithm of Brownian motion, it has been shown that, as $t \rightarrow \infty$, estimate $\hat{\alpha}_t$ converges almost surely to the true underlying parameter. In Ref. 4, Legland considered the identification problem for a more general class of systems governed by stochastic differential equation of the form

$$dy(t) = h(\alpha, x(t)) dt + dV(t), \quad t \geq 0,$$

where α is unknown and x is a diffusion process. Utilizing the maximum likelihood approach along with forward and backward Zakai equations, a numerical scheme has been developed for computing α given the output history $\{y(s) : s \leq t\}$.

In this paper, we consider the identification problem for a class of systems governed by Ito equations of the form

$$dx(t) = a(t, x(t), \alpha) dt + b(t, x(t), \alpha) dW(t), \quad t \in I \equiv [0, T], \quad (1a)$$

$$x(0) = x_0, \quad (1b)$$

and

$$dy(t) = h(x(t), \alpha) dt + \sigma_o(t, y(t)) d\tilde{W}(t), \quad t \in I, \quad (2a)$$

$$y(0) = 0, \quad (2b)$$

where W and \tilde{W} are two independent standard Wiener processes taking values from R^n and R^m , respectively, and α is an unknown parameter taking values from a compact convex set \mathcal{P} contained in some finite-dimensional space. The drift and diffusion coefficients are Borel functions as described below,

$$a : I \times R^n \times \mathcal{P} \rightarrow R^n, \quad h : R^n \times \mathcal{P} \rightarrow R^m,$$

$$b : I \times R^n \times \mathcal{P} \rightarrow R^{(n \times n)}, \quad \sigma_o : I \times R^m \rightarrow R^{(m \times m)}.$$

Further regularity properties of these functions will be presented in the sequel as required. We assume that all the random processes and vectors described above are defined on a complete probability space $(\Omega_o, \mathcal{B}_o, P_o)$. Then, loosely speaking, our problem is to identify the unknown parameter α on the basis of the output information $\{y(s) : s \geq 0\}$.

The paper is organized as follows. In Section 2, we present some of the notations that have been used in the sequel along with the necessary assumptions required to prove the existence result and to obtain the corresponding necessary conditions of optimality. In Section 3, we formulate the nonlinear filtering problem and present some of the well-known results. In Section 4, we formulate the identification problem as a deterministic control problem; then, following standard partial differential equation arguments, we show that the identification problem has a solution. Finally, in Section 5, we use standard variational arguments and make use of the Gateaux differentiability of the unnormalized density on the parameter set to obtain the necessary conditions for optimal identification.

2. Notations and Assumptions

Notations. Let $\eta(t)$, $t \geq 0$, be any random process, and let $\sigma\{\eta(s), s \leq t\}$ denote the σ -field generated by η up to time t . Define

$$\begin{aligned} \mathcal{F}_t^y &\equiv \sigma\{y(s), s \leq t\}, & \mathcal{F}_t^W &\equiv \sigma\{W(s), s \leq t\}, \\ \mathcal{F}_t &\equiv \mathcal{F}_t^y \vee \mathcal{F}_t^W \vee \sigma\{x_0\} \subset \mathcal{B}_o. \end{aligned}$$

Let $C(R^n)$ [resp. $C_b(R^n)$] denote the space of continuous [resp. bounded continuous] functions on R^n . Let $\mathcal{B}(R^n)$ denote the Borel field of subsets of R^n and $\mathcal{F} \equiv \bigcup_{t>0} \mathcal{F}_t$. Let Ω [resp. Ω_T] denote the space of continuous functions on $R_o \equiv [0, \infty)$ [resp. $[0, T]$] with values in $R^{(n+m)}$, and let \mathcal{A} [resp. \mathcal{A}_T] denote the Borel σ -algebra on Ω [resp. Ω_T]. We call (Ω, \mathcal{A}) the canonical sample space for the process $\{x(t), y(t)\}$, $t \geq 0$.

Let $L_2(I; R^n)$ denote the equivalence classes of measurable functions $f: I \rightarrow R^n$ such that $\int_I |f(t)|^2 dt < \infty$. For any Banach space E , we shall use $L_\infty(I; E)$ to denote the space of strongly measurable E -valued functions on I with the norm

$$\|f\|_\infty \equiv \text{ess sup}\{\|f(t)\|_E; t \in I\}.$$

Let $C(I; E)$ denote the space of strongly continuous E -valued functions on I furnished with the uniform topology

$$\|f\|_C \equiv \sup\{\|f(t)\|_E; t \in I\}.$$

For any pair of Banach spaces E and F , we use $\mathcal{L}(E, F)$ to denote the space of bounded linear operators from E to F . Let

$$H \equiv L_2(R^n), \quad V \equiv H^1 \equiv \{f \in H: \partial f / \partial x_i \in H, 1 \leq i \leq n\},$$

with V' being its dual. We use $\langle \cdot, \cdot \rangle$ to denote the pairing of V and V' . Further notations will be introduced in the sequel as required.

Assumptions

- (A1) $a(t, x, \alpha)$ is measurable in t and continuous in x and α . Further, for all $(t, \alpha) \in [0, T] \times \mathcal{P}$, $a(t, \cdot, \alpha)$ is bounded and satisfies uniform Lipschitz and growth conditions on R^n .
- (A2) The matrix function $b(t, x, \alpha)$ is measurable in t , continuous in x and α ; and for all $(t, \alpha) \in [0, T] \times \mathcal{P}$, $b(t, \cdot, \alpha)$ is bounded and satisfies uniform Lipschitz and growth condition on R^n . Further, there exists a constant $\gamma > 0$ such that

$$(bb')(t, x, \alpha) \equiv \sigma(t, x, \alpha) \geq \gamma I,$$

for all $(t, x, \alpha) \in [0, T] \times R^n \times \mathcal{P}$, where I denotes the identity matrix.

- (A3) For all $(t, \alpha) \in [0, T] \times \mathcal{P}$, the functions $(\partial/\partial x_j)\sigma_{ij}$ and $(\partial^2/\partial x_i \partial x_j)\sigma_{ij}$, $1 \leq i, j \leq n$, are bounded and satisfy a Holder condition on R^n .
- (A4) For every $(t, x) \in [0, T] \times R^n$, the mappings $\alpha \rightarrow \sigma(t, x, \alpha)$ and $\alpha \rightarrow a(t, x, \alpha)$ are once Gateaux differentiable on \mathcal{P} .
- (A5) For every $\alpha \in \mathcal{P}$, $h(\cdot, \alpha) \in C_b^2(R^n)$ and the map $\alpha \rightarrow h(\cdot, \alpha)$ is once Gateaux differentiable. Further, the mappings $\alpha \rightarrow h(x, \alpha)$ and $\alpha \rightarrow (\partial/\partial x_i)h(x, \alpha)$, $1 \leq i \leq n$, are continuous on \mathcal{P} for each $x \in R^n$.
- (A6) The matrix-valued functions $\sigma_o(t, y)$ is measurable in t , satisfies uniform Lipschitz and growth conditions on R^m , and
- (i) $(\sigma_o(t, y)\xi \cdot \xi) \geq \beta |\xi|^2$, $\beta > 0$, $\xi \in R^m$, $y \in R^m$,
- (ii) $E \int_T \text{tr}(\sigma_o(t, y)\sigma'_o(t, y)) dt < \infty$,

where the dot denotes the scalar product in R^m and $\text{tr}(B)$ denotes the trace of B .

Note that, under the given assumptions, the systems (1) and (2) have strong solutions for initial state x_0 with $E|x_0|^2 < \infty$ and $y(0) = 0$; see, for example, Ref. 7.

In the next section, we present some of the well-known results in nonlinear filtering theory (Refs. 8-9). These results are used to prove the existence of a solution for the identification problem and to derive the corresponding necessary conditions of optimality.

3. Nonlinear Filtering Problem

In this section, we formulate the filtering problem for the systems (1) and (2) and present the corresponding Kushner and Zakai equations (which

are parametrized by α). Let μ^1 and μ^2 be the measures induced on the canonical sample space (Ω, \mathcal{A}) by the system (1)–(2) and the system

$$dx(t) = a(t, x(t), \alpha) dt + b(t, x(t), \alpha) dW(t), \quad t \geq 0, \tag{3a}$$

$$dy(t) = a_o(t, y(t)) d\tilde{W}(t), \quad t \geq 0, \tag{3b}$$

respectively. For each $t \in [0, T]$, let μ_i^t , $i = 1, 2$, denote the restriction of the measure μ^i to \mathcal{A}_t . Then, under the given assumptions, the measures μ_1^t and μ_2^t are absolutely continuous with respect to one another. Further, the Radon–Nikodym derivative of μ_1^t with respect to μ_2^t is given by

$$d\mu_1^t/d\mu_2^t \equiv \rho_t^\alpha \equiv \exp \left\{ -1/2 \int_0^t |\sigma_o^{-1}(s, y(s))h(x(s), \alpha)|^2 ds + \int_0^t (\sigma_o^{-1}(s, y(s))h(x(s), \alpha) \cdot \sigma_o^{-1}(s, y(s)) dy(s)) \right\}, \tag{4}$$

for all $\alpha \in \mathcal{P}$, where the dot denotes the scalar product in any finite-dimensional space. It is known (Ref. 10) that if, for each $t \in [0, T]$,

$$E_2 \rho_t^\alpha \equiv \int_\Omega \rho_t^\alpha d\mu_2^t = 1,$$

then the process

$$\left\{ W(t), \int_0^t \sigma_o^{-1}(s, y(s)) dy(s), t \in [0, T] \right\}$$

is a standard Wiener process on the probability space $(\Omega, \mathcal{A}, \mu^2)$. For any bounded measurable function f on R^n , the optimal estimate (in the mean-square sense), relative to \mathcal{F}_t^y , is given by

$$\hat{f}(t) \equiv E_1 \{ f(x(t)) \mid \mathcal{F}_t^y \},$$

where E_1 denotes the expectation with respect to μ^1 . Using the fact that the measures μ^1 and μ^2 are absolutely continuous with respect to one another, it follows from (4) and the Bayes formula that

$$\hat{f}(t) = E_2 \{ \rho_t^\alpha f(x(t)) \mid \mathcal{F}_t^y \} / E_2 \{ \rho_t^\alpha \mid \mathcal{F}_t^y \}. \tag{5}$$

Clearly, $\hat{f}(t)$ depends on α . This dependence will be indicated by writing $\hat{f}^\alpha(t)$ instead of $\hat{f}(t)$. Let $\pi^\alpha(t) \equiv \pi^\alpha(t, \cdot)$ denote the conditional density of

$x(t)$ relative to \mathcal{F}_t^y , $t \geq 0$. It is known that $\pi^\alpha(t)$, $t \geq 0$, satisfies the following Kushner equation (Ref. 8):

$$d\pi^\alpha(t) = A^*(t, \alpha)\pi^\alpha(t) dt + (h^\alpha - \hat{h}^\alpha(t))\pi^\alpha(t)\Gamma_o^{-1}(t)[dy(t) - \hat{h}^\alpha(t) dt], \quad (6a)$$

$$\pi^\alpha(0) = p_0, \quad \alpha \in \mathcal{P}, \quad (6b)$$

for $t \geq 0$, where p_0 denotes the initial density, $\Gamma_o(t) \equiv (\sigma_0 \sigma_0')(t, y(t))$, and

$$A^*(t, \alpha)f \equiv - \sum_{i=1}^n (\partial/\partial x_i)(a_i(t, x, \alpha)f) + \sum_{i,j=1}^n (\partial^2/\partial x_i \partial x_j)(\sigma_{ij}(t, x, \alpha)f), \quad (7)$$

with

$$\sigma(t, x, \alpha) \equiv (bb')(t, x, \alpha).$$

Define $\varphi^\alpha(t) \equiv \varphi^\alpha(t, \cdot)$, $t \geq 0$, so that

$$\pi^\alpha(t, \cdot) = \varphi^\alpha(t, \cdot) / \int_{R^n} \varphi^\alpha(t, x) dx, \quad (8)$$

where φ^α is known as the unnormalized conditional density and satisfies the following Zakai equation (Ref. 9):

$$d\varphi^\alpha(t) = A^*(t, \alpha)\varphi^\alpha(t) dt + \Gamma_o^{-1}(t)h^\alpha\varphi^\alpha(t) \cdot dy(t), \quad t \geq 0, \quad (9a)$$

$$\varphi^\alpha(0) = p_0, \quad \alpha \in \mathcal{P}. \quad (9b)$$

Since

$$E_2\{\rho_t^\alpha f(x(t)) \mid \mathcal{F}_t^y\} = \int_{R^n} \varphi^\alpha(t, x)f(x) dx, \quad (10)$$

$$E_2\{\rho_t^\alpha \mid \mathcal{F}_t^y\} = \int_{R^n} \varphi^\alpha(t, x) dx, \quad (11)$$

it follows from (5) that

$$\begin{aligned} \hat{f}^\alpha(t) &= \int_{R^n} \varphi^\alpha(t, x)f(x) dx / \int_{R^n} \varphi^\alpha(t, x) dx \\ &\equiv \langle \varphi^\alpha(t), f \rangle / \langle \varphi^\alpha(t), 1 \rangle. \end{aligned} \quad (12)$$

Let $X \equiv C(I; R^n)$, $I \equiv [0, T]$, and let $\Sigma \in \mathcal{F}_T^y$. Then, it follows from (4) that

$$\int_{X \times \Sigma} d\mu_t^1 = \int_{X \times \Sigma} \rho_t^\alpha d\mu_t^2.$$

Since under the measure μ^2 , the process y is independent of x , it follows that μ^2 is given by the product of the two measures ν^x and $\tilde{\nu}$ which are defined on $\mathcal{B}(X)$ and \mathcal{F}_T^y , respectively. Then, it follows by Fubini's theorem that

$$\begin{aligned} \lambda_t^\alpha(\Sigma) &\equiv \mu_t^1(X \times \Sigma) = \int_{X \times \Sigma} \rho_t^\alpha \, d\mu_t^2 \\ &= \int_{\Sigma} \left(\int_X \rho_t^\alpha \, d\nu_t^x(x) \right) d\tilde{\nu}(y) = \int_{\Sigma} E_{\nu^x} \{ \rho_t^\alpha \mid y \} \, d\tilde{\nu}(y). \end{aligned}$$

Clearly, $\lambda_t^\alpha(\Sigma)$ defines a measure on \mathcal{F}_T^y . Let \tilde{y} be a realization of the process $\{y(s), s \geq 0\}$. Then, for the one point set $\Sigma \equiv y'_0 \equiv \{y(s), s \leq t\}$, one can verify that

$$\lambda_t^\alpha(\{y'_0\}) = E_{\nu^x} \{ \rho_t^\alpha \mid y'_0 \} = E_2 \{ \rho_t^\alpha \mid y'_0 \}$$

is defined \mathcal{F}_t^y -almost surely. We denote this by $l_t^\alpha(y)$.

In the completely observed case where both x and y are observable, the likelihood ratio is given by the Radon–Nikodym derivative ρ_t^α (Ref. 3). On the other hand, for the partially observed case one should consider $l_t^\alpha(y)$, as defined above, to be the likelihood ratio. It is known (Ref. 3) that the maximum likelihood estimate of α for the completely observed case is obtained by maximizing ρ_t^α (or $\ln \rho_t^\alpha$). For the partially observed case, this is obtained by maximizing

$$l_t^\alpha(y) = E_2 \{ \rho_t^\alpha \mid y'_0 \}.$$

Note, however, that

$$E_2 \{ \rho_t^\alpha \mid y'_0 \} = \langle \varphi^\alpha(t), 1 \rangle, \tag{13}$$

where $\varphi^\alpha(t)$, $t \geq 0$, is the solution of the Zakai equation (9) corresponding to the realization y .

4. Formulation of Identification Problem

In this section, we use the pathwise description of the Zakai equation (see, for example, Refs. 11–12) to formulate the identification problem as a deterministic control problem. Then, following similar arguments as those of Ahmed (Refs. 13–16), we show that this problem has a solution. First, let us verify how the above identification problem can be treated as an identification problem of systems governed by differential equations on Banach spaces.

Let the state process $x(t)$, $t \geq 0$, and the output process $y(t)$, $t \geq 0$, be governed by (1) and (2), respectively, with α being unknown. Let $\varphi^\alpha(t)$,

$t \geq 0$, $\alpha \in \mathcal{P}$, be the strong solution of Zakai equation (9). Clearly, for each $\alpha \in \mathcal{P}$, the solution $\varphi^\alpha(t)$, $t \geq 0$, is \mathcal{F}_t^y -adapted. Utilizing the maximum likelihood approach, given the history $\{y(s), s \leq t\}$, the unknown parameter α is determined by maximizing (13) over \mathcal{P} , subject to the constraint (9). Clearly, the choice $\alpha \equiv \alpha_t$ is dependent on the available information $\{y(s), s \leq t\}$.

It is interesting to note that, when the process $x(t)$, $t \geq 0$, is completely observable and governed by

$$dx(t) = a(t, x(t), \alpha) dt + b(t, c(t)) dW(t), \quad t \geq 0, \tag{14a}$$

$$x(0) = x_0, \tag{14b}$$

with α being unknown, the likelihood ratio is given by (Ref. 3)

$$\begin{aligned} \tilde{J}_t^\alpha \equiv \rho_t^\alpha = \exp \left\{ - (1/2) \int_0^t |b^{-1}(s, x(s))a(s, x(s), \alpha)|^2 ds \right. \\ \left. + \int_0^t b^{-1}(s, x(s))a(s, x(s), \alpha) \cdot b^{-1}(s, x(s)) dx(s) \right\}. \end{aligned}$$

Using Ito's lemma, one can easily verify that ρ_t^α , $t \geq 0$, satisfies the following integral equation:

$$\rho_t^\alpha = 1 + \int_0^t \rho_s^\alpha b^{-1}(s, x(s))a(s, x(s), \alpha) \cdot b^{-1}(s, x(s)) dx(s), \quad \text{a.s.}$$

In this case, the identification problem can be stated as follows. Given the path $\{x(s), s \leq t\}$, find an $\alpha^0 \in \mathcal{P}$ such that $\tilde{J}_t(\alpha^0) \geq \tilde{J}_t(\alpha)$, for all $\alpha \in \mathcal{P}$. For the case where the drift coefficient a is linear in α , i.e., $a(t, x, \alpha) \equiv K(t, x)\alpha$, where $K(t, x) \in R^{(n \times m)}$, Lipster and Shirayayev (Ref. 9) obtained an explicit expression for the optimal parameter α_t^0 , $t \geq 0$. This is given by

$$\begin{aligned} \alpha_t^0 = & \left(\int_0^t K'(s, x(s))[(bb')(s, x(s))]^{-1}K(s, x(s)) ds \right)^{-1} \\ & \times \int_0^t K'(s, x(s))((bb')(s, x(s)))^{-1} dx(s). \end{aligned}$$

From the above expression, it is clear that α_t^0 , $t \geq 0$, is \mathcal{F}_t^y -adapted. For the case where process $x(t)$, $t \geq 0$, is partially observable, the identification problem becomes much more difficult. In this case, one can treat this problem as an identification problem of infinite-dimensional systems by considering the unnormalized density $\varphi^\alpha(s)$, $s \leq t$ [see Eq. (9)], to be the state and the likelihood ratio,

$$J_t(\alpha) \equiv E_2\{\rho_t^\alpha \mid \mathcal{F}_t^y\} \equiv \langle \varphi^\alpha(t), 1 \rangle, \tag{15}$$

to be the objective functional. Here, the output process $\{y(s), s \leq t\}$ is considered to be the input to the Zakai equation (state equation). By maximizing $J_t(\alpha)$ over \mathcal{P} , one obtains the maximum likelihood estimate α_t° , which is clearly a functional of the observed history $\{y(s), s \leq t\}$.

In view of the above discussion, we can formulate the identification problem as follows. Define $p^\alpha(t) \equiv p^\alpha(t, \cdot)$, $t \geq 0$, such that

$$\varphi^\alpha(t) = p^\alpha(t) \exp(h^\alpha \cdot Z(t)), \quad t \geq 0, \tag{16}$$

where $\varphi^\alpha(t)$, $t \geq 0$, $\alpha \in \mathcal{P}$, is the solution of (9), the dot denotes the scalar product in R^m , and $Z(t)$, $t \geq 0$, is given by

$$Z(t) \equiv \int_0^t (\sigma_o \sigma_o')^{-1}(s, y(s)) dy(s) \equiv \int_0^t \Gamma_o^{-1}(s) dy(s). \tag{17}$$

Using (16)–(17) and utilizing Ito’s lemma, one can convert the Zakai equation (9), which is driven by the process y , into a parabolic partial differential equation with coefficients parametrized by the output process $\{y(s), s \leq t\}$; see Ref. 12. This equation is given by

$$(d/dt)p^\alpha(t) = F(t, \alpha)p^\alpha(t), \quad t \geq 0, \tag{18a}$$

$$p^\alpha(0) = p_0, \quad \alpha \in \mathcal{P}, \tag{18b}$$

where the operator F is given by

$$\begin{aligned} F(t, \alpha)u &= \exp(-h^\alpha \cdot Z(t))(A^*(t, \alpha)u \exp(h^\alpha \cdot Z(t))) \\ &\quad - (1/2)(\Gamma_o^{-1}(t)h^\alpha \cdot h^\alpha)u. \end{aligned} \tag{19}$$

Let $H \equiv L_2(R^n)$, and consider the Sobolev space

$$V \equiv M^1 \equiv \{f \in H: \partial f / \partial x_i \in H, 1 \leq i \leq n\},$$

with $V' \equiv (H^1)'$ being its dual. Let $\mathcal{L}(V, V')$ denote the class of all bounded linear operators from V to V' . Then, for any $u, v \in V$, the operator F gives rise to the following bilinear form:

$$\begin{aligned} \langle F(t, \alpha)u, v \rangle &= (1/2) \left(Z(t) \cdot \sum_{i,j=1}^n \int_{R^n} \sigma_{ij}^\alpha (\partial h^\alpha / \partial x_i) (\partial u / \partial x_j) v \, dx \right) \\ &\quad - (1/2) \sum_{i,j=1}^n \int_{R^n} \sigma_{ij}^\alpha (\partial u / \partial x_j) (\partial v / \partial x_i) \, dx \\ &\quad + \sum_{i=1}^n \int_{R^n} \tilde{a}_i^\alpha (\partial v / \partial x_i) u \, dx + \sum_{i=1}^n \int_{R^n} \tilde{a}_i^\alpha u v \, dx, \end{aligned} \tag{20}$$

where $\langle \cdot, \cdot \rangle$ denotes the pairing of V and V' . Further, the coefficients \tilde{a}_i^α and \tilde{a}_i^α are given by

$$\tilde{a}_i^\alpha \equiv a_i^\alpha - (1/2) \sum_{j=1}^n \partial \sigma_{ij}^\alpha / \partial x_j - (1/2) \sum_{j=1}^n \sigma_{ij}^\alpha (\partial h^\alpha / \partial x_j) \cdot Z(t), \quad (21)$$

$$\begin{aligned} \tilde{a}_i^\alpha \equiv & (1/2) Z(t) \cdot \sum_{j=1}^n (\partial \sigma_{ij}^\alpha / \partial x_j) (\partial h^\alpha / \partial x_i) - a_i^\alpha (\partial h^\alpha / \partial x_i) \cdot Z(t) \\ & + (1/2) \sum_{j=1}^n \sigma_{ij}^\alpha ((\partial h^\alpha / \partial x_i) \cdot Z(t)) ((\partial h^\alpha / \partial x_j) \cdot Z(t)) \\ & - (1/2) (\Gamma_o^{-1}(t) h^\alpha \cdot h^\alpha), \end{aligned} \quad (22)$$

with

$$a_i^\alpha \equiv a_i(t, x, \alpha),$$

$$\sigma_{ij}^\alpha \equiv \sigma_{ij}(t, x, \alpha),$$

and $Z(t), t \geq 0$, is given by (17).

Using (13), the stochastic identification problem can be formulated as an identification problem of a deterministic infinite-dimensional system as follows.

Problem (P). Given the system

$$(d/dt)p^\alpha(t) = F(t, \alpha)p^\alpha(t), \quad t \geq 0, \quad (23a)$$

$$p^\alpha(0) = p_0, \quad \alpha \in \mathcal{P}, \quad (23b)$$

find an $\alpha^\circ \in \mathcal{P}$ such that $J_t(\alpha^\circ) \geq J_t(\alpha)$, for all $\alpha \in \mathcal{P}$, where

$$J_t(\alpha) \equiv \langle \varphi^\alpha(t), 1 \rangle \equiv \langle p^\alpha(t) \exp(h^\alpha \cdot Z(t)), 1 \rangle. \quad (24)$$

In the remainder of this section we will show that Problem (P), as stated above, has a solution. For this, we need the following result which shows that the initial-value problem (23) has a unique weak solution and it satisfies certain bounds if the initial distribution p_0 has certain properties.

Lemma 4.1. Suppose that Assumptions (A1)–(A6) hold. Then:

- (i) for every $p_0 \in H$ and $\alpha \in \mathcal{P}$, the Cauchy problem (23) has a unique solution $p^\alpha \in L_2((0, t); V) \cap C([0, t]; H)$;
- (ii) for every $p_0 \in H$ satisfying

$$|p_0(x)| \leq \beta \exp(-\eta|x|^2),$$

for some $\beta, \eta > 0$, there exist $\gamma \equiv \gamma(\beta, \eta)$ and $\delta, 0 < \delta \leq \eta$ (possibly depending on β, η , and t) such that

$$|p^\alpha(s, x)| \leq \gamma \exp(-\delta|x|^2),$$

for all $0 \leq s \leq t$ and $\alpha \in \mathcal{P}$. (25)

Proof. The first part is a special case of Ref. 16, Theorem 1.1. The second part follows from the fact that, under the given assumptions, the fundamental solution of the initial-value problem (23), denoted $S^\alpha(x, t; \xi, \tau), t > \tau$, satisfied the following estimate (see Ref. 4):

$$|S^\alpha(x, t; \xi, \tau)| \leq [k_1/(t-\tau)^{n/2}] \exp\{-k_2|x-\xi|^2/(t-\tau)\},$$

for all $\alpha \in \mathcal{P}, x, \xi \in R^n$, and $0 \leq \tau \leq t$, where k_1 and k_2 are certain positive constants depending on the coefficients of the operator F and the parameter set \mathcal{P} . Using this estimate along with the assumptions on p_0 , one can verify (25). □

Defining

$$q^\alpha(t, x) \equiv p^\alpha(t, x) \exp\{(\delta/2)|x|^2\},$$

one can easily verify that q^α satisfies the Cauchy problem

$$(d/dt)q^\alpha = G(t, \alpha)q^\alpha, \quad t > 0, \tag{26a}$$

$$q^\alpha(0) = q_0, \tag{26b}$$

where

$$q_0(x) = p_0(x) \exp\{(\delta/2)|x|^2\}$$

and

$$\begin{aligned} G(t, \alpha)f = & \exp\{-h^\alpha \cdot Z(t) + (\delta/2)|x|^2\} A^*(t, \alpha) \\ & \times (f \exp\{h^\alpha \cdot Z(t) - (\delta/2)|x|^2\}) \\ & - (1/2)(\Gamma_\sigma^{-1}(t)h^\alpha \cdot h^\alpha). \end{aligned} \tag{27}$$

Under Assumptions (A1)–(A6), one can verify that the operator G satisfies the following properties:

- (P1) For each $\alpha \in \mathcal{P}$, and for any $u, v \in V$, the mapping $t \rightarrow \langle G(t, \alpha)u, v \rangle$ is measurable and there exists a constant $c > 0$ such that

$$|\langle G(t, \alpha)u, v \rangle| \leq c\|u\|_V \|v\|_V, \quad t \geq 0.$$

- (P2) There exist constant $\gamma > 0$ and $\tilde{\gamma} \in R$ such that

$$-\langle G(t, \alpha)u, v \rangle + \tilde{\gamma} \|u\|_H^2 \geq \gamma \|u\|_V^2, \quad \alpha \in \mathcal{P}, t \geq 0.$$
- (P3) For any sequence $\{\alpha^n\}$ that converged to α^o in \mathcal{P} ,

$$G(t, \alpha^n) \rightarrow G(t, \alpha^o), \quad t \geq 0,$$

in the strong operator topology of $\mathcal{L}(V, V')$.

- (P4) The mapping $\alpha \rightarrow G(t, \alpha), t \geq 0$, is once Gateaux differentiable in the strong operator topology of $\mathcal{L}(V, V')$ in the sense that

$$\lim_{\epsilon \downarrow 0} \|\{[G(t, \alpha^\epsilon)u - G(t, \alpha^o)u]/\epsilon\} \tilde{G}(t, \alpha^o; \alpha - \alpha^o)u\|_{V'} = 0,$$

for all $t \geq 0, \alpha, \alpha^o \in \mathcal{P}, 0 \leq \epsilon \leq 1$, and $u \in V$, where $\alpha^\epsilon \equiv \alpha^o + \epsilon(\alpha - \alpha^o)$ and $\tilde{G}(t, \alpha^o; \alpha - \alpha^o), t \geq 0$, denotes the Gateaux differential of G at the point α^o in the direction $\alpha - \alpha^o$.

From the above discussion, it is clear that $q_0 \in H$ and $q^\alpha \in L_2((0, t); V) \cap C([0, t]; H)$ for all $\alpha \in \mathcal{P}$ and $t < \infty$. The objective functional (24) can then be written in terms of q^α as

$$J_t(\alpha) = \langle q^\alpha(t), \eta^\alpha(t) \rangle_H, \tag{28}$$

where

$$\eta^\alpha(t) \equiv \eta^\alpha(t, x) = \exp\{h^\alpha(x) \cdot Z(t)\} \exp\{-(\delta/2)|x|^2\}, \quad x \in R^n.$$

Note that under Assumption (A5), $\eta^\alpha(t) \in H$ for each $t < \infty$. In fact, $\eta^\alpha(t) \in H$ even for h^α satisfying a linear growth condition. In view of the above discussion, the optimization problem (23)–(24) is equivalent to the problem (26)–(28). We recall that our problem is to find an $\alpha \in \mathcal{P}$ that maximizes (28), subject to the differential constraint (26).

The following result claims that Problem (P) has a solution.

Theorem 4.1. Existence. Consider Problem (P), and suppose that our basic assumptions hold and the \mathcal{P} is compact. Then, the mapping $\alpha \rightarrow J_t(\alpha), t < \infty$, where $J_t(\alpha)$ is given by (28), is continuous on \mathcal{P} and Problem (P) has a solution.

Proof. If $J_t(\alpha), t < \infty$, is infinite for some $\alpha \in \mathcal{P}$, there is nothing to prove. Thus, we assume that $J_t(\alpha) < \infty$, for all $\alpha \in \mathcal{P}$. For the proof of continuity, let $\alpha^n, \alpha^o \in \mathcal{P}$ such that $\alpha^n \rightarrow \alpha^o$, and let q^n and q^o denote the solutions of (26) corresponding to α^n and α^o , respectively. Defining

$$z^n(t) \equiv q^n(t) - q^o(t), \quad t \geq 0,$$

one can easily verify that $z^n(t)$, $t \geq 0$, satisfies the following differential equation:

$$(d/dt)z^n(t) = G(t, \alpha^n)z^n(t) + (G(t, \alpha^n) - G(t, \alpha^o))q^o(t), \quad t \geq 0, \tag{29a}$$

$$z^n(0) = 0. \tag{29b}$$

Scalar multiplying the above equation on both sides by z^n , integrating over $[0, t]$, and using Property (P2), we have

$$\begin{aligned} &|z^n(t)|_H^2 - 2\tilde{\gamma} \int_0^t |z^n(\theta)|_H^2 d\theta + 2\gamma \int_0^t |z^n(\theta)|_V^2 d\theta \\ &\leq 2 \int_0^t \langle (G(\theta, \alpha^n) - G(\theta, \alpha^o))q^o(\theta), z^n(\theta) \rangle_{V'-V} d\theta. \end{aligned}$$

Using the Schwartz inequality, it follows that

$$\begin{aligned} &|z^n(t)|_H^2 - 2\tilde{\gamma} \int_0^t |z^n(\theta)|_H^2 d\theta + 2\gamma \int_0^t |z^n(\theta)|_V^2 d\theta \\ &\leq 2 \left(\int_0^t \| (G(\theta, \alpha^n) - G(\theta, \alpha^o))q^o(\theta) \|_{V'}^2 d\theta \right)^{1/2} \\ &\times \left(\int_0^t |z^n(\theta)|_V^2 d\theta \right)^{1/2}. \end{aligned} \tag{30}$$

Using the elementary inequality

$$ab \leq (1/2\epsilon)a^2 + (\epsilon/2)b^2,$$

$a, b \in R$, and $\epsilon > 0$, and taking $\epsilon = \gamma$, it follows by the Gronwall lemma that

$$\begin{aligned} &|z^n(t)|_H^2 + \gamma \int_0^t |z^n(\theta)|_V^2 d\theta \\ &\leq [\exp\{2\tilde{\gamma}t/\gamma\}] \int_0^t \| (G(\theta, \alpha^n) - G(\theta, \alpha^o))q^o(\theta) \|_{V'}^2 d\theta. \end{aligned}$$

Therefore, $z^n \in L_2((0, t); V) \cap L_\infty([0, t]; H)$, and by Property (P3) and the dominated convergence theorem, it follows that

$$\lim_n \sup_t |z^n(t)|_H = 0,$$

and hence $q^n(t) \rightarrow q^o(t)$ in H , uniformly in t .

Since $\alpha \rightarrow h(x, \alpha)$ is continuous on \mathcal{P} , for each $x \in R^n$, and since h^α has at most linear growth, it follows by the dominated convergence theorem that

$$\lim_{n \rightarrow \infty} |\eta^{\alpha^n}(t) - \eta^{\alpha^0}(t)|_H = 0, \quad \text{for each } t < \infty,$$

whenever $\alpha^n \rightarrow \alpha^0$. This fact and the continuity of q^α on \mathcal{P} , as shown above, imply the continuity of $J_t(\alpha)$ on \mathcal{P} . Since \mathcal{P} is compact, $J_t(\alpha)$ attains its maximum on \mathcal{P} . This completes the proof. \square

In the next section, we utilize standard variational arguments (see, for example, Refs. 13–16) and make use of Gateaux differentiability of q^α on \mathcal{P} to derive the necessary conditions of optimality for Problem (P).

5. Necessary Conditions for Optimal Identification

In this section, we present the necessary conditions of optimality for the identification problem [Problem (P)] as stated in the previous section. In our derivations, we shall follow similar arguments as those of Ahmed (Refs. 13–15) and make use of the Gateaux differentiability of q^α [see Eq. (26)] on the parameter set \mathcal{P} .

Let $\alpha^\epsilon \equiv \alpha^0 + \epsilon(\alpha - \alpha^0)$, $\epsilon \in [0, 1]$, and let $q^\epsilon(t) \equiv q(t, \alpha^\epsilon)$ and $q^0(t) \equiv q(t, \alpha^0)$, $t \geq 0$, denote the solutions of the initial-value problem (26) corresponding to α^ϵ and α^0 , respectively. Let

$$\tilde{q}^0(t) \equiv \tilde{q}(t, \alpha^0, \alpha - \alpha^0) \equiv \lim_{\epsilon \downarrow 0} [q^\epsilon(t) - q^0(t)]/\epsilon, \quad t \geq 0,$$

denote the Gateaux differential of q at α^0 in the direction $\alpha - \alpha^0$. The following result shows that the Gateaux differential \tilde{q}^0 exists and it is the solution of a related differential equation.

Lemma 5.1. Consider the system (26) and suppose that Assumptions (A1)–(A6) hold and the \mathcal{P} is compact and convex. Then, the map $\alpha \rightarrow q^\alpha$ is Gateaux differentiable on \mathcal{P} . Further, at each point $\alpha^0 \in \mathcal{P}$, the Gateaux differential of q in the direction $\alpha - \alpha^0$, denoted by $\tilde{q}(t, \alpha^0, \alpha - \alpha^0)$, $t \geq 0$, is given by the weak solution of the following differential equation:

$$(d/dt)\xi(t) = G(t, \alpha^0)\xi(t) + \tilde{G}(t, \alpha^0, \alpha - \alpha^0)q^0(t), \quad t \geq 0, \quad (31a)$$

$$\xi(0) = 0, \quad (31b)$$

where q^0 is the solution of (26) corresponding to α^0 and \tilde{G} is the Gateaux differential of G in the sense of Property (P4).

Proof. Let $\alpha^\circ, \alpha \in \mathcal{P}$. Since \mathcal{P} is convex, we have

$$\alpha^\epsilon \equiv \alpha^\circ + \epsilon(\alpha - \alpha^\circ) \in \mathcal{P}, \quad 0 \leq \epsilon \leq 1.$$

Defining

$$\tilde{q}^\epsilon(t) \equiv (1/\epsilon)(q^\epsilon(t) - q^\circ(t)), \quad t \geq 0,$$

and using (26), one can easily verify that

$$\begin{aligned} (d/dt)\tilde{q}^\epsilon(t) &= G(t, \alpha^\epsilon)\tilde{q}^\epsilon(t) \\ &\quad + (1/\epsilon)(G(t, \alpha^\epsilon) - G(t, \alpha^\circ))q^\circ(t), \quad t \geq 0, \end{aligned} \tag{32a}$$

$$\tilde{q}^\epsilon(0) = 0. \tag{32b}$$

By arguments similar to those of Theorem 4.1, we arrive at the following estimate:

$$\begin{aligned} &|\tilde{q}^\epsilon(t)|_H^2 + \gamma \int_0^t |\tilde{q}^\epsilon(\theta)|_V^2 d\theta \\ &\leq [\exp\{2\tilde{\gamma}t/\gamma\}] \int_0^t \|[G(\theta, \alpha^\epsilon) - G(\theta, \alpha^\circ)]/\epsilon\}q^\circ(\theta)\|_{V'} d\theta, \end{aligned}$$

for all $0 \leq t < \infty$. Hence, it follows from the above inequality and Property (P4) that the set $\{\tilde{q}^\epsilon, \epsilon \in [0, 1]\}$ is contained in a bounded subset of $L_2((0, t); V) \cap L_\infty([0, t]; H)$. Hence, from every sequence $\tilde{q}^n \equiv \tilde{q}^{\epsilon_n}$, with $\epsilon_n \in [0, 1]$ and $\epsilon_n \rightarrow 0$, one can extract a subsequence relabeled as $\{\tilde{q}^n\}$ and $\tilde{q}^\circ \in L_2((0, t); V) \cap L_\infty([0, t]; H)$ such that $\tilde{q}^n \rightarrow \tilde{q}^\circ$ weakly in $L_2((0, t); V)$. Hence, the Gateaux differential of q exists and is given by $\tilde{q}(t, \alpha^\circ, \alpha - \alpha^\circ) \equiv \tilde{q}^\circ(t), t \geq 0$. It remains to show that \tilde{q}° satisfies (31). Indeed, since $G(t, \alpha^n) \rightarrow G(t, \alpha^\circ)$ in the strong operator topology of $\mathcal{L}(V, V')$ and $\tilde{q}^n \rightarrow \tilde{q}^\circ$ weakly in $L_2((0, t); V)$, then $G(\cdot, \alpha^n)\tilde{q}^n \rightarrow G(\cdot, \alpha^\circ)\tilde{q}^\circ$ weakly in $L_2((0, t); V')$. Hence, by Property (P4), it follows from (32) that $(d/dt)\tilde{q}^n \in L_2((0, t); V')$ for all n and $(d/dt)\tilde{q}^n \rightarrow \psi$ in $L_2((0, t); V')$, for a suitable ψ in $L_2((0, t); V')$, and that ψ is the distributional derivative of \tilde{q}° . Hence, \tilde{q}° satisfies the differential equation

$$(d/dt)\tilde{q}^\circ(t) = G(t, \alpha^\circ)\tilde{q}^\circ(t) + \tilde{G}(t, \alpha^\circ, \alpha - \alpha^\circ)q^\circ(t),$$

in the sense of distribution in V' . Since $\tilde{q}^\circ \in L_2((0, t); V)$ and $(d/dt)\tilde{q}^\circ \in L_2((0, t); V')$, it is clear that $\tilde{q}^\circ \in C([0, t]; H)$ and $\tilde{q}^\circ(0)$ is well defined and equals $\tilde{q}^n(0) = 0$ for all n . Hence, \tilde{q}° satisfies the differential equation (31) and one may identify \tilde{q}° as ξ . This completes the proof. \square

With the help of the above lemma, we now prove the following necessary conditions of optimality for Problem (P).

Theorem 5.1. Necessary Conditions of Optimality. Consider Problem (P) given by (23) and (24), or equivalently (26) and (28), and suppose that Theorem 4.1 holds. Then, in order that α° be the maximum likelihood estimate of the unknown parameter α , it is necessary that it satisfies the optimality conditions given by the system equation

$$(d/ds)q(s) = G(s, \alpha^\circ)q(s), \quad 0 < s \leq t < \infty, \tag{33a}$$

$$q(0) = q_0, \tag{33b}$$

the adjoint equation

$$-(d/ds)r(s) = G^*(s, \alpha^\circ)r(s), \quad 0 < s \leq t < \infty, \tag{34a}$$

$$r(t) = \eta^{\alpha^\circ}(t) \equiv \eta^\circ(t), \tag{34b}$$

and the inequality

$$\int_0^t \langle \tilde{G}(s, \alpha^\circ, \alpha - \alpha^\circ)q^\circ(s), r^\circ(s) \rangle ds + \langle q^\circ(t), \tilde{\eta}(t, \alpha^\circ, \alpha - \alpha^\circ) \rangle \leq 0, \tag{35}$$

for all $\alpha \in \mathcal{P}$. Here, G^* is the formal adjoint of G ; \tilde{G} and $\tilde{\eta}$ are the Gateaux differentials of G and η , respectively; and q° and r° are the solutions of (33) and (34).

Proof. The proof follows from standard variational arguments as in Refs. 13–16. Since $\alpha \rightarrow q^\alpha$ has a Gateaux differential on \mathcal{P} , it follows that J , as defined by (28), also has a Gateaux differential. Then, in order that J attains its maximum at $\alpha^\circ \in \mathcal{P}$, it is necessary that

$$J'_t(\alpha^\circ; \alpha - \alpha^\circ) \equiv \lim_{\epsilon \downarrow 0} (1/\epsilon) \{ J_t(\alpha^\circ + \epsilon(\alpha - \alpha^\circ)) - J_t(\alpha^\circ) \} \leq 0, \tag{36}$$

for all $\alpha \in \mathcal{P}$. Using the result of Lemma 5.1, it follows from (28) and (36) that

$$J'_t(\alpha^\circ, \alpha - \alpha^\circ) = \langle \tilde{q}^\circ(t), \eta^\circ(t) \rangle + \langle q^\circ(t), \tilde{\eta}(t, \alpha^\circ; \alpha - \alpha^\circ) \rangle \leq 0, \tag{37}$$

for all $\alpha \in \mathcal{P}$, where \tilde{q}° denotes the Gateaux differential of q as defined by Lemma 5.1. Inequality (37) can be further simplified by introducing the adjoint variable r , which is the solution of the following differential equation:

$$-(d/ds)r(s) = G^*(s, \alpha^\circ)r(s), \quad 0 \leq s < t < \infty, \tag{38a}$$

$$r(t) = \eta^\circ(t) = \eta^\circ(t). \tag{38b}$$

Reversing the flow of time, $s \rightarrow t - s$, and noting that $\eta^\circ(t) \in H$, $t \geq 0$, it follows from Lemma 4.1 that (38) has a unique solution

$r \in L_2((0, t); V) \cap C([0, t]; H)$. Using (37), (38), and Lemma 5.1, one can easily verify that

$$\langle \tilde{q}^\circ(t), \eta^\circ(t) \rangle = \int_0^t \langle \tilde{G}(s, \alpha^\circ, \alpha - \alpha^\circ) q^\circ(s), r^\circ(s) \rangle ds. \tag{39}$$

Now, Inequality (35) follows from (37) and (39). This completes the proof. \square

Remark 5.1. Here we have not considered the question of consistency of the estimated parameter α° . This question was settled in Ref. 3 for the case where the state process is completely observable and the drift coefficient is linear in α . The consistency question was also settled in Refs. 1–2 for the case where both state and observed processes are governed by linear time-invariant stochastic systems. For partially observed nonlinear stochastic systems, this remains as open problem.

As a final remark, it should be noted that similar results can be obtained for the general case where the state process x is governed by a stochastic differential equation of the form

$$\begin{aligned} dx(t) &= a(t, x, \alpha) dt + b(t, x, \alpha) dW(t) \\ &\quad + \int_{R_0^c} c(t, x, \alpha, \xi) \beta(dt, d\xi), \quad t \geq 0, \\ x(0) &= x_0, \end{aligned}$$

where $R_0^c \equiv R' \setminus \{0\}$ and β is a counting measure obeying a generalized Poisson distribution with certain mean.

References

1. TUNGAIT, J. K., *Identification and Model Approximation for Continuous-Time Systems on Finite Parameter Sets*, IEEE Transactions on Automatic Control, Vol. AC-25, pp. 1202–1206, 1980.
2. TUNGAIT, J. K., *Continuous-Time System Identification on Compact Parameter Sets*, IEEE Transactions on Information Theory, Vol. IT-31, pp. 652–659, 1985.
3. LIPSTER, R. S., and SHIRYAYEV, A. N., *Statistics of Random Processes*, Vols. 1–2, Springer-Verlag, Berlin, Germany, 1978.
4. LEGLAND, F., *Nonlinear Filtering and Problem of Parametric Estimation*, Stochastic Systems: The Mathematics of Filtering and Identification and Applications, Edited by M. Hazewinkel and J. Willems, D. Reidel Publishing Co., Boston, Massachusetts, pp. 613–620, 1980.

5. KUMAR, P. R., and VARAIYA, P., *Stochastic Systems: Estimation, Identification, and Adaptive Control*, Prentice-Hall, Englewood Cliffs, New Jersey, 1986.
6. TUNGAIT, J. K., *Global Identification of Continuous Time Systems with Unknown Noise Covariance*, IEEE Transactions on Information Theory, Vol. IT-28, pp. 531–536, 1982.
7. AHMED, N. U., *Elements of Finite-Dimensional Systems and Control Theory*, Longman Scientific and Technical, London, England, 1988.
8. KUSHNER, H. J., *Dynamical Equations for Optimal Nonlinear Filtering*, Journal of Differential Equations, Vol. 3, pp. 179–190, 1967.
9. ZAKAI, M., *On the Optimal Filtering of Diffusion Processes*, Zeitschrift für Wahrscheinlichkeits Theorie, Verwindette, Vol. 11, pp. 203–243, 1969.
10. GIRSANOV, I. V., *On Transforming a Certain Class of Stochastic Process by Absolutely Continuous Substitution of Measures*, Theory of Probability and Applications, Vol. 5, pp. 285–301, 1960.
11. FLEMING, W. H., and PARDOUX, E., *Optimal Control for Partially Observed Diffusions*, SIAM Journal of Control and Optimization, Vol. 20, pp. 261–285, 1982.
12. DAVIS, M. H., *Pathwise Nonlinear Filtering*, Stochastic Systems: The Mathematics of Filtering and Identification and Applications, Edited by M. Hazewinkel and J. Willems, D. Reidel Publishing Co., Boston, Massachusetts, pp. 505–529, 1980.
13. AHMED, N. U., and TEO, K. L., *Optimal Control of Distributed-Parameter Systems*, North-Holland, New York, New York, 1981.
14. AHMED, N. U., *Identification of Linear Operators in Differential Equations on Banach Space*, Operator Methods for Optimal Control Problems, Edited by S. J. Lee, Marcel Dekker, New York, New York, pp. 1–35, 1987.
15. AHMED, N. U., *Identification of Operators in Systems Governed by Evolution Equations on Banach Space*, IFIP Conference on Optimal Control of Systems Governed by Partial Differential Equations, Santiago de Compostela, Spain, pp. 610–630, 1987.
16. AHMED, N. U., *Optimization and Identification of Systems Governed by Evolution Equations on Banach Space*, Pitman Research Notes in Mathematical Sciences, Longman, Boston, Massachusetts, Vol. 184, 1988.
17. FRIEDMAN, A., *Partial Differential Equations of Parabolic Type*, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.