Journal of Multivariate Analysis **80**, 322–343 (2002)
doi:10.1006/jmva.2001.1989, available online at http://www.idealibrary.com on IDEAL®

Parameter Estimation for Controlled Semilinear Stochastic Systems: Identifiability and Consistency[1]

B. Goldys

The University of New South Wales, Sydney, Australia
E-mail: B.Goldys@unsw.edu.au

and

B. Maslowski

Czech Academy of Sciences, Praha, Czech Republic
E-mail: maslow@math.cas.cz

Received August 4, 1999; published online August 21, 2001

We consider a controlled stochastic semilinear evolution equation with the drift depending on the unknown parameter. We show that the maximum likelihood estimator is strongly consistent for a class of bounded predictable controls.

AMS 1991 subject classifications: 62M09; 62M40; 93E10; 93E35.

Key words and phrases: semilinear stochastic evolution equation; maximum likelihood estimator; identifiability; invariant measure; strong law of large numbers; transition semigroup; infinitesimal generator.

1. INTRODUCTION

The aim of this paper is to study the maximum likelihood estimator of the parameter $\alpha_0 \in \mathbb{R}^d$ in a controlled stochastic semilinear equation

$$\begin{cases} dX_t = (AX_t + f(\alpha_0, X_t) - u_t)\,dt + \sqrt{Q}\,dW_t, & t \geqslant 0, \\ X_0 = x, \end{cases} \tag{1.1}$$

considered in a separable Hilbert space H. The process W is a cylindrical Wiener process on H and (u_t) is a uniformly bounded progressively measurable process. We assume that $f: H \to H$ is Lipschitz and Gateaux

[1] This work was partially supported by the Small ARC Grant Scheme and by GAČR Grant 201/98/1454.

0047-259X/01 $35.00

differentiable and A generates a strongly continuous semigroup on H. It is well known (see monograph [14] for a thorough discussion of this subject) that a large variety of stochastic partial differential equations may be studied as special cases of (1.1), see also an example at the end of this paper.

Estimation of the drift parameter in a stochastic differential equation of type (1.1) with $u=0$ has been studied for a long time in the case when $H=\mathbb{R}^d$; see [28] and references therein for an extensive study of the this problem. The first proofs of consistency of the Maximum Likelihood Estimators of α_0 in finite dimensional case have been obtained in the seminal papers [29, 30]. The estimation problem was also studied in the context of adaptive control of diffusion processes in [7–9].

Most of the results on parameter estimation for stochastic evolution equations in infinite dimensional spaces we are aware of concern the linear systems. For example, the least squares method has been used in [15, 16] to define a strongly consistent family of estimators for controlled linear systems of stochastic PDE's with distributed and boundary controls, respectively. The maximum likelihood method has been employed in [2–4] to estimate unknown parameters (which can be vectors in $\mathbb{R}^d$ or functions) in stochastic parabolic PDEs; see also the recent paper [6], where the unknown parameter is a function (in general discontinuous) involved in the second order term of the parabolic operator. The hyperbolic case was treated in [5]. Consistency of the maximum likelihood Estimators (MLE) in a different sense from ours was recently studied in interesting papers [22, 23] for certain linear stochastic parabolic PDEs. A study of the Bayes estimators for the parameters of linear stochastic partial differential equations has been recently initiated in [31].

It seems that few results are available in the case of nonlinear stochastic infinite-dimensional systems. For an uncontrolled potential type parameter-dependent system of stochastic reaction-diffusion equations a minimum contrast estimator is proposed and its strong consistency is proved in [26]; the method is based on an earlier result for linear systems obtained in [24].

The aim of this paper is to show that if α_0 is identifiable and additionally f is Hölder continuous in α then the maximum likelihood estimator α_t of α_0 is strongly consistent, that is,

$$\mathbb{P}(\alpha_t \to \alpha_0)=1,$$

under some reasonable assumptions on the control (u_t). Our work is motivated by by the recent works on the ergodic adaptive control of stochastic partial differential equations, see [17]; see Remark 4.4 for more details.

2. SETTING OF THE PROBLEM AND SOME AUXILIARY FACTS

We consider an equation

$$\begin{cases} dX_t = (AX_t + f(\alpha_0, X_t) - u_t)\, dt + \sqrt{Q}\, dW_t, & t \geqslant 0 \\ X_0 = x, \end{cases} \tag{2.1}$$

on a separable Hilbert space H with the norm $|\cdot|$ and the inner product $\langle \cdot, \cdot \rangle$. The following is a standing hypothesis for the rest of the paper.

HYPOTHESIS 2.1. (i) *A generates a C_0-semigroup (S_t) on H. Moreover, there exists $\omega \in \mathbb{R}$ such that*

$$\|S_t\| \leqslant e^{\omega t}, \qquad t \geqslant 0. \tag{2.2}$$

(ii) *W is a standard cylindrical Wiener process on H defined on a stochastic basis $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ and $Q = Q^*$ is a bounded operator on H with bounded inverse.*

(iii) *$\mathscr{A}$ is a compact subset of $\mathbb{R}^d$ and $f\colon \mathscr{A} \times H \to H$ is such that $f(\alpha, \cdot)$ is Gateaux differentiable for each $\alpha \in \mathscr{A}$. Moreover, there exists $K > 0$ such that*

$$|f(\alpha, x)| \leqslant K(1 + |x|), \qquad x \in H, \tag{2.3}$$

and

$$\langle Ax + f(\alpha, x + y), x \rangle \leqslant -K\, |x|^2 + a(|y|)\, |x|, \qquad x \in \mathscr{D}(A),\ y \in H, \tag{2.4}$$

where $a\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous increasing function.

(iv) *There exists $T > 0$ such that*

$$\int_0^T \|S_t\|_{HS}^2\, dt < \infty,$$

where $\|\cdot\|_{HS}$ denotes the Hilbert–Schmidt norm of an operator.

(v) *Controls u are taken from the set of admissible controls $\mathscr{U}$ which consists of progressively measurable processes $u\colon \mathbb{R}_+ \times \Omega \to H$ such that $\mathbb{P}(u_t \in B(r_0)) = 1$ for all $t \geqslant 0$, where $B(r_0) \subset H$ stands for a centered ball with the radius r_0.*

Remark 2.2. (a) Condition (iv) which may seem unnecessarily restrictive is a direct consequence of Condition (ii). If the noise is cylindrical then (iv) is necessary for the existence of an H-valued solution to (1.1).

(b) Assumption (v) which allows bounded controls only, excludes unstable systems stabilised by the adaptive control. This condition is technical and it would be interesting to remove it. This will be a subject of a future paper.

Let (Z_t) be a solution to the linear equation

$$\begin{cases} dZ_t = AZ_t\,dt + \sqrt{Q}\,dW_t, \\ Z_0 = x,\ t \geqslant 0. \end{cases} \tag{2.5}$$

In view of condition (iii) we can assume in (2.2) without loss of generality that $\omega < 0$. Taking (iv) into account we find in this case that the process (Z_t) has a unique invariant measure $\mu = N(0, Q_\infty)$, where the operator

$$Q_\infty = \int_0^\infty S_t Q S_t^*\,dt$$

is of trace class. Therefore we shall always assume that the following hypothesis holds.

HYPOTHESIS 2.3. (i) *There exists* $\omega > 0$ *such that*

$$\langle Ax, x \rangle \leqslant -\omega\,|x|^2, \qquad x \in \mathscr{D}(A). \tag{2.6}$$

(iii) *The function a introduced in Hypothesis* 2.1 *enjoys the property*

$$\int_H a\left(|y| + \frac{r_0}{\omega}\right)\mu(dy) < \infty. \tag{2.7}$$

The parameter $\alpha \in \mathscr{A}$ remains fixed throughout the rest of this section and for simplicity of notation it is omitted in all formulae.

If Hypothesis 2.1 holds then the equation

$$\begin{cases} dY_t^x = (AY_t^x + f(Y_t^x))\,dt + \sqrt{Q}\,dW_t, \qquad t \geqslant 0, \\ Y_0^x = x, \end{cases} \tag{2.8}$$

has a unique mild solution, see Theorem 7.4 in [14]. Let $P_t^f \phi(x) = \mathbb{E}\phi(Y_t^x)$ be a transition semigroup of the Markov process (Y_t^x). By Hypothesis 2.1 the semigroup (P_t^f) has a unique invariant measure μ^f [19] and therefore, the semigroup (P_t^f) extends to a C_0-semigroup of contractions on $L^2(H, \mu^f)$ with the generator L^f. The domain of L^f in $L^2(H, \mu^f)$ will be denoted by $\mathscr{D}(L^f)$.

Let $\mathscr{K} = \{k_n : n \geqslant 1\} \subset \mathscr{D}(A^*)$ be a countable set such that $\mathscr{K}_l = \mathrm{lin}(\mathscr{K})$ is dense in $\mathrm{dom}(A^*)$ when endowed with the graph norm. Let $\mathscr{G} = \{g_n : n \geqslant 1\} \subset \bigcup_{m=1}^\infty C_0^2(\mathbb{R}^m)$ has the property, that $\mathscr{G}_l = \mathrm{lin}\{g_n : n \geqslant 1\}$ is dense in each $C_0^2(\mathbb{R}^m)$, $m \geqslant 1$, in the topology of uniform convergence on compacts,

$$\mathscr{D}_0 = \{ g \circ P : g \in \mathscr{G}_l,\ P(H) \subset \mathscr{K}_l, \dim P(H) < \infty \},$$

where P stands for an orthogonal projection on H and $C_0^2(\mathbb{R}^n)$ is the space of twice differentiable functions on $\mathbb{R}^m$ with compact supports. Let $v: H \to \mathbb{R}$ be a bounded Borel function. If Hypothesis 2.1 holds then by the Girsanov Theorem the equation

$$\begin{cases} dY_t^{v,x} = (AY_t^{v,x} + f(Y_t^{v,x}) + v(Y_t^{v,x}))\,dt + \sqrt{Q}\,dW_t, \\ Y_0^{v,x} = x \in H, \end{cases}$$

has a unique weak solution. For every $\phi \in \mathscr{D}_0$ a differential operator

$$L^{f+v}\phi(x) = \tfrac{1}{2}\operatorname{tr}(QD^2\phi(x)) + \langle x, A^*D\phi(x)\rangle + \langle f(x) + v(x), D\phi(x)\rangle$$

is well defined and $L^{f+v}\phi \in L^2(H, \mu^f)$. The next proposition is a version of the results proved in [12, 13]. For the reader's convenience we sketch its proof.

PROPOSITION 2.4. *Assume Hypotheses* 2.1 *and* 2.3 *and let* $v: H \to \mathbb{R}$ *be a bounded Borel function. Then the operator* $(L^{f+v}, \mathscr{D}_0)$ *in* $L^2(H, \mu^f)$ *has a unique extension to a generator of a* C_0*-semigroup* (P_t^{f+v})*, still denoted by* L^{f+v} *and* $\mathscr{D}(L^{f+v}) = \mathscr{D}(L^f)$*. Moreover, the semigroup* (P_t^{f+v}) *may be identified with the transition semigroup of the process* Y^v.

Proof. Step 1. In this part of the proof we construct a core for the generator L^f in $L^2(H, \mu^f)$. Let us recall first some properties of the semigroup (P_t^f) acting the spaces of continuous functions. It is well known, see [14], that the semigroup (P_t^f) is Feller, that is, $P_t^f(C_b(H)) \subset C_b(H)$, where $C_b(H)$ stands for the Banach space of bounded continuous functions on H with the norm

$$\|\phi\|_\infty = \sup_{x \in H} |\phi(x)|.$$

Following [10] we say that a sequence $(\phi_n) \subset C_b(H)$ is $\mathscr{K}$-convergent to $\phi \in C_b(H)$, $\phi_n \xrightarrow{\mathscr{K}} \phi$, if

$$\sup_{n \geqslant 1} \|\phi_n\|_\infty < \infty, \tag{2.9}$$

and

$$\phi_n \to \phi, \qquad \text{uniformly on compacts.} \tag{2.10}$$

The domain $\mathscr{D}_0(L^f)$ of the operator L^f in $C_b(H)$ is defined as the set of all functions $\phi \in C_b(H)$ such that

$$\frac{P_t^f \phi - \phi}{t} \xrightarrow[t \to 0]{\mathscr{K}} \psi \qquad \text{(say).} \tag{2.11}$$

It is easy to see that $\mathscr{D}_0(L^f) \subset \mathscr{D}(L^f)$. Let

$$\mathscr{D}_1 = \{ g \circ P : g \in \bigcup_{m=1}^{\infty} C_b^2(\mathbb{R}^m),\ P(H) \subset \mathscr{D}(A^*), \dim P(H) < \infty \},$$

where $C_b^2(\mathbb{R}^m)$ is a space of twice differentiable bounded functions on $\mathbb{R}^m$. Let $C_1(H)$ denote the space of continuous functions ϕ: $H \to \mathbb{R}$ such that

$$\|\phi\|_1 = \sup_{x \in H} \frac{|\phi(x)|}{1 + |x|} < \infty.$$

The space $C_1(H)$ with the norm $\|\cdot\|_1$ is a Banach space and (P_t^f) is a semigroup of bounded operators on $C_1(H)$. If for a certain $\phi \in C_1(H)$

$$\frac{P_t^f \phi - \phi}{t} \xrightarrow[t \to 0]{\mathscr{K}} \psi \in C_1(H)$$

then we say that $\phi \in \mathscr{D}_1(L^f)$, the domain of the generator L^f of the semigroup (P_t^f) in $C_1(H)$, and we put $\psi = L^f \phi$. It has been proved in [20] that for every $\phi \in \mathscr{D}_1(L^f)$ there exists a sequence $(\phi_n) \subset \mathscr{D}_1$, such that

$$\phi_n \xrightarrow{\mathscr{K}} \phi \qquad \text{and} \qquad L^f \phi_n \xrightarrow{\mathscr{K}} L^f \phi. \tag{2.12}$$

It is easy to see that for every $\phi \in \mathscr{D}_1$ we can find $(\phi_n) \subset \mathscr{D}_0$ such that the above convergence holds. Therefore, for every $\phi \in \mathscr{D}_1(L^f)$ there exists a sequence $(\phi_n) \subset \mathscr{D}_0$ such that (2.12) holds. Finally, it follows that for $\phi \in \mathscr{D}_0(L^f)$, the domain of L^f in $C_b(H)$ (2.12) holds as well.

Step 2. In this part of the proof we use the same notation $\|\cdot\|$ for the norm in the space $L^2(H, \mu^f)$ and in the space $L^2(H, \mu^f; H)$ of H-valued functions. It follows from Step 1 that $\mathscr{D}_0$ is a core for the generator L^f of the semigroup (P_t^f) considered in $L^2(H, \mu^f)$. Then proceeding in the same way as in [13] one can show that for every $T > 0$ and $\phi \in \mathscr{D}_0$

$$\int_0^T \|Q^{1/2} D P_t^f \phi\|^2 \, dt = \|\phi\|^2 - \|P_T^f \phi\|^2. \tag{2.13}$$

Let

$$B\phi(x) = \langle v(x), D\phi(x)\rangle, \qquad \phi \in \mathscr{D}_0.$$

Since Q is boundedly invertible, we have

$$B\phi(x) = \langle Q^{-1/2}v(x), Q^{1/2}D\phi(x)\rangle, \qquad \phi \in \mathscr{D}_0.$$

It follows from (2.13) and the Hölder inequality that for $\phi \in \mathscr{D}_0$

$$\int_0^T \|BP_t^f \phi\| \, dt \leqslant c\sqrt{T}\left(\int_0^T \|Q^{1/2} DP_t^f \phi\|^2 \, dt\right)^{1/2} \leqslant c\sqrt{T}\,\|\phi\|,$$

where $c = \|Q\|^{-1/2}\,\|v\|_\infty$. For T sufficiently small we have $c\sqrt{T} < 1$, and therefore by the result in [32] the operator $(L^f + B, \mathscr{D}_0)$ extends to a generator L^{f+v} of a C_0-semigroup (P_t^{f+v}) on $L^2(H, \mu^f)$ and

$$\mathscr{D}(L^f) = \mathscr{D}(L^{f+v}).$$

It follows that the graph norms of L^f and L^{f+v} are equivalent, hence $\mathscr{D}_0$ is a core for L^{f+v} which implies that the extension of $(L^f + B, \mathscr{D}_0)$ to a generator of C_0 semigroup is unique. Finally, we note (but omit the details of the proof) that taking into account the Girsanov Theorem one can show that $P_t^{f+v}\,\phi(x) = \mathbb{E}\phi(Y_t^{v,x})$ for every $\phi \in L^2(H, \mu^f)$. ∎

Lemma 2.5. *Assume Hypotheses* 2.1 *and* 2.3. *Then the semigroup* (P_t^{f+v}) *has a unique invariant measure* μ^{f+v}. *Moreover, if* ν *is a probability measure on* H, *such that*

$$\int_H L^{f+v}\psi(x)\, \nu(dx) = 0, \qquad \psi \in \mathscr{D}_0, \tag{2.14}$$

then $\nu = \mu^{f+v}$.

Proof. We need to show that ν is an invariant measure for (P_t^{f+v}), that is,

$$\langle \nu, P_t^{f+v}\phi - \phi\rangle = 0 \tag{2.15}$$

for a measure determining set of bounded functions ϕ. To this end it is enough to show that for every $\beta > 0$

$$\int_0^\infty e^{-\beta t}\langle \nu, P_t^{f+v}\phi - \phi\rangle \, dt = 0. \tag{2.16}$$

Since for $\phi \in \mathscr{D}(L^{f+v})$

$$P_t^{f+v}\phi - \phi = \int_0^t L^{f+v} P_s^{f+v} \phi \, ds,$$

we find that (2.16) holds for a certain $\phi \in \mathscr{D}(L^{f+v})$ if and only if

$$\langle v, L^{f+v} J_\beta \phi \rangle = 0, \tag{2.17}$$

where $J_\beta = (\beta - L^{f+v})^{-1}$. Note that for every $\psi \in \mathscr{D}_0$ the function

$$(\beta - L^{f+v})\, \psi(x) = (\beta - L^f)\, \psi(x) - \langle v(x), D\psi(x) \rangle$$

is bounded and $(\beta - L^{f+v})\, \mathscr{D}_0$ is dense in $L^2(H, \mu^f)$ by Proposition 2.4. Hence (2.17) follows which proves (2.15) for every $\phi \in (\beta - L^{f+v})\, \mathscr{D}_0$. ∎

3. IDENTIFIABILITY OF α

Definition 3.1. The parameter α_0 is said to be identifiable if for every $\alpha \in \mathscr{A}$, $\alpha \neq \alpha_0$ there exists $x \in H$ such that

$$f(\alpha, x) \neq f(\alpha_0, x). \tag{3.1}$$

Note that (3.1) is, in some sense, a necessary and sufficient condition for the identifiability. Indeed, assume that $f(\alpha_1, \cdot) = f(\alpha_0, \cdot)$ for some $\alpha_1 \neq \alpha_0$. Then α_1 and α_0 cannot be distinguished using the observations $\{X_t : t > 0\}$.

For a given control $u \in \mathscr{U}$ we define a measure

$$\mu_t(\omega)(\Gamma) = \frac{1}{t} \int_0^t I_\Gamma(u_s(\omega))\, ds, \qquad \Gamma \in \mathscr{B}(B(r_0)),$$

where $\mathscr{B}(B(r_0))$ is the σ-algebra of Borel subsets of $B(r_0)$. In the sequel we omit ω and write simply $\mu_t(\Gamma)$.

For any $t \geq 1$ and $\zeta = (\zeta_1, \zeta_2) \in L^2_{loc}(\mathbb{R}_+, B(r_0)) \times L^2_{loc}(\mathbb{R}_+, H)$ we define a measure v_t^ζ by the equation

$$\int_{B(r_0) \times H} \phi(y, x)\, v_t^\zeta(dy, dx) = \frac{1}{t} \int_0^t \phi(\zeta_1(s), \zeta_2(s))\, ds, \qquad \phi \in C_b(B(r_0) \times H). \tag{3.2}$$

The existence of v_t^ζ follows from the fact that the functional defined by the right hand side of (3.2) is bounded and nonnegative on $C_b(B(r_0) \times H)$ and tends to zero if $\phi_n \downarrow 0$. Therefore, the representing measure v_t^ζ is nonnegative, $\sigma-$additive and clearly $v_t^\zeta(B(r_0) \times H) = 1$.

LEMMA 3.2. *Assume Hypotheses 2.1 and 2.3. Moreover, assume that the set of measures $\{\mu_t : t \geqslant 1\}$ is tight $\mathbb{P}$-as. Then the family of measures $\{\nu_t^{(u, X)} : t \geqslant 1\}$ is $\mathbb{P}$-a.s. tight. In particular the lemma holds if $u_s = h(\alpha_s, X_s)$, where (α_s) is an $\mathcal{A}$-valued predictable process and $h\colon \mathcal{A} \times H \to B(r_0)$ is continuous.*

Proof. For every $\zeta = (\zeta_1, \zeta_2) \in L^2_{loc}(\mathbb{R}_+, B(r_0)) \times L^2_{loc}(\mathbb{R}_+, H)$ and compact sets $K_1 \subset B(r_0)$ and $K_2 \subset H$ we have

$$\begin{aligned} &\nu_t^\zeta((B(r_0) \times H) \backslash (K_1 \times K_2)) \\ &\quad = \frac{1}{t} \int_0^t I_{(B(r_0) \times H) \backslash (K_1 \times K_2)}(\zeta_1(s), \zeta_2(s))\, ds \\ &\quad \leqslant \frac{1}{t} \int_0^t I_{B(r_0) \backslash K_1}(\zeta_1(s))\, ds + \frac{1}{t} \int_0^t I_{H \backslash K_2}(\zeta_2(s))\, ds. \end{aligned} \tag{3.3}$$

Thus by the Prokhorov Theorem it suffices to show that there exists a sequence (K_n) of compact sets in H such that

$$\lim_{n \to \infty} \sup_{t \geqslant 1} \left(\frac{1}{t} \int_0^t I_{H \backslash K_n}(X_s)\, ds \right) = 0, \qquad \mathbb{P}\text{-a.s.} \tag{3.4}$$

Step 1. We show first that

$$\limsup_{t \to \infty} \left(\frac{1}{t} \int_0^t |X_s|^2\, ds \right) < \infty, \qquad \mathbb{P}\text{-a.s.} \tag{3.5}$$

Setting

$$Z_t = \int_0^t S_{t-s} \sqrt{Q}\, dW_s$$

and

$$\tilde{Z}_t = -\int_0^t S_{t-s} u_s\, ds + Z_t, \qquad Y_t = X_t - \tilde{Z}_t,$$

we obtain

$$Y_t = S_t x + \int_0^t S_{t-s} f(\alpha_0, Y_s + \tilde{Z}_s)\, ds, \qquad t \geqslant 0. \tag{3.6}$$

Approximating Y by strong solutions to the equation

$$\frac{d}{dt} Y_t^\lambda = A Y_t^\lambda + f_\lambda(Y_t^\lambda + \tilde{Z}_t),$$

where $f_\lambda = \lambda(\lambda I - A)^{-1} f$ and using (2.4) we obtain

$$|Y_t| \leqslant e^{-\omega t} |x| + \int_0^t e^{-k(t-s)} a(|\tilde{Z}_s|)\, ds; \tag{3.7}$$

see [19] for a similar proof. Since the Ornstein–Uhlenbeck process Z^x is strongly Feller, it satisfies the Ergodic Theorem; see [26]. Furthermore, we have

$$\left| \int_0^t S_{t-s} u_s\, ds \right| \leqslant \int_0^\infty e^{-\omega(t-s)} r_0\, ds = \frac{r_0}{\omega}, \tag{3.8}$$

hence by the Young Inequality we find that

$$\begin{aligned}
\frac{1}{T} \int_0^T |X_t|^2\, dt &\leqslant \frac{2}{T} \int_0^T |Y_t|^2\, dt + \frac{2}{T} \int_0^T |\tilde{Z}_t|^2 dt \\
&\leqslant \frac{4}{T} |x|^2 \int_0^\infty e^{-2\omega t}\, dt + \frac{4}{T} \int_0^\infty e^{-2kt}\, dt \int_0^T a(|\tilde{Z}_t|)\, dt \\
&\quad + \frac{4}{T} \int_0^T |Z_t|^2\, dt + \frac{4 r_0^2}{\omega^2 T} \\
&\leqslant C_1 + \frac{C_2}{T} \int_0^T a\left(|Z_t| + \frac{r_0}{\omega} \right) dt + \frac{4}{T} \int_0^T |Z_t|^2\, dt
\end{aligned} \tag{3.9}$$

for some constants C_1, C_2 independent of T. Thereby by (2.7)

$$\begin{aligned}
&\limsup_{T \to \infty} \left(\frac{1}{T} \int_0^T |X_t|^2\, dt \right) \\
&\qquad \leqslant C_1 + C_2 \int_H a\left(|x| + \frac{r_0}{\omega} \right) \mu(dx) + 4 \int_H |x|^2 \mu(dx) < \infty
\end{aligned} \tag{3.10}$$

$\mathbb{P}$-a.s. Note that from (3.10) it follows that

$$\begin{aligned}
\limsup_{T \to \infty} \left(\frac{1}{T} \int_0^T \int_t^{t+1} |X_s|^2\, ds\, dt \right) &= \limsup_{T \to \infty} \left(\frac{1}{T} \int_0^T \int_0^T I_{[t,t+1]}(s)\, |X_s|^2\, dt\, ds \right) \\
&\leqslant \limsup_{T \to \infty} \left(\frac{1}{T} \int_0^T |X_s|^2\, ds \right) < \infty, \quad \mathbb{P}\text{-a.s.}
\end{aligned} \tag{3.11}$$

Step 2. We show that there exists a measurable set $\hat{\Omega} \subset \Omega$ such that $\mathbb{P}(\hat{\Omega}) = 1$ and for every $\omega \in \hat{\Omega}$ there exists a sequence $(C_n(\omega))$ of compact sets in H with the property

$$\lim_{n \to \infty} \sup_{t \geqslant 1} \left(\frac{1}{t} \int_0^t I_{H \setminus C_n(\omega)}(Z_s(\omega))\, ds \right) = 0. \tag{3.12}$$

Let

$$\kappa_t(\Gamma) = \frac{1}{t} \int_0^t I_\Gamma(Z_s)\, ds, \qquad \Gamma \in \mathscr{B}(H),\ t > 0. \tag{3.13}$$

The Ergodic Theorem for the Ornstein–Uhlenbeck process Z yields for $t \to \infty$

$$\frac{1}{t} \int_0^t \phi(Z_s)\, ds = \int_H \phi(y)\, \kappa_t(dy) \to \int_H \phi(y)\, \mu(dy), \qquad \mathbb{P}\text{-a.s.} \tag{3.14}$$

for each $\phi \in BUC(H)$, where $BUC(H)$ stands for a subspace of $C_b(H)$ consisting of functions which are uniformly continuous on H and endowed with the sup-norm. It is easy to see that there exists a countable subset $\mathscr{M}$ of the unit ball in $BUC(H)$ with the following property: for every ϕ in the unit ball there exists a sequence $(\phi_n) \subset \operatorname{lin}(\mathscr{M})$ such that $\phi_n \xrightarrow{\mathscr{K}} \phi$. Therefore, we can find a measurable set $\hat{\Omega} \subset \Omega$ such that $\mathbb{P}(\hat{\Omega}) = 1$ and for the sample points from $\hat{\Omega}$ and each $\phi \in \operatorname{lin}(\mathscr{M})$ (3.14) holds. It follows that $\kappa_t \to \mu$ in the narrow topology. Therefore, by the Prokhorov Theorem for each sample point in $\hat{\Omega}$ there exists a sequence (C_n) of compact sets in H satisfying (3.12).

Step 3. Fix a sample point from $\hat{\Omega}$. Then

$$X_{t+1} = S_1 x + J(f(\alpha_0, X_{\cdot + t}) + u_{\cdot + t}) + Z_{t+1} - S_1 Z_t, \qquad t \geqslant 0, \tag{3.15}$$

where J: $L^1((0, 1), H) \to H$ is given by the formula

$$Jh = \int_0^1 S_{1-s} Q^{1/2} h(s)\, ds.$$

Note that J is a compact operator and for $s, t \geqslant 0$

$$|Q^{-1/2}(f(\alpha_0, X_{s+t}) + u_{s+t})| \leqslant K(1 + |X_{s+t}|) + r_0. \tag{3.16}$$

Set

$$A_n^t = \{s \in [0, t] : |X_s| > n\},$$
$$B_n^t = \{s \in [0, t] : Z_s \notin C_n\},$$
$$C_n^t = \{s \in [0, t] : \|Q^{-1/2}(f(\alpha_0, X_{\cdot + s}) + u_{\cdot + s})\|_{L^1((0,1); H)} > n\}.$$

By (3.12) we have for $n \to \infty$

$$\sup_{t \geqslant 0} \left(\frac{1}{t} |B_n^t| \right) = \sup_{t \geqslant 0} \left(\frac{1}{t} \int_{B_n^t} I_{H \setminus C_n}(Z_s)\, ds \right)$$
$$= \sup_{t \geqslant 0} \left(\frac{1}{t} \int_0^t I_{H \setminus C_n}(Z_s)\, ds \right) \to 0, \tag{3.17}$$

where $|\cdot|$ above stands for the Lebesgue measure. It is also easy to see that (3.5) implies

$$\lim_{n \to \infty} \sup_{t \geqslant 0} \left(\frac{1}{t} |A_n^t| \right) = 0. \tag{3.18}$$

Indeed, otherwise there would exist $\varepsilon > 0$, $t_n \to \infty$ and $k_n \to \infty$ such that

$$\frac{1}{t_n} |A_{k_n}^{t_n}| > \varepsilon \tag{3.19}$$

and this inequality would imply for $n \to \infty$

$$\frac{1}{t_n} \int_0^{t_n} |X_s|\, ds \geqslant \frac{1}{t_n} \int_{A_{k_n}^{t_n}} |X_s|\, ds \geqslant \frac{k_n}{t_n} |A_{k_n}^{t_n}| \geqslant \varepsilon k_n \to \infty. \tag{3.20}$$

It follows from the proof of (3.5) that (3.5) and (3.11) hold for all sample points from $\hat{\Omega}$. Hence, (3.20) contradicts (3.5). Similarly, it can be shown that (3.11) and (3.16) imply

$$\lim_{n \to \infty} \sup_{t \geqslant 1} \left(\frac{1}{t} |C_n^t| \right) = 0. \tag{3.21}$$

Now, we define sets

$$K_n = S_1 \overline{B(n)} + J\overline{B_1(n)} + C_n + S_1 C_n, \qquad n \geqslant 1,$$

where

$$B_1(n) = \{\psi \in L^1((0, 1), H) : \|\psi\|_{L^1((0,1), H)} < n\}.$$

Clearly, each K_n is compact in H and

$$\frac{1}{t}\int_0^t I_{H\setminus K_n}(X_{s+1})\,ds \leqslant \frac{1}{t}\int_0^t (I_{H\setminus\overline{B(n)}}(X_s)$$
$$+ I_{L^1((0,1),H)\setminus\overline{B_1(n)}}(Q^{-1/2}(f(\alpha_0, X_{\cdot+s}) + u_{\cdot+s}))$$
$$+ I_{H\setminus C_n}(Z_{s+1}) + I_{H\setminus C_n}(Z_s))\,ds$$
$$\leqslant \frac{1}{t}(|A_n^t| + |C_n^t| + |B_n^{t+1}| + |B_n^t|) \to 0, \qquad n\to\infty, \tag{3.22}$$

uniformly in $t\geqslant 1$, by (3.17), (3.18), and (3.21). Since the sample point in $\hat{\Omega}$ was arbitrary, we obtain (3.4). If $u_s = h(\alpha_s, X_s)$ then it remains to show that the family of measures $\{\mu_t : t\geqslant 1\}$ is $\mathbb{P}$-a.s. tight. Take a sample point from $\hat{\Omega}$ as above and set $\hat{K} = h(\mathscr{A}, K_n)$. Then $\hat{K}_n \subset H$ is compact for each n and

$$\mu_t(B(r_0)\setminus\hat{K}_n) = \frac{1}{t}\int_0^t I_{B(r_0)\setminus\hat{K}_n}(h(\alpha_s, X_s))\,ds = \frac{1}{t}\int_0^t I_{H\setminus K_n}(X_s)\,ds,$$

hence $\{\mu_t : t\geqslant 1\}$ is relatively compact by (3.22). ∎

PROPOSITION 3.3. *Assume Hypotheses* 2.1 *and* 2.3. *Moreover, assume that* α_0 *is identifiable and the family of measures* $\{\mu_t : t\geqslant 1\}$ *is* $\mathbb{P}$*-a.s. tight. Then*

$$\liminf_{t\to\infty} \frac{1}{t}\int_0^t |Q^{-1/2}(f(\alpha, X_s) - f(\alpha_0, X_s))|^2\,ds > 0 \qquad \mathbb{P}\text{-}a.s. \tag{3.23}$$

for each $\alpha \neq \alpha_0$.

Proof. Let

$$\phi(x) = g_m(\langle x, k_1\rangle, ..., \langle x, k_n\rangle) \tag{3.24}$$

for a certain $g_m \in \mathscr{G}$ and $k_1, ..., k_n \in \mathscr{K}$. By the Ito formula we obtain for $t\geqslant 0$

$$\frac{1}{t}\left(\phi(X_t) - \phi(x) - \int_0^t \langle Q^{1/2} D\phi(X_s), dW_s\rangle\right)$$
$$= \frac{1}{t}\int_0^t (L^{f_0}\phi(X_s) - \langle u_s, D\phi(X_s)\rangle)\,ds \qquad \mathbb{P}\text{-a.s.}, \tag{3.25}$$

where f_0 above stands for $f(\alpha_0, \cdot)$. Therefore,

$$\frac{1}{t}\left(\phi(X_t) - \phi(x) - \int_0^t \langle Q^{1/2}\, D\phi(X_s), dW_s\rangle\right)$$

$$= \int_{B(r_0)\times H} \psi(y, x)\, \nu_t^{(u, X)}(dy, dx), \qquad \mathbb{P}\text{-a.s.}, \tag{3.26}$$

where

$$\psi(y, x) = L^{f_0}\phi(x) - \langle y, D\phi(x)\rangle. \tag{3.27}$$

The function ψ is continuous on $B(r_0)\times H$ and

$$|\psi(y, x)| \leqslant k(1 + |x|), \qquad x \in H, \tag{3.28}$$

for a certain $k > 0$. Using the Strong Law of Large Numbers for Martingales we find that the left hand side of (3.26) tends to zero $\mathbb{P}$-a.s. as $t \to \infty$ and thereby

$$\lim_{t\to\infty} \int_{B(r_0)\times H} \psi(y, x)\, \nu_t^{(u, X)}(dy, dx) = 0, \qquad \mathbb{P}\text{-a.s.}. \tag{3.29}$$

We can find a measurable $\Omega_1 \subset \Omega$ such that $\mathbb{P}(\Omega_1) = 1$ and (3.29) holds for every $\omega \in \Omega_1$ and any ϕ defined by (3.24). Take a sample path $(u(\omega), X(\omega)) = (u_.(\omega), X_.(\omega))$ for a fixed $\omega \in \Omega_1$. By Lemma 3.2 the family of measures $\{\nu_t^{(u(\omega), X(\omega))} : t \geqslant 1\}$ is tight. Let $\nu^{(u(\omega), X(\omega))}$ denote any limit point of this set. The measure $\nu^{(u(\omega), X(\omega))}$ may be disintegrated to the form

$$\nu^{(u(\omega), X(\omega))} = \nu_1(dy \mid x)\, \nu_2(dx),$$

where the dependence on ω on the right hand side is suppressed.

Our next aim is to show that

$$\overline{\operatorname{supp}(\nu_2)} = H. \tag{3.30}$$

Since H is a Polish space the function $x \to \nu_1(U|x)$ is Borel-measurable for every Borel set $U \subset H$ and we can define a bounded Borel function

$$v(x) = \int_{B(r_0)} y\nu_1(dy \mid x).$$

By (3.29)

$$\int_{B(r_0)\times H} (L^{f_0}\phi(x) - \langle y, D\phi(x)\rangle)\, \nu^{(u(\omega), X(\omega))}(dy, dx)$$

$$= \int_H (L^{f_0}\phi(x) - \langle v(x), D\phi(x)\rangle)\, \nu_2(dx) = 0, \tag{3.31}$$

for every ϕ given by (3.24). Therefore, by Proposition 2.5 (3.31) holds for every $\phi \in \mathscr{D}_0$, hence by Lemma 3.1 ν_2 is an invariant measure for the equation

$$dX_t = (AX_t + f(\alpha_0, X_t) + v(X_t))\, dt + \sqrt{Q}\, dW_t. \tag{3.32}$$

By the Girsanov Theorem the measure ν_2 is equivalent to μ and $\overline{\text{supp}(\mu)} = H$, hence (3.30) follows. Note that the function

$$x \to |Q^{-1/2}(f(\alpha, x) - f(\alpha_0, x))|^2$$

is continuous on H, has at most quadratic growth and for $\alpha \neq \alpha_0$ is strictly positive on a nonempty open set. For any $\omega \in \Omega_1$ and any sequence $t_n \to \infty$ we can find a subsequence (denoted again by (t_n)) and a limit point $\nu^{(u, X)}$ of the set $\{\nu_{t_n} : n \geqslant 1\}$ such that

$$\lim_{n\to\infty} \frac{1}{t_n} \int_0^{t_n} |Q^{-1/2}(f(\alpha, X_s) - f(\alpha_0, X_s))|^2\, ds$$

$$= \int_{B(r_0)\times H} |Q^{-1/2}(f(\alpha, x) - f(\alpha_0, x))|^2\, \nu^{(u, X)}(dy, dx),$$

where $\nu^{(u, X)}(dy, dx) = \nu_1^x(dy)\, \nu_2(dx)$. Hence by (3.30)

$$\lim_{n\to\infty} \frac{1}{t_n} \int_0^{t_n} |Q^{-1/2}(f(\alpha, X_s) - f(\alpha_0, X_s))|^2\, ds$$

$$= \int_H |Q^{-1/2}(f(\alpha, x) - f(\alpha_0, x))|^2\, \nu_2(dx) > 0.$$

Since $\omega \in \Omega_1$ was arbitrary and $\mathbb{P}(\Omega_1) = 1$, we obtain (3.23). ∎

4. CONSISTENCY OF MLE ESTIMATOR

Following the finite dimensional case (cf. [7, 8]) we can use Proposition 1 to establish strong consistency of the family of maximum likelihood estimators of α. Let

$$l_t^\alpha = e^{m_t^\alpha - 1/2 a_t^\alpha}, \tag{4.1}$$

where

$$m_t^\alpha = \int_0^t \langle Q^{-1/2}(f(\alpha, X(t)) - f(\alpha_0, X(s))), dW(s) \rangle, \tag{4.2}$$

and

$$a_t^\alpha = \int_0^t |Q^{-1/2}(f(\alpha, X(t)) - f(\alpha_0, X(s)))|^2 \, ds. \tag{4.3}$$

We define the estimator α_t as a maximiser of the function $\alpha \to l_t^\alpha$:

$$\alpha_t = \operatorname{argmax}_{\alpha \in \mathscr{A}} l_t^\alpha. \tag{4.4}$$

The fact that α_t can be realized as an adapted, measurable process follows from Lemma 4.2 below. We will need in this section an additional

HYPOTHESIS 4.1. (a) *$\mathscr{A} \subset \mathbb{R}^d$ is a compact set.*

(b) *There exist $p > 0$, $k < \infty$ and $\theta \in (0, 1]$ such that*

$$|f(\alpha_1, x) - f(\alpha_2, x)| \leqslant k \, |\alpha_1 - \alpha_2|^\theta \, (1 + |x|^p), \qquad \alpha_1, \alpha_2 \in \mathscr{A}, \; x \in H.$$

(c) *For every $r > 0$*

$$\int_H a^r \left(\frac{r_0}{\omega} + |y| \right) \mu(dy) < \infty.$$

LEMMA 4.2. *Assume Hypotheses* 2.1, 4.1. *Then the map $(t, \alpha) \to m_t^\alpha$ has a continuous modification on $(0, \infty) \times \mathscr{A}$, the continuity in α being uniform with respect to $t \in [\delta, \infty)$ for each $\delta > 0$.*

Proof. Obviously, $t \to m_t^\alpha$ is a.s. continuous for every $\alpha \in \mathscr{A}$. thus it suffices to prove a.s. continuity of the function $\alpha \to \frac{1}{t} m_t^\alpha$, uniformly in $t \in [\delta, \infty)$. We will prove first that for every $q \geqslant 0$

$$\sup_{t \geqslant 0} \mathbb{E} \, |X_t|^{2q} < \infty. \tag{4.5}$$

This clearly true if $f=0$. For $f\neq 0$ we will use the notation $X_t=Y_t+\tilde{Z}_t$ as in the previous section. Then, similarly as in Lemma 3.2 assumption (iii) of Hypothesis 2.1 yields for $t\geqslant 0$

$$\mathbb{E}\,|Y_t|^{2q}\leqslant c_1\,|x|^{2q}\,e^{-2kqt}+c_2\int_0^t e^{-2kq(t-s)}\mathbb{E}\,|Y_s|^{2q-1}\,a\left(\frac{r_0}{\omega}+|Z_s|\right)ds \tag{4.6}$$

for any $q\geqslant\frac{1}{2}$ and suitable constants c_1, c_2. for $q=\frac{1}{2}$ we get immediately

$$\sup_{t\geqslant 0}\,\mathbb{E}\,|Y_t|<\infty, \tag{4.7}$$

since

$$\lim_{s\to\infty}\,\mathbb{E}a\left(\frac{r_0}{\omega}+|Z_s|\right)=\int_H a\left(\frac{r_0}{\omega}+|y|\right)\mu(dy)<\infty.$$

For $q>\frac{1}{2}$ we can use induction with the step $\frac{1}{2}$. By the Hölder inequality applied to (4.6) we obtain

$$\begin{aligned}\mathbb{E}\,|Y_t|^{2q}\leqslant{}& c_1\,|x|^{2q}\,e^{-2kqt}\\&+c_2\int_0^t e^{-2kq(t-s)}(\mathbb{E}\,|Y_s|^{2q-1/2})^{1/b}\left(\mathbb{E}a^{b'}\left(\frac{r_0}{\omega}+|Z_s|\right)\right)1/b'\,ds,\end{aligned} \tag{4.8}$$

where $b=\frac{4q-1}{2(2q-1)}$ and $b'=\frac{b}{b-1}$. Since $\mathbb{E}a^b(|Z_s|)$ is bounded by Hypothesis 4.1 we obtain

$$\sup_{t\geqslant 0}\,\mathbb{E}\,|X_t|^{2q}\leqslant c_3(\sup_{t\geqslant 0}\,\mathbb{E}\,|Y_t|^{2q}+\sup_{t\geqslant 0}\,\mathbb{E}\,|\tilde{Z}_t|^{2q})<\infty, \tag{4.9}$$

for a certain $c_3>0$. Now we have for $r\geqslant 1$ by Lemma 4.12 of [25]

$$\begin{aligned}\mathbb{E}\left|\frac{1}{t}\,m_t^{\alpha_1}-\frac{1}{t}\,m_t^{\alpha_2}\right|^{2r}&\leqslant\frac{1}{t^{2r}}\,\mathbb{E}\left|\int_0^t\langle Q^{-1/2}(f(\alpha_1,X_s)-f(\alpha_2,X_s)),dW_s\rangle\right|^{2r}\\&\leqslant\frac{(2r-1)^r\,r^r t^{r-1}}{t^{2r}}\int_0^t\mathbb{E}\,|Q^{-1/2}(f(\alpha_1,X_s)-f(\alpha_2,X_s))|^{2r}\,ds\\&\leqslant C_1 t^{-r-1}\int_0^t|\alpha_1-\alpha_2|^{2\theta r}\,\mathbb{E}(1+|X_s|^p)^{2r}\,ds\\&\leqslant C_2 t^{-r}\,|\alpha_1-\alpha_2|^{2\theta r},\end{aligned} \tag{4.10}$$

for some $C_1, C_2 > 0$. Taking r large enough so that $2\theta r > d$, and using the Kolmogorov–Totoki criterion for the existence of versions with continuous sample paths (see, e.g., [18] we conclude the proof. ▮

THEOREM 4.3. *Let Hypotheses* 2.1 *and* 4.1 *hold. Assume that* (3.1) *is satisfied and the control* (u_t) *satisfies the conditions from Proposition* 3.3. *Then the family of estimators* (α_t) *defined by* (4.4) *is strongly consistent, that is*

$$\mathbb{P}(\lim_{t\to\infty} \alpha_t = \alpha_0) = 1.$$

Proof. Since the process $t \to m_t^\alpha$ is a martingale we can find a Wiener process β^α such that $m_t^\alpha = \beta^\alpha(a_t^\alpha)$ for $t \geqslant 0$. By the Strong Law Large Numbers $\frac{1}{t}\beta^\alpha(t) \to 0$ $\mathbb{P}$-a.s. for $t \to \infty$. Furthermore, by (2.3) we have

$$\mathbb{P}\left(\limsup_{t\to\infty} \frac{1}{t}\, a_t^\alpha < \infty\right) = 1, \qquad \alpha \in \mathscr{A}.$$

It follows that for every $\alpha \in \mathscr{A}$

$$\lim_{t\to\infty} \frac{1}{t}\, m_t^\alpha = \lim_{t\to\infty} \frac{\beta^\alpha(a_t^\alpha)}{a_t^\alpha}\, \frac{a_t^\alpha}{t} = 0, \qquad P\text{-a.s.} \tag{4.11}$$

By Lemma 4.2 convergence in (4.11) is uniform with respect to $\alpha \in \mathscr{A}$, hence

$$\frac{1}{t}\, m_t^\alpha \to 0, \qquad \mathbb{P}\text{-a.s.}. \tag{4.12}$$

Then the definition of l_t^α and α_t yields

$$m_t^{\alpha_t} - \frac{1}{2}\, a_t^{\alpha_t} \geqslant m_t^{\alpha_0} - \frac{1}{2}\, a_t^{\alpha_0}$$

and therefore -a.s.

$$\lim_{t\to\infty} \frac{1}{t}\, a_t^{\alpha_t} = \lim_{t\to\infty} \frac{1}{t} \int_0^t |Q^{-1/2}(f(\alpha_s, X_s) - f(\alpha_0, X_s))|^2\, ds = 0,$$

which together with Proposition 3.3 completes the proof of the theorem. ▮

Remark 4.4. Note that our result together with the result from [17] proves the self-optimizing property of the adaptive control defined in [17] for the case considered there. Indeed, assume that the conditions from [17] (that is, (A1)–(A5) and (3.9)) and, in addition, our Hypothesis 4.1 are satisfied. Moreover, assume (3.1). The adaptive control can be defined in the feedback form

$$u_t = DH(Dv_{\alpha_t}(X_t)), \tag{4.13}$$

where H is the Hamiltonian of the problem, $(v_\alpha, \lambda_\alpha)$ is the solution of the ergodic Hamilton–Jacobi–Bellman equation corresponding to the parameter α (cf. [17, 21] for details) and α_t is the MLE defined above. Our Hypotheses 2.1, 2.3, and 4.1 are satisfied in that case and from Proposition 3.4 and Corollary 3.1 of [17] it follows that the function $(\alpha, x) \to DH(Dv_\alpha(x))$ is continuous on $\mathscr{A} \times H$. Therefore, the control u_t (4.13) satisfies conditions of Proposition (3.3) and we obtain $\alpha_t \to \alpha_0$ $\mathbb{P}$-a.s. by Theorem 4.3. By Proposition 3.5 of [17] it follows that the adaptive control (u_t) is self-optimizing.

4.1. Example. We provide below a simple example of a nonlinear stochastic partial differential equations for which the theory of maximum likelihood estimators developed above is valid.

Let A be realization in $H = L^2(0, 1)$ of a differential operator

$$A_0 x(\zeta) = \frac{\partial}{\partial \zeta}\left(a \frac{\partial x}{\partial \zeta}\right)(\zeta) + b(\zeta) \frac{\partial x}{\partial \zeta}(\zeta) + c_0(\zeta)\, x(\zeta), \qquad \zeta \in [0, 1],$$

endowed with the Dirichlet boundary conditions. We assume that a is Lipschitz on [0,1] and $b, c_0 \in L^\infty(0, 1)$. Moreover, we assume that

$$a^2(\zeta) \geqslant m > 0, \qquad \zeta \in [0, 1].$$

Under these assumptions A generates an analytic C_0-semigroup of contractions on $L^2(0, 1)$, see for example [27] for the results on generation of semigroups by differential operators. Under the present assumptions it also known, see [1], that A generates a semigroup $S(t)$ of Hilbert–Schmidt operators on $L^2(0, 1)$ and for all $T > 0$

$$\int_0^T \|S(t)\|_{HS}^2 \, ds < \infty.$$

Let $\mathscr{A} \subset \mathbb{R}^d$ be a compact set and let

$$f_0 \colon \mathscr{A} \times \mathbb{R} \to \mathbb{R}$$

be such a mapping that f_0 is continuously differentiable in the second argument and

$$\sup_{\zeta, \alpha} \left| \frac{\partial f_0}{\partial \zeta} (\alpha, \zeta) \right| < \infty.$$

We will assume also that

$$|f_0(\alpha_1, \zeta) - f_0(\alpha_2, \zeta)| \leqslant k \, |\alpha_1 - \alpha_2|^\theta \, (1 + |\zeta|), \qquad \alpha_1, \alpha_2 \in \mathscr{A}, \zeta \in [0, 1],$$

and

$$\zeta f_0(\alpha, \zeta) \leqslant -k_1 \zeta^2 + k_2, \qquad \alpha \in \mathscr{A}, \zeta_1, \zeta_2 \in [0, 1],$$

where $k_1 + m\pi^2 > \|c_0\|_\infty$. Then we define the Nemytskii operator f: $L^2(0, 1) \to L^2(0, 1)$:

$$f(x)(\zeta) = f_0(x(\zeta)), \qquad \zeta \in [0, 1].$$

We assume also that the random field $\{u_t(\zeta) : t \geqslant 0, \zeta \in [0, 1]\}$ satisfies the following conditions: the process

$$[0, \infty) \ni t \to u_t(\cdot) \in L^2(0, 1) \text{ is progressively measurable}$$

and

$$\sup_{t \geqslant 0} \int_0^1 |u_t(\zeta)|^2 \, d\zeta \leqslant r_0^2.$$

It is easy to check that Hypotheses 2.1, 2.3 and 4.1 are satisfied in the present case and therefore the stochastic partial differential equation

$$\begin{cases} \dfrac{\partial X}{\partial t} (t, \zeta) = \dfrac{\partial}{\partial \zeta} \left(a \dfrac{\partial X}{\partial \zeta} \right) (t, \zeta) + b(\zeta) \dfrac{\partial X}{\partial \zeta} (t, \zeta) \\ \qquad + c_0(\zeta) \, X(t, \zeta) + f_0(\alpha_0, X(t, \zeta)) - u_t(\zeta) + \dfrac{\partial^2 W}{\partial t \, \partial \zeta}, \\ X(0, \zeta) = x(\zeta), \zeta \in [0, 1], X(t, 0) = X(t, 1) = 0, \end{cases}$$

has a unique solution in the sense defined in Section 2. In the above equation W stands for the cylindrical Wiener process on $L^2(0, 1)$, that is,

$$\mathbb{E} W(t_1, \zeta_1) \, W(t_2, \zeta_2) = \min(t_1, t_2) \min(\zeta_1, \zeta_2), \qquad t_i \geqslant 0, \zeta_i \in [0, 1].$$

As a consequence we find the strong consistency of the maximum likelihood estimator α_t of the parameter α_0 provided the family of measures $\{\mu_t: t \geqslant 1\}$ is relatively compact on $L^2(0, 1)$. In particular, this condition is satisfied if

$$u_t(\zeta) = K(\alpha_t, X(t, \zeta)),$$

where the function $K: \mathscr{A} \times \mathbb{R} \to \mathbb{R}$ is uniformly bounded and continuous.

REFERENCES

1. S. Agmon, On eigenfunctions and on the eigenvalues of general elliptic boundary value problem, *Comm. Pure Appl. Math.* **15** (1962), 119–147.
2. S. I. Aihara, Regularized maximum likelihood estimate for infinite-dimensional parameter in stochastic parabolic systems, *SIAM J. Control Optim.* **30** (1992), 745–764.
3. S. I. Aihara and A. Bagchi, Parameter identification for stochastic diffusion equations with unknown boundary conditions, *Appl. Math. Optim.* **17** (1988), 15–36.
4. S. I. Aihara and Y. Sunahara, Identification of an infinite-dimensional parameter for stochastic diffusion equations, *SIAM J. Control Optim.* **26** (1988), 1062–1075.
5. S. I. Aihara and A. Bagchi, Parameter identification for hyperbolic stochastic systems, *J. Math. Anal. Appl.* **166** (1991), 485–499.
6. S. I. Aihara, Identification of a discontinuous parameter in stochastic parabolic systems, *Appl. Math. Optim.* **37** (1998), 43–69.
7. V. S. Borkar, "Optimal Control of Diffusion Processes," Longman, Harlow/New York, 1989.
8. V. S. Borkar and M. K. Ghosh, Ergodic control of multidimensional diffusions. II. Adaptive control, *Appl. Math. Optim.* **21** (1990), 191–220.
9. V. S. Borkar and S. M. Mundra, Bayesian parameter estimation and adaptive control of Markov processes with time-averaged cost, *Appl. Math. (Warsaw)* **25** (1998), 339–358.
10. S. Cerrai and F. Gozzi, Strong solutions of Cauchy problems associated to weakly continuous semigroups, *Differential Integral Equations* **8** (1995), 465–486.
11. A. Chojnowska-Michalik and B. Goldys, Existence, uniqueness and invariant measures for stochastic semilinear equations on Hilbert spaces, *Probab. Theory Related Fields* **102** (1995), 331–356.
12. A. Chojnowska-Michalik and B. Goldys, "On Generators of Infinite-Dimensional Diffusions," Report S96-9, School of Mathematics, The University of New South Wales, 1996.
13. G. Da Prato and B. Goldys, Invariant measures of non symmetric dissipative stochastic systems, *Scuola Norm. Sup. Pisa* **1** (1999).
14. G. Da Prato and J. Zabczyk, "Stochastic Equations in Infinite Dimensions," Cambridge Univ. Press, Cambridge, UK, 1992.
15. T. E. Duncan, B. Pasik-Duncan, and B. Goldys, Adaptive control of linear stochastic evolution systems, *Stochastics Stochastics Rep.* **35** (1991), 129–142.
16. T. E. Duncan, B. Maslowski, and B. Pasik-Duncan, Adaptive boundary and point control of linear stochastic distributed parameter systems, *SIAM J. Control Optim.* **32** (1994), 648–672.
17. T. E. Duncan, B. Maslowski, and B. Pasik-Duncan, Adaptive control for semilinear stochastic systems, submitted for publication.
18. K. D. Elworthy, "Stochastic Differential Equations on Manifolds," London Math. Soc. Lecture Notes, Vol. 70, Cambridge Univ. Press, Cambridge, UK, 1982.

19. D. Gątarek and B. Goldys, On invariant measures for diffusions on Banach spaces, *Potential Anal.* **7** (1996), 539–553.
20. B. Goldys and M. Kocan, Diffusion semigroups in spaces of continuous functions with mixed topology, *J. Differential Equations* **173** (2001), 17–39.
21. B. Goldys and B. Maslowski, Ergodic control of stochastic semilinear equations and Hamilton–Jacobi equations, *J. Math. Anal. Appl.* **234** (1999), 592–631.
22. M. Huebner, R. Khasminskii, and B. L. Rozovskii, Two examples of parameter estimation for stochastic partial differential equations, *in* "Stochastic Processes: A Festschrift in Honour of Gopinath Kallianpur," pp. 149–160, Springer-Verlag, New York, 1993.
23. M. Huebner and B. L. Rozovskii, On asymptotic properties of maximum likelihood estimators for parabolic stochastic PDE's, *Probab. Theory Related Fields* **103** (1995), 143–163.
24. T. Koski and W. Loges, On minimum contrast estimation for Hilbert space-valued stochastic differential equations, *Stochastics* **16** (1986), 217–225.
25. R. S. Liptser and A. N. Shiryaev, "Statistics of Random Processes. I. General Theory," Springer-Verlag, New York/Berlin, 1977.
26. B. Maslowski, On ergodic behaviour of solutions to systems of reaction-diffusion equations with correlated noise, *in* "Stochastic Processes Related Topics," Vol. 61, pp. 93–102, Akademie-Verlag, Berlin, 1991.
27. A. Pazy, "Semigroups of Linear Operators and Applications to Partial Differential Equations," Springer-Verlag, New York/Berlin, 1983.
28. B. L. S. Rao Prakasa, "Statistical Inference for Diffusion Type Processes," Arnold Sevenoaks, 1999.
29. B. L. S. Rao Prakasa, Asymptotic theory of estimation in nonlinear stochastic differential equations, *Sankhyā Ser. A* **43** (1981), 170–183.
30. B. L. S. Rao Prakasa, Asymptotic theory for nonlinear least squares estimator for diffusion processes, *Statistics* **14** (1983), 195–209.
31. B. L. S. Rao Prakasa, Bayes estimation for some stochastic partial differential equations, *J. Statist. Plann. Inference*, in press.
32. J. Voigt, On the perturbation theory for strongly continuous semigroups, *Math. Ann.* **229** (1997), 163–171.