Convergence Rates of Continuous-Time Stochastic ELS Parameter Estimation

HAN-FU CHEN AND JOHN B. MOORE

Abstract—**Discrete-time convergence rates for extended least squares (ELS) algorithms are generalized to the continuous-time case. An essential difference in the estimation is the appropriate prefiltering, while in the theory, the existence of solutions of the stochastic equations is a concern.**

I. INTRODUCTION

Least squares (LS) estimation of continuous-time stochastic signal models with additive "white" noise is considered in [1]-[3]. Ergodicity assumptions are involved in the theory of [1]. In [2], [3] ergodicity is not assumed and convergence rates are given. In [1] there is a conjecture that corresponding convergence rates for extended least-squares algorithms in the colored noise case cannot be obtained. Some of the foundations are laid for such results in [2], [3]. In [4], ELS estimation is prescribed for signal models with appropriate prefiltering, including possibly the domination of certain noise signals by additive noise. Also, a weighting coefficient selection scheme is built into the ELS estimation to improve convergence properties and avoid finite escape times. The theory of [1]-[4] relies on Martingale convergence theorems.

Here, we generalize the work of [6] which gives rates of convergence for the discrete-time stochastic colored noise case to the continuous-time framework. A key ingredient is the prefiltering in the estimation, and for the theory, a key issue is the possibility of the existence of finite escape times on sample paths.

II. PROBLEM STATEMENT

Consider the dynamic system described by the following multivariable stochastic integral equation:

$$A(S)y_t = B(S)u_t + C(S)v_t \tag{2.1}$$

where $A(S)$, $B(S)$, and $C(S)$ are matrix polynomials in the integral operator S as

$$A(S) = I + A_1S + \cdots + A_pS^p,\ B(S) = B_1S + B_2S^2 + \cdots + B_qS^q,\ C(S) = I + C_1S + \cdots + C_rS^r.$$

Without loss of generality, $C(S)$ can be "minimum phase" in that $C(s)$ is full rank in Re $s > 0$ with s the Laplace transform variable. For subsequent theory, we impose the stronger condition that

$C(S)$ is strictly minimum phase, in that

$$C(s) \text{ is full rank in Re } s \geq 0. \tag{2.2}$$

Note that $Sx_t = \int_0^t x_\lambda\, d\lambda$, $S^2x_t = \int_0^t \int_0^\tau x_\lambda\, d\lambda\, d\tau$. Here v_t satisfies

$$W(S)v_t = w_t \tag{2.3}$$

where (w_t, F_t) is a Wiener process, with F_t a family of nondecreasing σ-algebras, and $W(S) = I + W_1S + \cdots + W_rS^r$ such that

$$C(S)W^{-1}(S) - \frac{1}{2}I \text{ is strictly positive real.} \tag{2.4}$$

In [4], it is shown how such signal models can arise from plants, with the following transfer function description under zero initial conditions

$$y^p(s) = [A^p(s)]^{-1}B^p(s)u^p(s) + [A^p(s)]^{-1}C^p(s)w^p(s)$$

where y_t^p are the outputs, u_t^p the inputs, w_t^p the noise disturbances, and $A^p(s)$ $B^p(s)$, $C^p(s)$ are polynomials in s. By filtering y_t^p, u_t^p through asymptotically stable filters with transfer function $W^{-1}(s^{-1})$ to yield y_t, u_t then the model for this system can be expressed in the form (2.1)-(2.3). Also in [4], it is shown that by adding appropriate disturbances at the plant output, the model has the form (2.1)-(2.3) with (2.4) satisfied. Details are not repeated here.

Let us consider the estimation of the parameters A_i, B_i, C_i on the basis of past measurements, $\{y_\tau, u_\tau;\ \tau \leqq t\}$, where

u_t is F_t-measurable and locally bounded in L_2. (That is

$$\int_0^t \|u_\lambda\|^2\, d\tau < \infty \quad \text{a.s. for all } t < \infty.) \tag{2.5}$$

First, let us denote

$$\theta^T = [-A_1 \cdots -A_p\ \ B_1 \cdots\ \ B_q\ C_1 \cdots C_r]$$

$$(\phi_t^0)^T = [y_t^T\ Sy_t^T \cdots S^{p-1}y_t^T\ \ u_t^T \cdots S^{q-1}u_t^T\ v_t^T \cdots S^{r-1}v_t^T] \tag{2.6}$$

so that the model (2.1) can be written as

$$y_t = S\theta^T\phi_t^0 + v_t. \tag{2.7}$$

The ELS based estimation of θ, yielding estimates θ_t involves the stochastic differential equation

$$d\theta_t = \bar{P}_t\phi_t W(S)(dy_t - \theta_t^T\phi_t dt)^T, \qquad \theta_0 \text{ arbitrary}$$

$$\bar{P}_t = \left(\int_0^t \phi_\tau\phi_\tau^T\, d\tau + a^{-1}I\right)^{-1} \qquad a \triangleq \text{dimension of } \phi_t \tag{2.8}$$

where

$$\phi_t^T = [y_t^T\ Sy_t^T \cdots S^{p-1}y_t^T\ \ u_t^T \cdots S^{q-1}u_t^T\ \ \hat{v}_t^T \cdots S^{r-1}\hat{v}_t^T]$$

$$\hat{v}_t = y_t - S(\theta_t^T\phi_t). \tag{2.9}$$

Notice the presence of $W(S)$ and that tr $\bar{P}_0^{-1} = 1$.

Assumption 1: The solution of (2.8) exists for almost all sample paths $\omega \in \Omega$ up to an escape time. (2.10)

Let us define for each ω

$$\sigma = \sup\left\{t : \int_0^t \phi_\lambda^T\bar{P}_\lambda\phi_\lambda\, d\lambda < \infty\right\}. \tag{2.11}$$

When σ is finite, it denotes a finite escape time of the process θ_t. With the above definitions and assumption, we now claim that with $I_{[t<\sigma]}$ the ω-dependent indicator function which is unity for $t < \sigma$ and zero otherwise, then

$$I_{[t<\sigma]} \in F_t \tag{2.12}$$

$$\sigma \text{ is a Markov time} \tag{2.13}$$

$$\int_0^t \|I_{[\lambda<\sigma]}\bar{P}_\lambda\phi_\lambda\|^2\, d\lambda < \infty \qquad \text{for all finite } t. \tag{2.14}$$

The result (2.12) follows since

$$[t<\sigma] \triangleq \left\{\omega : \int_0^t \phi_\lambda^T\bar{P}_\lambda\phi_\lambda\, d\lambda < \infty\right\} \in F_t.$$

The result (2.13) follows since $[t < \sigma] = [\sigma \leqq t]^c$ where c denotes the complement so that $[t \geqq \sigma] \in F_t$, or equivalently σ is a Markov time [5]. The result (2.14) follows from the property (2.11), noting that $\bar{P}_\lambda^2 \leqq$

Manuscript received August 11, 1986.
H.-F. Chen is with the Institute of Systems Science, Academia Sinica, Beijing, China.
J. B. Moore is with the Department of Systems Engineering, Research School of Physical Sciences, Australian National University, Canberra, A.C.T., Australia.
IEEE Log Number 8612539.

0018-9286/87/0300-0267$01.00 © 1987 IEEE

$\bar{P}_\lambda^{1/2}\bar{P}_0\bar{P}_\lambda^{1/2} \leqq (\operatorname{tr} \bar{P}_0)\bar{P}_\lambda$. It now proves convenient to consider the following modified form of (2.8):

$$d\theta_t = I_{[t<\sigma]}P_t\phi_t W(S)(dy_t - \theta_t^T\phi_t dt)^T, \qquad \theta_0 \text{ arbitrary}$$

$$P_t = \left(\int_0^t I_{[t<\sigma]}\phi_\tau\phi_\tau^T d\tau + a^{-1}I\right)^{-1}, \; a = \text{dimension of } \phi_t. \qquad (2.8)'$$

Remarks:

1) Should ϕ_t in (2.8)' be independent of θ_t, then (2.11)–(2.14) hold and we could conclude that θ_t exists for all t as the unique strong solution of

$$\theta_t = \theta_0 + \int_0^t I_{[\lambda<\sigma]}P_\lambda\phi_\lambda W(S)(dy_\lambda - \theta_\lambda^T\phi_\lambda\, d\lambda)^T. \qquad (2.15)$$

Since here for $C(S) \neq I$, ϕ_t is θ_t dependent, (2.8) is highly nonlinear, and we cannot conclude the same properties for θ_t without a formidable analysis of the nature of the nonlinearities. Assumption (2.10) in essence is that θ_t of (2.8)' exists for all t as the unique strong solution of (2.15).

2) In the case $C(S) = I$, the "white" noise case, $\phi_t = \phi_t^0$ is independent of v_t, $\hat{v}_t$ and hence of θ_t. Moreover, under (2.5) from (2.10), $\sigma = \infty$ and there is no finite escape time almost surely. Then $I_{[t<\sigma]} \equiv 1$ and (2.8) is a linear stochastic differential equation with a unique strong solution for all t.

3) For the case $W(S) = I$, the condition (2.4) can only be satisfied with $C(S) = I$ (in contrast to the discrete-time case). See also [2].

III. Main Results

Lemma 3.1: Consider the plant (signal model) (2.1)–(2.3) [or (2.6), (2.7)], under conditions (2.4), (2.5). Consider also the ELS estimation scheme (2.8)–(2.10). Then there are constants $\epsilon > 0$ and $k_1 > 0$ such that

$$\int_0^t [f_\lambda^T g_\lambda - \epsilon(g_\lambda^T g_\lambda + f_\lambda^T f_\lambda)]\, d\lambda + k_1 \geqq 0 \qquad \text{for } t<\sigma \qquad (3.1)$$

where, denoting $\tilde{\theta}_t = \theta - \theta_t$, $\tilde{v}_t = v_t - \hat{v}_t$

$$g_t = \tilde{\theta}_t^T\phi_t, \; Sf_t = [C(S) - W(S)]\tilde{v}_t + \frac{1}{2}Sg_t. \qquad (3.2)$$

Moreover, under condition (2.2), for some k_2, k_3, with $\tilde{\phi}_t = \phi_t^0 - \phi_t$

$$\int_0^t \|\tilde{\phi}_\lambda\|^2\, d\lambda \leqq k_1 \int_0^t \|g_\lambda\|^2\, d\lambda + k_3. \qquad (3.3)$$

Proof: Simple manipulations yield for $t < \sigma$

$$d\tilde{v}_t = -\theta^T\tilde{\phi}_t dt - g_t dt, \text{ or } C(S)\tilde{v}_t = -Sg_t$$

$$d\tilde{\phi}_t = F\tilde{\phi}_t dt - Gg_t dt \qquad (3.4)$$

where $\tilde{\phi}_t^T = [0 \cdots 0\tilde{v}_t^T \cdots S^{r-1}\tilde{v}_t^T]$ and

$$F = \operatorname{diag}\left\{0, \left|\begin{matrix} -c_1 & -c_2 & \cdots & -c_r \\ & I & & 0 \end{matrix}\right|\right\}, \; G^T = [0 \cdots 0 \; I \; \cdots 0].$$

Also,

$$f_t = -[C(S) - W(S)]C^{-1}(S)g_t + \frac{1}{2}g_t$$

$$= \left[W(S)C^{-1}(S) - \frac{1}{2}I\right]g_t. \qquad (3.5)$$

Now (3.1) follows from (3.5) and application of the strict positive real condition (2.4). Also from (3.4) we see that ϕ_t are the states of a linear system with characteristic equation $\det C(s^{-1}) = 0$, so that result (3.3) follows under (2.2).

Lemma 3.2: With the definition of P_t in (2.8)' and σ in (2.11),

$$\int_0^t I_{[\lambda<\sigma]}\phi_\lambda^T P_\lambda\phi_\lambda\, d\lambda = \ln(\det P_t^{-1}) + a \ln a \qquad (3.6)$$

Proof: From (2.8)

$$\det P_{t+dt}^{-1} = \det P_t^{-1} \det (I + I_{[t<\sigma]}P_t\phi_t\phi_t^T dt)$$

$$= \det P_t^{-1}(1 + I_{[t<\sigma]}\phi_t^T P_t\phi_t dt)$$

so that

$$\frac{d(\det P_t^{-1})}{\det P_t^{-1}} = I_{[t<\sigma]}\phi_t^T P_t\phi_t dt$$

and the result (3.6) follows.

Lemma 3.3: Let the measurable process M_t be adapted to F_t and define

$$T = \sup\left\{t : q_t \triangleq \int_0^t \|M_\lambda\|^2\, d\lambda < \infty\right\}.$$

Then for all $\eta > 0$, as $t \to T$

$$\int_0^t M_\lambda\, dw_\lambda = 0[q_t^{1/2} \ln^{1/2+\eta}(q_t + e)] \text{ a.s.} \qquad (3.7)$$

Proof: As for σ in (2.13), T is a Markov time and $I_{[t<T]} \in F_t$. Now for all t

$$\int_0^t \frac{\|I_{[\lambda<T]}\|^2 d\lambda}{[q_\lambda^{1/2}\ln^{1/2-\eta}(q_\lambda+e)]^2} = \frac{-1}{2\eta \ln^{2\eta}(q_\lambda+e)}\Bigg|_0^t < \infty. \qquad (3.8)$$

Let us define

$$x_t = \int_0^t \frac{I_{[\lambda<T]}M_\lambda\, dw_\lambda}{q_\lambda^{1/2}\ln^{1/2\eta}(q_\lambda+e)} \qquad (3.9)$$

and recall that under (3.8), via Martingale convergence, [4]

$$\lim_{t\to\infty} x_t = x \text{ a.s.} \qquad (3.10)$$

for some random variable x. Also, applying the Itô formula [5]

$$q_t^{1/2}\ln^{1/2+\eta}(q_t+e)x_t = \int_0^t x_\lambda d[q_\lambda^{1/2}\ln^{1/2-\eta}(q_\lambda+e)] + \int_0^t I_{[\lambda<T]}M_\lambda\, dw_\lambda. \qquad (3.11)$$

Now

$$\frac{1}{q_t^{1/2}\ln^{1/2+\eta}(q_t+e)}\int_0^t I_{[\lambda<T]}M_\lambda\, dw_\lambda$$

$$= (x_t - x) - \frac{1}{q_t^{1/2}\ln^{1/2-\eta}(q_t+e)}\int_0^t (x_\lambda - x)\, d[q_\lambda^{1/2}\ln^{1/2+\eta}(q_\lambda+e)]$$

$$\xrightarrow{\text{a.s.}} \begin{cases} o(1) \text{ as } t\to\infty \text{ if } q_t\to\infty \\ 0(1) \text{ as } t\to\infty \text{ otherwise.}\end{cases}$$

The result (3.7) follows.

Theorem 3.1: For the signal model/conditions (2.1)–(2.7) and ELS estimation scheme (2.8)'–(2.10), then as $t \to \sigma$

$$\|\theta - \theta_t\|^2 = 0\left(\frac{\ln \lambda_{\max} P_t^{-1}}{\lambda_{\min} P_t^{-1}}\right) \text{ a.s.} \qquad (3.12a)$$

$$= 0\left(\frac{\ln \lambda_{\max}(P_t^0)^{-1}}{\lambda_{\min}(P_t^0)^{-1} - \ln \lambda_{\max}(P_t^0)^{-1}}\right) \text{ a.s.} \qquad (3.12b)$$

where P_t^0 is given from (2.8)' with Φ_t replacing ϕ_t. Moreover, on the set

$$H_1 = \left\{\omega : \limsup_{t\to T} \frac{\ln \lambda_{\max} P_t^{-1}}{\lambda_{\min} P_t^{-1}} < \infty\right\} \qquad (3.13a)$$

$$H_2 = \left\{ \omega : \limsup_{t \to T} \frac{\ln \lambda_{\max} (P_t^0)^{-1}}{\lambda_{\min} (P_t^0)^{-1} - \ln \lambda_{\max} (P_t^0)^{-1}} < \infty \right\} \quad (3.13b)$$

then there is almost surely no finite escape time ($\sigma = \infty$).

Proof: i) From (2.3), $W(s) d\tilde{v}_t^T = W(S)(dv_t - d\hat{v}_t)^T = [dw_t - W(S) d\hat{v}_t^T]$ so that from (2.8)′, (3.2), (3.4)

$$d\tilde{\theta}_t = -I_{[t<\sigma]} P_t \phi_t \left\{ \frac{1}{2} g_t dt + f_t dt + dw_t \right\}^T. \quad (3.14)$$

Applying the Itô formula, then almost surely

$$d(\operatorname{tr} \tilde{\theta}_t^T P_t^{-1} \tilde{\theta}_t) = -I_{[t<\sigma]} \{2 \operatorname{tr} \tilde{\theta}_t^T \phi_t \left(\frac{1}{2} g_t dt + f_t dt + dw_t \right)^T$$

$$- \operatorname{tr} \tilde{\theta}_t^T \phi_t \phi_t^T \tilde{\theta}_t dt - \phi_t^T P_t \phi_t dt \}$$

$$= -I_{[t<\sigma]} \{2 f_t^T g_t + 2 dw_t^T g_t - \phi_t^T P_t \phi_t dt\}$$

$$\leqq -I_{[t<\sigma]} \{2\epsilon (\|g_t\|^2 + \|f_t\|^2) dt + 2 g_t^T dw_t - \phi_t^T P_t \phi_t dt\}$$

where the inequality follows from application of (3.1), under (2.4). Integrating and reorganizing gives, with $t \Lambda \sigma$ denoting min (t, σ)

$$0 \leqq \operatorname{tr} \tilde{\theta}_t^T P_t^{-1} \tilde{\theta}_t + \epsilon \int_0^{t\Lambda \delta} (\|g_\lambda\|^2 + \|f_\lambda\|^2) \, d\lambda$$

$$\leqq \operatorname{tr} \tilde{\theta}_0^T P_0^{-1} \tilde{\theta}_0 + 2k_1 + \int_0^{t\Lambda\sigma} \phi_\lambda^T P_\lambda \phi_\lambda \, d\lambda$$

$$- \left[\epsilon \int_0^{t\Lambda\sigma} \|g_\lambda\|^2 \, d\lambda + 2 \int_0^{t\Lambda\sigma} g_\lambda^T \, dw_\lambda \right]. \quad (3.15)$$

Now applying Lemma 3.3, as $t \to \sigma$, the last integral is dominated by the second last integral so that the square bracketed form becomes negative if $\int_0^\infty \|g_\lambda\|^2 \, d\lambda = \alpha$ and is of 0(1) if $\int_0^\infty \|g_\lambda\|^2 \, d\lambda < \infty$. Noting this and applying Lemma 3.2, as $t \to \sigma$ (3.15) leads to

$$\int_0^{t\Lambda\sigma} (\|g_\lambda\|^2 + \|f_\lambda\|^2) \, d\lambda + \operatorname{tr} \tilde{\theta}_t^T P_t^{-1} \tilde{\theta}_t$$

$$= 0(1) + \ln (\det P_t^{-1}) \text{ a.s.} \quad (3.16)$$

Now $\det P_t^{-1} \geqq \lambda_{\max}(P_t^{-1})(a)^{1-a} \geqq \operatorname{tr} (P_t^{-1}) a^{-a}$ so that

$$\ln \operatorname{tr} P_t^{-1} - a \ln a \leqq \ln \det P_t^{-1} \leqq a \ln \operatorname{tr} P_t^{-1}. \quad (3.17)$$

Noting that $\|\tilde{\theta}_t\|^2 \leqq [\lambda_{\min} P_t^{-1}]^{-1} \operatorname{tr} \theta_t^T P_t^{-1} \theta_t$, then (3.16), (3.17) lead to (3.12a).

ii) From (3.16), as $t \to \sigma$

$$\int_0^t \|g_\lambda\|^2 \, d\lambda = 0 \, (\ln \operatorname{tr} P_t^{-1}).$$

Hence, by Lemma 3.1, as $t \to \sigma$

$$\int_0^t \|\tilde{\theta}_\lambda\|^2 \, d\lambda = 0 \, (\ln \operatorname{tr} P_t^{-1}). \quad (3.18)$$

Thus, we have as $t \to \sigma$

$$\operatorname{tr} P_t^{-1} \leqq 2 \operatorname{tr} (P_t^0)^{-1} + 2 \int_0^t \|\tilde{\theta}_\lambda\|^2 \, d\lambda$$

$$= 2 \operatorname{tr} (P_t^0)^{-1} + 0 \, (\ln \operatorname{tr} P_t^{-1})$$

which means that as $t \to \sigma$

$$\operatorname{tr} P_t^{-1} = 0 \, [\operatorname{tr} (P_t^0)^{-1}], \; \lambda_{\max} P_t^{-1} = 0[\lambda_{\max} P_t^0)^{-1}]. \quad (3.19)$$

Further, for any $x \in R^a$ with $\|x\| = 1$, we know that for $t < \sigma$

$$\int_0^t \|x^T \phi_\lambda^0\|^2 \, d\lambda \leqq 2 \int_0^t \|x^T \phi_\lambda\|^2 \, d\lambda + 2 \int_0^t \|x^T \tilde{\phi}_\lambda\|^2 \, d\lambda.$$

Selecting x *so that the first integral on the right-hand side is* $\lambda_{\min}(P_t^{-1})$, then the right-hand side is not less than $\lambda_{\min}(P_t^0)^{-1}$. Hence, for $t < \sigma$ and applying (3.18)

$$\lambda_{\min} \, (P_t^0)^{-1} \leqq 2 \, \{\lambda_{\min} \, (P_t^{-1}) + 0 \, [\ln \lambda_{\max} \, (P_t)^{-1}].$$

Thus, as $t \to \sigma$ and applying (3.19)

$$\lambda_{\min} \, (P_t)^{-1} \geqq 0[(\lambda_{\min} \, P_t^{0-1}) - \ln \lambda_{\max} \, (P_t^0)^{-1}]. \quad (3.20)$$

Application of (3.19), (3.20) in (3.12a) gives the result (3.12b).

iii) Now on the set H_i of (3.13), from (3.12) $\|\theta_t\|$ does not diverge to ∞. Consequently, the system generating ϕ_t can be viewed as a linear time-varying system with parameters θ, θ_t which do not diverge to ∞ on H. Thus, under (2.5) ϕ_t can grow no faster than exponentially and there is then no finite escape time on H.

Corollary 3.1: Consider the plant (signal model) (2.1) with $C(S) = I$, (2.3) with $W(s) = I$, so that v_t is a Wiener process (and trivially (2.2), (2.4) are satisfied). Then $\phi_t = \phi_t^0$ *and on the least squares estimation of* θ via (2.8), (2.9) under (2.5) (2.10), there is almost surely no finite escape time, so $\sigma = \infty$ and $I_{[t<\sigma]} = 1$. Moreover, there is almost sure convergence of the parameter estimates θ_t given from (with $\bar{P}_t = P_t$)

$$\|\theta - \theta_t\|^2 = 0 \left(\frac{\ln \lambda_{\max} P_t^{-1}}{\lambda_{\min} P_t^{-1}} \right) \quad \text{as } t \to \infty. \quad (3.21)$$

Proof: Since $\phi_t = \phi_t^0$ is derived from a linear time-invariant system with parameters θ, ϕ_t^0 can grow no faster than exponentially under (2.5), and there is no finite escape time. The result (3.21) now follows from an application of Theorem 3.1.

Remarks:

1) Under a plant stability assumption there is no finite escape time for ϕ_t^0 and $\lambda_{\max}(P_t^0)^{-1} \leqq 0(t)$. With u_t, v_t suitably exciting so that $\lambda_{\min}(P_t^0)^{-1} \geqq 0(t)$, then (3.18), (3.21) become

$$\|\theta - \theta_t\|^2 = 0(t^{-1} \ln t) \quad \text{as } t \to \infty. \quad (3.22)$$

2) The result of Lemma 3.3 is perhaps of independent interest to any ELS stochastic analysis.

IV. Conclusion

Convergence (and divergence) rates for ELS estimation of a class of linear continuous-time stochastic signal models have been demonstrated. Assuming the nonlinear stochastic differential equations involved have solutions which exist up until (sample path dependent) finite escape times, conditions are derived which exclude the existence of finite escape times.

The material of this note complements derivations of a companion paper which employ weighting coefficient selections to avoid finite escape times, and gives global convergence rates for continuous-time ELS based stochastic adaptive control [4].

References

[1] J. H. Van Schuppen, "Convergence results for continuous time adaptive stochastic filtering algorithms," *J. Math. Anal. Appl.*, vol. 96, no. 1, pp. 209-225, 1983.

[2] H. F. Chen, *Recursive Estimation and Control for Stochastic Systems.* New York: Wiley, 1985.

[3] ——, "Quasi-least-squares identification and its strong consistency," *Int. J. Contr.*, vol. 34, no. 5, pp. 921-936, 1981.

[4] J. B. Moore, "Convergence of continuous time stochastic adaptive ELS estimation and control schemes," submitted for publication.

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

[6] H. F. Chen and L. Guo, "Convergence rate of least squares identification and adaptive control for stochastic systems," *Int. J. Contr.*, pp. 1459-1476, Nov. 1986.