SIAM J. CONTROL AND OPTIMIZATION
Vol. 25, No. 3, May 1987

004

ASYMPTOTICALLY OPTIMAL ADAPTIVE CONTROL WITH CONSISTENT PARAMETER ESTIMATES*

H. F. CHEN† AND L. GUO†

Abstract. For the discrete-time linear stochastic systems with unknown coefficients we give an adaptive control by which both strong consistency of the parameter estimates and asymptotic optimality for the tracking system are achieved simultaneously. This is done by disturbing the signal that is to be tracked, and the disturbance consists of a sequence of random vectors with covariance matrices tending to zero. The main result is essentially based on some criteria for consistency of parameter estimate for the system without monitoring, which are also demonstrated in the paper. The existence of adaptive control is also discussed.

Key words. stochastic system, parameter estimate, strong consistency, adaptive tracking, asymptotic optimality

AMS(MOS) subject classification. 93C40

1. Introduction. Since Åström and Wittenmark [1] introduced the self-tuning regulator, much work has been devoted in recent years to the parameter-adaptive control and to the related parameter estimation problem. For the adaptive tracking problem, Goodwin, Ramadge and Caines [15] and Sin and Goodwin [21] have established the global convergence of the system and the asymptotic optimality of the tracking error by use of the stochastic gradient and the modified least squares algorithms, respectively. On the other hand, for linear stochastic systems without monitoring there are different conditions guaranteeing the strong consistency of estimates for the unknown system coefficients by invoking various approaches such as the probabilisitic method (Ljung [18], Moore [20], Solo [22]), the ordinary differential equation method (Ljung [19], Kushner and Clark [17]) and the combined treatment (Chen [5], [6], [8]). But the crucial point in these different conditions is almost the same fact—the persistent excitation condition, which means that for the matrix $\sum_{i=1}^{n} \varphi_i\varphi_i^T$ consisting of the stochastic regressors φ_i the ratio of its maximum to minimum eigenvalues is bounded. Unfortunately, it does not always take place for the system with asymptotically optimal adaptive control given in Goodwin, Ramadge and Caines [15] and Sin and Goodwin [21], as shown in Becker, Kumar and Wei [2].

In order to get the consistent estimate for unknown parameters the adaptive control law is disturbed by a random noise introduced artificially (see Caines and Lafortune [3], Chen [7], Chen and Caines [9]). With such a treatment it turns out that the estimate is strongly consistent but the tracking error differs from its minimal value by an additional term caused by the random noise added to the adaptive control law.

However, all these facts do not mean that there is no adaptive control law forcing the long run average of the tracking errors to be minimal and, at the same time, making the parameter estimate strongly consistent, since the asymptotically optimal adaptive control law is not unique.

In this paper, we first give an adaptive control by which both strong consistency of the estimates and optimality for the tracking system are achieved simultaneously. The main idea is that the asymptotically optimal adaptive control is disturbed by a random vector sequence with vanishing covariance matrices, in contrast to the work of Caines and Lafortune [3], Chen and Caines [9] and Chen [7], where the disturbance

* Received by the editors March 27, 1985; accepted for publication (in revised form) March 5, 1986.
† Institute of Systems Science, Academia Sinica, Beijing, People's Republic of China.

is of constant covariance matrix. As a result the matrix $\sum_{i=1}^n \varphi_i\varphi_i^\tau$ mentioned above is ill-conditioned; hence no persistent excitation-like condition can be applied to guarantee consistancy for estimates. However, recently the authors have obtained some new results (Chen and Guo [10], [11], [12]), establishing the strong consistency of parameter estimates for systems with $\sum_{i=1}^n \varphi_i\varphi_i^\tau$ ill-conditioned, and it appears that they are suitable to the analysis of the case of adaptive control with vanishing disturbances and make the system asymptotically optimal and the parameter estimates strongly consistent.

2. Statement of the problem. Let $(\Omega, \mathscr{F}, P)$ be a probability space with a family $\{\mathscr{F}_n\}$ of nondecreasing sub-σ-algebras. Consider the following stochastic control system:

$$y_n + A_1 y_{n-1} + \cdots + A_p y_{n-p} = B_1 u_{n-1} + \cdots + B_q u_{n-q} + w_n + C_1 w_{n-1} + \cdots + C_r w_{n-r} \tag{2.1}$$

where y_n, u_n and w_n are the m-, l- and m-dimensional output, input and driven noise, respectively, and $p \geqq 1$, $q \geqq 1$, $y_n = w_n = 0$, $u_n = 0$ for $n < 0$. A_i, B_j, C_k $(i = 1 \dots p$, $j = 1 \dots q$, $k = 1 \dots r)$ are the unknown matrices.

Assume that u_n and w_n are $\mathscr{F}_n$-measurable and

$$E(w_n \mid \mathscr{F}_{n-1}) = 0, \qquad E(\|w_n\|^2 \mid \mathscr{F}_{n-1}) \leqq c_0 r_{n-1}^\varepsilon \tag{2.2}$$

with constants $c_0 > 0$, $\varepsilon \in [0, 1)$ and r_{n-1} defined later on by (2.9).

Let z be the shift-back operator and set

$$A(z) = I + A_1 z + \cdots + A_p z^p, \tag{2.3}$$

$$B(z) = B_1 + B_2 z + \cdots + B_q z^{q-1}, \tag{2.4}$$

$$C(z) = I + C_1 z + \cdots + C_r z^r, \tag{2.5}$$

$$\theta^\tau = [-A_1 \cdots -A_p B_1 \cdots B_q C_1 \cdots C_r]. \tag{2.6}$$

Denote by θ_n the nth estimate for θ, and let θ_n be given by

$$\theta_{n+1} = \theta_n + \frac{\varphi_n}{r_n}(y_{n+1}^\tau - \varphi_n^\tau \theta_n) \tag{2.7}$$

with

$$\varphi_n^\tau = [y_n^\tau, y_{n-1}^\tau, \cdots, y_{n-p+1}^\tau, u_n^\tau \cdots u_{n-q+1}^\tau, y_n^\tau - \varphi_{n-1}^\tau\theta_{n-1}, \cdots, y_{n-r+1}^\tau - \varphi_{n-r}^\tau \theta_{n-r}], \tag{2.8}$$

$$r_n = 1 + \sum_{i=1}^n \|\varphi_i\|^2, \qquad r_0 = 1. \tag{2.9}$$

The initial values θ_0 and φ_0 are arbitrarily chosen.

Under reasonable conditions Goodwin, Ramadge and Caines [15] proved the global convergence and asymptotical optimality of the tracking system with u_n defined from

$$\theta_n^\tau \varphi_n = y_{n+1}^*, \tag{2.10}$$

i.e.,

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=1}^n \|u_i\|^2 < \infty, \qquad \overline{\lim_{n\to\infty}} \frac{1}{n} \sum_{i=1}^n \|y_i\|^2 < \infty \quad \text{a.s.} \tag{2.11}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)(y_i - y_i^*)^\tau = R \qquad \text{a.s.} \tag{2.12}$$

However, in this case the estimate θ_n may be inconsistent (Becker, Kumar and Wei [2]). It can be easily explained by the following example. Let $y_n^* \equiv 0$ and $\theta_0^\tau \theta_0 > \theta^\tau \theta$. Then we have

$$\theta_n^\tau \varphi_n \equiv 0, \qquad \theta_n^\tau(\theta_{n+1} - \theta_n) = \theta_n^\tau \frac{\varphi_n}{r_n} y_{n+1}^\tau \equiv 0;$$

hence

$$\theta_n^\tau \theta_n = \theta_{n-1}^\tau \theta_{n-1} + (\theta_n - \theta_{n-1})^\tau (\theta_n - \theta_{n-1})$$

$$= \theta_0^\tau \theta_0 + \sum_{i=1}^n (\theta_i - \theta_{i-1})^\tau (\theta_i - \theta_{i-1}) \geqq \theta_0^\tau \theta_0 > \theta^\tau \theta.$$

In order to achieve strongly consistent parameter estimates, Caines and Lafortune [3], Chen [7] and Chen and Caines [9] added a disturbance with covariance matrix $R_1 > 0$ to the reference sequence $\{y_n^*\}$. In this case θ_n tends to θ but the long run average of the tracking errors differs from its minimum value R by an additional term R_1:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)(y_i - y_i^*)^\tau = R + R_1 \quad \text{a.s.}$$

It is natural to ask: Is it possible to achieve simultaneously both asymptotic optimality of the adaptive tracking system and the strong consistency of parameter estimates? To answer this question is the topic of the paper.

3. Main result. We first define adaptive control for the tracking system. Let $\{\varepsilon_i\}$ be an m-dimensional i.i.d. sequence which is independent of $\{w_n\}$ with properties $E\varepsilon_i\varepsilon_i^\tau = I$, $E\|\varepsilon_i\|^5 < \infty$.

Without loss of generality we assume $\mathscr{F}_n = \sigma\{w_i, i \leqq n; \varepsilon_j, j \leqq n\}$.

Unlike (2.10) we define adaptive control from the equation

$$\theta_n^\tau \varphi_n = y_{n+1}^* + v_n \tag{3.1}$$

where $\{y_n^*\}$ is a bounded deterministic reference sequence and

$$v_1 = 0, \quad v_n = \frac{\varepsilon_n}{\log^{1/8} n} \quad \forall n \geqq 2. \tag{3.2}$$

(The existence of u_n satisfying (3.1) or (2.10) is discussed in Appendix 1.)

The disturbance v_n in (3.1) is designed to have a vanishing covariance matrix in order to make tracking error asymptotically minimal, but for this the system loses the persistent excitation property which is of crucial importance in the analysis of Caines and Lafortune [3], Chen [7] and Chen and Caines [9]. To overcome this difficulty is the main task of the present paper.

We need the following conditions:

(A_1) $C(z) - \frac{1}{2}I$ is strictly positive real;

(A_2) B_1 if of full rank and zeros of det $B_1^+ B(z)$ lie outside the closed unit disk;

(A_3) $B_1^+ A(z)$ and $B_1^+ B(z)$ are left-coprime and $B_1^+ B_q$ is of full rank;

(A_4) $\{w_i\}$ is a mutually independent sequence with $Ew_i = 0$; $\sup_i E\|w_i\|^{4+\delta} < \infty$ for some $\delta > 0$ and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n w_i w_i^\tau = R > 0 \quad \text{a.s.} \tag{3.3}$$

THEOREM 1. *For system* (2.1)-(2.2) *let the parameter estimate be given by* (2.7)-(2.9) *and let the control be defined by* (3.1) *and* (3.2). *If conditions* A_1-A_4 *are fulfilled, then the tracking system is asymptotically optimal and the estimate is strongly consistent, i.e.,* (2.11) *and* (2.12) *take place and* $\theta_n \xrightarrow[n\to\infty]{} \theta$. *a.s.*

The proof of this theorem is given in § 5. For this we need some criteria for strong consistency of parameter estimate for systems without monitoring.

4. Parameter estimation for systems without monitoring. In contrast to φ_n we define an estimate-free vector φ_n^0:

$$\varphi_n^0 = [y_n^\tau \cdots y_{n-p+1}^\tau, u_n^\tau \cdots u_{n-q+1}^\tau, w_n^\tau \cdots w_{n-r+1}^\tau]^\tau \tag{4.1}$$

and set

$$\xi_n = y_n - w_n - \theta_{n-1}^\tau \varphi_{n-1}, \tag{4.2}$$

$$\varphi_n^\xi = [0 \cdots 0, 0 \cdots 0, \xi_n^\tau \cdots \xi_{n-r+1}^\tau]^\tau. \tag{4.3}$$

Then we have

$$\varphi_n = \varphi_n^0 + \varphi_n^\xi, \tag{4.4}$$

$$y_{n+1} = \theta^\tau \varphi_n^0 + w_{n+1}, \tag{4.5}$$

and

$$\begin{aligned}\theta_{n+1} &= \theta_n + \frac{\varphi_n}{r_n}(\varphi_n^{0\tau}\theta + w_{n+1}^\tau - \varphi_n^\tau \theta_n)\\ &= \theta_n + \frac{\varphi_n}{r_n}(\varphi_n^\tau\theta - \varphi_n^{\xi\tau}\theta + w_{n+1}^\tau - \varphi_n^\tau\theta_n);\end{aligned}$$

hence

$$\tilde{\theta}_{n+1} = \left(I - \frac{\varphi_n\varphi_n^\tau}{r_n}\right)\tilde{\theta}_n + \frac{\varphi_n\varphi_n^{\xi\tau}}{r_n}\theta - \frac{\varphi_n}{r_n}w_{n+1}^\tau \tag{4.6}$$

with

$$\tilde{\theta}_n = \theta - \theta_n. \tag{4.7}$$

Let the matrix $\Phi(n, i)$ be recursively defined by

$$\Phi(n+1, i) = \left(I - \frac{\varphi_i\varphi_i^\tau}{r_i}\right)\Phi(n, i), \qquad \Phi(i, i) = I. \tag{4.8}$$

Then from (4.6) it follows that

$$\tilde{\theta}_{n+1} = \Phi(n+1, 0)\tilde{\theta}_0 + \sum_{j=0}^{n} \Phi(n+1, j+1)\frac{\varphi_j\varphi_j^{\xi\tau}}{r_j}\theta - \sum_{j=0}^{n}\Phi(n+1, j+1)\frac{\varphi_j}{r_j}w_{j+1}^\tau, \tag{4.9}$$

from which we see that the behavior of $\Phi(n, 0)$ is of great importance for consistency of parameter estimates.

LEMMA 1. *For the system and algorithm defined by* (2.1)-(2.2) *and* (2.7)-(2.9) *if condition* A_1 *holds, then*

$$\sum_{n=0}^{\infty} \frac{\|\xi_{n+1}\|^2}{r_n} < \infty \quad a.s.; \tag{4.10}$$

moreover, if conditions A_2 *and* A_4 *hold and* (2.10) *or* (3.1) *is satisfied, then* $r_n \to \infty$, *and*

$$\frac{1}{n}\sum_{i=0}^{n} \|\xi_{i+1}\|^2 \xrightarrow[n\to\infty]{} 0. \tag{4.11}$$

Proof. In Chen and Caines [9] and Chen [7], (4.10) and (4.11) are proved for v_n with the constant covariance matrix, but they can be verified by the same argument used there. □

LEMMA 2. *For the system and algorithm defined by* (2.1)-(2.2) *and* (2.7)-(2.9) *if condition* A_1 *holds then* $\Phi(n,0) \xrightarrow[n\to\infty]{} 0$ *implies* $\theta_n \xrightarrow[n\to\infty]{} \theta$ *a.s. for any initial value* θ_0. *For the special case of* $r=0$, *the converse assertion is also true, i.e., if* $\theta_n \xrightarrow[n\to\infty]{} \theta$ *a.s. for any* θ_0 *then* $\Phi(n,0) \xrightarrow[n\to\infty]{} 0$ *a.s.*

Proof. The first step is to show

$$\sum_{i=0}^{n-1} \frac{\|\Phi(n,i+1)\varphi_i\|^2}{r_i} \leqq d \tag{4.12}$$

for any vector sequence $\{\varphi_n\}$ with $\Phi(n,i)$ and r_n related by (2.9) and (4.8), where d is the dimension of φ_n.

Then by (2.2) and (4.12) we can prove that the last term of (4.9) goes to zero if $\Phi(n,0) \xrightarrow[n\to\infty]{} 0$. Finally, by (4.10) and (4.12) the second term on the right-hand side of (4.9) also converges to zero if $\Phi(n,0) \xrightarrow[n\to\infty]{} 0$. This is just a sketch proof for the first conclusion. For detailed proof we refer to Chen and Guo [12]. The second conclusion can be easily seen from (4.9). □

LEMMA 3. *If* $r_n \xrightarrow[n\to\infty]{} \infty$, $\overline{\lim}_{n\to\infty} r_n/r_{n-1} < \infty$ *and there exist quantities* N_0 *and* M *possibly depending on* ω *such that*

$$\frac{\lambda_{\max}^n}{\lambda_{\min}^n} \leqq M(\log r_n)^{1/4} \quad a.s. \quad \forall n \geqq N_0; \tag{4.13}$$

then $\Phi(n,0) \xrightarrow[n\to\infty]{} 0$, *where* $\lambda_{\max}^n$ *and* $\lambda_{\min}^n$ *denote the maximum and minimum eigenvalue of the matrix* $\sum_{i=1}^n \varphi_i\varphi_i^\tau + (1/d)I$ *respectively and* d *denotes the dimension of* φ_n.

Proof. We only give a sketch of the proof and refer readers interested in details to Chen and Guo [10], [11].

The key point is to find a function $m(t)$ such that $m(t) \xrightarrow[t\to\infty]{} \infty$ and

$$\|\Phi(m(N+k\alpha), m(N+(k-1)\alpha))\| \leqq \sqrt{1-\frac{\beta^2}{c_1 k}} \quad \forall k \geqq 1$$

for some N, $\alpha>0$, $\beta>0$ and $c_1>0$. If it has been done, then

$$\begin{aligned}\|\Phi(m(N+k\alpha),0)\| &\leqq \prod_{i=1}^{k} \|\Phi(m(N+i\alpha), m(N(i-1)\alpha))\| \cdot \|\Phi(m(N),0)\| \\ &\leqq \left[\prod_{i=1}^{k}\left(1-\frac{\beta^2}{c_1 i}\right)\right]^{1/2} \xrightarrow[k\to\infty]{} 0;\end{aligned}$$

hence $\Phi(n,0) \xrightarrow[n\to\infty]{} 0$ since $\|\Phi(j+1,j)\| \leqq 1$ for all j.

It appears that the following defined function can serve as the desired one:

$$m(t) = \max[n: t_n \leqq t],$$

$$t_n \triangleq \sum_{i=2}^{n-1} \frac{\|\varphi_i\|^2}{r_i(\log r_i)^{1/4}}. \qquad \square$$

Remark 1. Lemma 3 is a purely algebraic result, namely, it is true for any vector sequence $\{\varphi_n\}$, only if $\Phi(n,0)$, φ_n and r_n are related by (2.9) and (4.8).

Remark 2. There exists an example (see Chen and Guo [13]) showing that Lemma 3 is no longer true if condition (4.13) is replaced by a more general one:

$$\lambda_{\max}^n/\lambda_{\min}^n \leqq M(\log r_n)^{1+a}, \qquad a>0.$$

This means that, in order for the estimate given by (2.7) to be consistent, the condition number of $\sum_{i=1}^n \varphi_i\varphi_i^\tau+(1/d)I$ is allowed to diverge at a rate of $(\log r_n)^{1/4}$, but not faster than $(\log r_n)^{1+a}$.

LEMMA 4. *Let* $\{\varphi_n^1\}$, $\{\varphi_n^2\}$ *and* $\{\psi_n\}$ *be the vector sequence satisfying conditions* $\varphi_n^1=\varphi_n^2+\psi_n$ *and*

$$\sum_{n=0}^{\infty}\frac{\|\psi_n\|^2}{r_{1n}}<\infty. \tag{4.14}$$

Then $\Phi_1(n,0)\xrightarrow[n\to\infty]{}0$ *if and only if* $\Phi_2(n,0)\xrightarrow[n\to\infty]{}0$, *where by definition*

$$\Phi_i(n+1,0)=\left(I-\frac{\varphi_n^i\varphi_n^{i\tau}}{r_{in}}\right)\Phi_i(n,0), \qquad \Phi_i(0,0)=I,$$

$$r_{in}=1+\sum_{j=1}^n\|\varphi_j^i\|^2, \quad r_{i0}=1, \quad i=1,2.$$

Proof. Without loss of generality we assume that $\|\varphi_0^1\|\neq 1$.

Suppose $\Phi_1(n,0)\xrightarrow[n\to\infty]{}0$; then from the following chain of equalities:

$$\det\Phi_1(n+1,0)=\det\prod_{i=0}^n\Phi_1(i+1,i)=\prod_{i=0}^n\det\left(I-\frac{\varphi_i^1\varphi_i^{1\tau}}{r_{1i}}\right)$$

$$=\prod_{i=1}^n\frac{r_{1i-1}}{r_{1i}}(1-\|\varphi_0^1\|^2)=\frac{1}{r_{1n}}(1-\|\varphi_0^1\|^2)$$

we see that $r_{1n}\xrightarrow[n\to\infty]{}\infty$.

By (4.14) and the Kronecker lemma we have

$$\frac{r_{2n}}{r_{1n}}=\frac{r_{1n}-2\sum_{i=1}^n\varphi_i^{1\tau}\psi_i+\sum_{i=1}^n\|\psi_i\|^2}{r_{1n}}\xrightarrow[n\to\infty]{}1 \tag{4.15}$$

and by (4.14)

$$\sum_{n=0}^{\infty}\frac{\|\psi_n\|^2}{r_{2n}}<\infty. \tag{4.16}$$

We immediately verify that

$$\Phi_2(n+1,0)=\Phi_1(n+1,0)+\sum_{j=0}^n\Phi_1(n+1,j+1)\frac{\varphi_j^1\psi_j^\tau}{r_{1j}}\Phi_2(j,0)$$

$$+\sum_{j=0}^n\Phi_1(n+1,j+1)\frac{\psi_j\varphi_j^{2\tau}}{r_{2j}}\Phi_2(j,0)$$

$$+\sum_{j=0}^n\frac{\Phi_1(n+1,j+1)\varphi_j^1}{r_{1j}^{1/2}}\left(\sqrt{\frac{r_{2j}}{r_{1j}}}-\sqrt{\frac{r_{1j}}{r_{2j}}}\right)\frac{\varphi_j^{2\tau}\Phi_2(j,0)}{r_{2j}^{1/2}}.$$

By using (4.12) and (4.14)–(4.16) it is not difficult to conclude that $\Phi_1(n,0)\xrightarrow[n\to\infty]{}0$ implies $\Phi_2(n,0)\xrightarrow[n\to\infty]{}0$. The converse implication is proved in a similar way. □

THEOREM 2. *For the system and algorithm defined by* (2.1)-(2.2) *and* (2.7)-(2.9) *if condition* A_1 *holds and if* $r_n \xrightarrow[n\to\infty]{} \infty$,

$$\overline{\lim_{n\to\infty}}\, r_n/r_{n-1} < \infty \text{ and } \lambda_{\max}^n/\lambda_{\min}^n \leqq M(\log r_n)^{1/4} \quad \forall n \geqq N$$

(*or* $r_n^0 \xrightarrow[n\to\infty]{} \infty$, $\overline{\lim}_{n\to\infty} r_n^0/r_{n-1}^0 < \infty$ *and* $\lambda_{\max}^{0n}/\lambda_{\min}^{0n} \leqq M(\log r_n^0)^{1/4}$ *for all* $n \geqq N$) *with* N *and* M *possibly depending on* ω, *then*

$$\theta_n \xrightarrow[n\to\infty]{} \theta \quad a.s.$$

for any initial value, where $\lambda_{\max}^{0n}$, $\lambda_{\min}^{0n}$ *denote the maximum and minimum eigenvalue of* $\sum_{i=1}^n \varphi_i^0\varphi_i^{0\tau} + (1/d)I$, *respectively, and* $r_n^0 = 1 + \sum_{i=1}^n \|\varphi_i^0\|^2$ *with* φ_n^0 *defined by* (4.1).

Proof. Since (4.4), (4.10) and Lemma 4 can be applied with $\varphi_n^1 = \varphi_n$, $\varphi_n^2 = \varphi_n^0$ and $\psi_n = \varphi_n^\xi$, hence $\Phi(n, 0) \xrightarrow[n\to\infty]{} 0$ if and only if $\Phi_0(n, 0) \xrightarrow[n\to\infty]{} 0$, where $\Phi_0(n, 0)$ is defined by (4.8) with φ_n and r_n replaced by φ_n^0 and r_n^0, respectively. Then the conclusions of the theorem immediately follow from Lemmas 2 and 3. □

5. Proof of Theorem 1. To begin with we prove the following lemmas.

LEMMA 5. *Let* $\{v_n\}$ *be defined by* (3.2) *and let* $H_N(z) = \sum_{i=0}^\infty H_i(N)z^i$ *be the matrix series in shift-back operator* z, *where the matrix coefficients* $H_i(N)$ *may depend upon* ω, *but there are constants* (*independent of* ω) $k_1 > 0$ *and* $k_2 > 0$ *such that*

$$\|H_i(N)\| \leqq k_1 \exp(-k_2 i) \quad \forall i \quad \forall N \geqq 0.$$

Then

$$\lim_{N\to\infty} \frac{\log^{1/4} N}{N} \sum_{n=1}^N (H_N(z)v_n)(H_N(z)v_n)^\tau = \lim_{N\to\infty} \frac{\log^{1/4} N}{N} \sum_{i=0}^N H_i(N)H_i^\tau(N)\left(\sum_{j=2}^N \frac{1}{\log^{1/4} j}\right).$$

Proof. See Appendix 2.

LEMMA 6. *Let condition* A_4 *except* (3.3) *be held and let* $H(z) = \sum_{i=0}^\infty H_i z^i$ *and* $G(z) = \sum_{i=0}^\infty G_i z^i$ *be matrix series in shift-back operator* z *with* $\|H_i\| + \|G_i\| \leqq k_1 \exp(-k_2 i)$ *for all* $i \geqq 0$, *for some constants* $k_1 > 0$, $k_2 > 0$. *Then there exists* $\gamma \in (0, 1)$ *such that for all* $l \geqq 0$, $m \geqq 0$,

$$\lim_{N\to\infty} \frac{1}{N^\gamma} \sum_{n=1}^N (H(z)w_{n+1-l})(G(z)v_{n-m})^\tau = 0 \quad a.s., \tag{5.1}$$

$$\lim_{N\to\infty} \frac{1}{N^\gamma} \sum_{n=1}^N (H(z)w_{n+1-l})\eta_n^\tau = 0 \quad a.s. \tag{5.2}$$

for any bounded deterministic sequence $\{\eta_n\}$, *and*

$$\overline{\lim_{N\to\infty}} \frac{1}{N} \sum_{n=1}^N \|H(z)w_{n+1-l}\|^2 < \infty \quad a.s. \tag{5.3}$$

Proof. See Appendix 2.

Set

$$H_1(z) = [B_1^+B(z)]^{-1}B_1^+A(z), \tag{5.4}$$

$$H_2(z) = H_1(z) - [B_1^+B(z)]^{-1}B_1^+C(z), \tag{5.5}$$

$$Y_n^* = [y_n^{*\tau} \cdots y_{n-p+1}^{*\tau}, (H_1(z)y_{n+1}^*)^\tau \cdots (H_1(z)y_{n-q+2}^*)^\tau]^\tau \tag{5.6}$$

and

$$Z_n=[v_{n-1}^\tau \cdots v_{n-p}^\tau, (H_1(z)v_n)^\tau \cdots (H_1(z)v_{n-q+1})^\tau]^\tau. \tag{5.7}$$

In the following, by $\lambda_{\min}(X)$ $(\lambda_{\max}(X))$ we mean the minimum (maximum) eigenvalue of the matrix X; we have the following.

LEMMA 7. *Under conditions of Theorem* 1 *if*

$$\varliminf_{N\to\infty} \lambda_{\min}\left(\frac{\log^{1/4} N}{N}\sum_{n=1}^{N}(Y_n^* Y_n^{*\tau}+Z_n Z_n^\tau)\right)\neq 0 \quad a.s. \tag{5.8}$$

then

$$\theta_n \xrightarrow[n\to\infty]{} \theta \quad a.s.$$

Proof. By Theorem 2 we only need to prove $\Phi_0(n,0)\xrightarrow[n\to\infty]{} 0$ a.s. From (2.1) we have

$$u_n=[B_1^+B(z)]^{-1}B_1^+A(z)y_{n+1}-[B_1^+B(z)]^{-1}B_1^+C(z)w_{n+1}; \tag{5.9}$$

then by (3.1), (4.2), (5.4), (5.5) and (5.9), φ_n^0 defined by (4.1) can be written as

$$\varphi_n^0=\varphi_n^1+\psi_n \tag{5.10}$$

where

$$\psi_n=[\xi_n^\tau, \cdots \xi_{n-p+1}^\tau, (H_1(z)\xi_{n+1})^\tau \cdots (H_1(z)\xi_{n-q+2})^\tau 0 \cdots 0]^\tau, \tag{5.11}$$

$$\varphi_n^1=\varphi_n^2+\varphi_n^3, \tag{5.12}$$

$$\varphi_n^2=[w_n^\tau \cdots w_{n-p+1}^\tau, (H_2(z)w_{n+1})^\tau \cdots (H_2(z)w_{n-q+2})^\tau, w_n^\tau \cdots w_{n-r+1}^\tau]^\tau, \tag{5.13}$$

$$\varphi_n^3=[Y_n^{*\tau}+Z_n^\tau, 0 \cdots 0]^\tau. \tag{5.14}$$

By (4.4), (4.10), similar to (4.15) we have

$$\frac{r_n^0}{r_n}=\frac{r_n-2\sum_{i=1}^n \varphi_i^\tau\varphi_i^\xi+\sum_{i=1}^n\|\varphi_i^\xi\|^2}{r_n}\xrightarrow[n\to\infty]{} 1. \tag{5.15}$$

Then by the Schwarz inequality it follows that

$$\sum_{n=0}^\infty \frac{\|H_1(z)\xi_{n+1-l}\|^2}{r_n^0}\leqq \sum_{n=0}^\infty \frac{1}{r_n^0}\sum_{i=0}^\infty \|H_{1i}\| \sum_{i=0}^\infty \|H_{1i}\|\cdot\|\xi_{n+1-l-i}\|^2$$

$$\leqq k_0\sum_{i=0}^\infty \|H_{1i}\|\sum_{n=0}^\infty \frac{\|\xi_{n+1-l-i}\|^2}{r_n^0}<\infty$$

where the last inequality is obtained because $\xi_i=0$ for $i<0$ and the coefficients in $H_1(z)=\sum_{i=0}^\infty H_{1i}z^i$ have the estimates $\|H_{1i}\|\leqq k_1\exp(-k_2 i)$, for all $i\geqq 0$, $(k_1>0, k_2>0)$ by condition A_2. Thus we have established

$$\sum_{n=0}^\infty \frac{\|\psi_n\|^2}{r_n^0}<\infty \tag{5.16}$$

and by Lemma 4 we conclude that $\Phi_0(n,0)\xrightarrow[n\to\infty]{} 0$ iff $\Phi_1(n,0)\xrightarrow[n\to\infty]{} 0$.

Next, we prove that

$$\varliminf_{N\to\infty} \lambda_{\min}\left(\frac{\log^{1/4} N}{N}\sum_{n=1}^N \varphi_n^1\varphi_n^{1\tau}\right)\neq 0 \quad \text{a.s.} \tag{5.17}$$

If for some $\omega\in\Omega$ (5.17) were not true, then we would find a subsequence of eigenvectors $\left[\begin{smallmatrix}\alpha_{N_k}\\ \beta_{N_k}\end{smallmatrix}\right]$ for matrix $(\log^{1/4} N_k/N_k)\sum_{n=1}^{N_k}\varphi_n^1\varphi_n^{1\tau}$ with $N_k \xrightarrow[k\to\infty]{} \infty$, $\alpha_{N_k}\in R^{mp+lq}$, $\beta_{N_k}\in R^{mr}$ and

$$\|\alpha_{N_k}\|^2+\|\beta_{N_k}\|^2=1 \tag{5.18}$$

such that

$$(\alpha_{N_k}^\tau,\beta_{N_k}^\tau)\frac{\log^{1/4} N_k}{N_k}\sum_{n=1}^{N_k}\varphi_n^1\varphi_n^{1\tau}\begin{pmatrix}\alpha_{N_k}\\ \beta_{N_k}\end{pmatrix}\xrightarrow[k\to\infty]{}0. \tag{5.19}$$

Without loss of generality we always assume that this fixed ω does not belong to a possible exceptional set of probability zero. Obviously, α_{N_k} and β_{N_k} would be ω-dependent but not necessarily measurable.

Utilizing Lemma 6 one can easily be convinced of the fact

$$\frac{\log^{1/4} N}{N}\sum_{n=1}^{N}\varphi_n^2\varphi_n^{3\tau}\xrightarrow[N\to\infty]{}0, \tag{5.20}$$

then (5.19) is reduced to

$$(\alpha_{N_k}^\tau,\beta_{N_k}^\tau)\frac{\log^{1/4} N_k}{N_k}\sum_{n=1}^{N_k}\varphi_n^2\varphi_n^{2\tau}\begin{pmatrix}\alpha_{N_k}\\ \beta_{N_k}\end{pmatrix}\xrightarrow[k\to\infty]{}0, \tag{5.21}$$

$$(\alpha_{N_k}^\tau,\beta_{N_k}^\tau)\frac{\log^{1/4} N_k}{N_k}\sum_{n=1}^{N_k}\varphi_n^3\varphi_n^{3\tau}\begin{pmatrix}\alpha_{N_k}\\ \beta_{N_k}\end{pmatrix}\xrightarrow[k\to\infty]{}0. \tag{5.22}$$

In view of Lemma 6, (5.8) implies

$$\varliminf_{N\to\infty}\lambda_{\min}\left(\frac{\log^{1/4} N}{N}\sum_{n=1}^{N}(Y_n^*+Z_n)(Y_n^*+Z_n)^\tau\right)\neq 0. \tag{5.23}$$

Paying attention to the fact that the last mr elements in φ_n^3 are zeros, by (5.22) and (5.23) we conclude that

$$\alpha_{N_k}\xrightarrow[k\to\infty]{}0; \tag{5.24}$$

hence, recalling (5.18) we have

$$\|\beta_{N_k}\|\xrightarrow[k\to\infty]{}1. \tag{5.25}$$

Let

$$x_n^1=[w_n^\tau\cdots w_{n-p+1}^\tau,(H_2(z)w_{n+1})^\tau\cdots(H_2(z)w_{n-q+2})^\tau]^\tau,$$
$$x_n^2=[w_n^\tau\cdots w_{n-r+1}^\tau]^\tau.$$

Then $\varphi_n^2=[x_n^{1\tau},x_n^{2\tau}]^\tau$ and (5.21) implies

$$\frac{1}{N_k}\sum_{n=1}^{N_k}\|\alpha_{N_k}^\tau x_n^1+\beta_{N_k}^\tau x_n^2\|^2\xrightarrow[k\to\infty]{}0. \tag{5.26}$$

Further, we have

$$\varlimsup_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\|x_n^1\|^2<\infty\quad\text{a.s.} \tag{5.27}$$

by Lemma 6, and

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}x_n^2x_n^{2\tau}=\begin{bmatrix}R & & 0\\ & \ddots & \\ 0 & & R\end{bmatrix}>0 \tag{5.28}$$

by ergodicity.

Thus from (5.24) and (5.26)-(5.28) it follows that

$$\lim_{k\to\infty}\frac{1}{N_k}\sum_{n=1}^{N_k}\beta_{N_k}^{\tau}x_n^2x_n^{2\tau}\beta_{N_k}=0,$$

which leads to $\beta_{N_k}\xrightarrow[k\to\infty]{}0$ by (5.28). Comparing it with (5.25) we obtain a contradiction, which shows the truth of (5.17). Therefore, there exist $\alpha_0>0$, N_0 such that

$$\lambda_{\min}\left(\sum_{i=1}^{n}\varphi_i^1\varphi_i^{1\tau}\right)\geqq\frac{n}{\log^{1/4}n}\alpha_0\quad\forall n\geqq N_0. \tag{5.29}$$

By (5.12) and Lemma 6 it follows that

$$\varliminf_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\|\varphi_i^1\|^2\geq\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\|w_i\|^2=\operatorname{tr}R>0 \tag{5.30}$$

and

$$\varlimsup_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\|\varphi_i^1\|^2<\infty\quad\text{a.s.} \tag{5.31}$$

From (5.30) and (5.31) it follows that there are positive quantities $\beta\geqq\alpha>0$ such that

$$\alpha n\leqq r_{1n}\leqq\beta n, \tag{5.32}$$

which, together with (5.29), yields

$$\lambda_{\max}^{1n}/\lambda_{\min}^{1n}\leqq M\log^{1/4}r_{1n}\quad\forall n\geqq N_0\quad\text{with some } M>0$$

where $\lambda_{\max}^{1n}$ and $\lambda_{\min}^{1n}$ denote, respectively, the maximum and minimum eigenvalues of $\sum_{i=1}^{n}\varphi_i^1\varphi_i^{1\tau}+(1/\alpha)I$. Then we obtain the required assertion $\Phi_1(n,0)\xrightarrow[n\to\infty]{}0$ by Lemma 3 and Remark 1. □

Proof of Theorem 1. Since $\sum_{i=2}^{n}(v_iv_i^\tau-(1/\log^{1/4}i)I)/i$ is a convergent martingale, by the Kronecker lemma it follows that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}v_iv_i^\tau=\lim_{n\to\infty}\frac{1}{n}\sum_{i=2}^{n}\frac{1}{\log^{1/4}i}I=0. \tag{5.33}$$

Similarly we have

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}w_iv_{i-1}^\tau=0. \tag{5.34}$$

From (3.1) and (4.2) it follows that

$$y_{n+1}=\xi_{n+1}+y_{n+1}^*+w_{n+1}+v_n. \tag{5.35}$$

Then (2.11) and (2.12) follow immediately from (4.11), (5.9), (5.33)-(5.35) and condition A_2.

Thus we only need to prove $\theta_n\xrightarrow[n\to\infty]{}\theta$ a.s. By Lemma 7 it suffices to verify that

$$\varliminf_{N\to\infty}\lambda_{\min}\left(\frac{\log^{1/4}N}{N}\sum_{n=1}^{N}Z_nZ_n^\tau\right)\neq0\quad\text{a.s.} \tag{5.36}$$

If (5.36) were not true, then there would exist a subsequence of eigenvectors $\alpha_{N_k} \in R^{mp+lq}$ for matrix $(\log^{1/4} N_k/N_k) \sum_{n=1}^{N_k} Z_n Z_n^\tau$ with $N_k \xrightarrow[k\to\infty]{} \infty$ and

$$\|\alpha_{N_k}\| = 1 \quad \forall k \geqq 1 \tag{5.37}$$

such that

$$\alpha_{N_k}^\tau \frac{\log^{1/4} N_k}{N_k} \sum_{n=1}^{N_k} Z_n Z_n^\tau \alpha_{N_k} \xrightarrow[k\to\infty]{} 0. \tag{5.38}$$

Without loss of generality we suppose $\alpha_{N_k} \xrightarrow[k\to\infty]{} \alpha$. Write α_{N_k} and α in the component form

$$\alpha_{N_k} = [\alpha_1^\tau(N_k) \cdots \alpha_{p+q}^\tau(N_k)]^\tau, \qquad \alpha = [\alpha_1^\tau \cdots \alpha_{p+q}^\tau]^\tau$$

with $\alpha_i(N_k)$, α_i being m-dimensional and $\alpha_{p+j}(N_k)$, α_{p+j} l-dimensional vectors, $i = 1 \cdots p$, $j = 1 \cdots q$.

Set

$$\begin{aligned} H_{N_k}(z) &= \alpha_1^\tau(N_k) z + \cdots + \alpha_p^\tau(N_k) z^p \\ &\quad + \alpha_{p+1}^\tau(N_k) H_1(z) + \cdots + \alpha_{p+q}^\tau(N_k) H_1(z) z^{q-1} \\ &\triangleq \sum_{i=0}^{\infty} h_i^\tau(N_k) z^i, \end{aligned} \tag{5.39}$$

$$\begin{aligned} H(z) &= \alpha_1^\tau z \cdots + \alpha_p^\tau z^p + \alpha_{p+1}^\tau H_1(z) + \cdots + \alpha_{p+q}^\tau H_1(z) z^{q-1} \\ &\triangleq \sum_{i=0}^{\infty} h_i^\tau z^i. \end{aligned} \tag{5.40}$$

We note that α_{N_k} and hence $H_{N_k}(z)$ may depend on ω, but by condition A_2 and (5.37) it is clear that there are constants $c_1 > 0$, $c_2 > 0$ such that $\|h_i(N_k)\| \leqq c_1 \exp(-c_2 i)$ for all $i \geqq 0$, for all $k \geqq 0$. Then Lemma 5 can be applied, and from (5.38), (5.39) we have

$$\begin{aligned} 0 &= \lim_{k\to\infty} \frac{\log^{1/4} N_k}{N_k} \sum_{n=1}^{N_k} [\alpha_1^\tau(N_k) v_{n-1} + \cdots + \alpha_p^\tau(N_k) v_{n-p} + \alpha_{p+1}^\tau(N_k) H_1(z) v_n + \cdots \\ &\qquad\qquad + \alpha_{p+q}^\tau(N_k) H_1(z) v_{n-q+1}]^2 \\ &= \lim_{k\to\infty} \frac{\log^{1/4} N_k}{N_k} \sum_{n=1}^{N_k} [(\alpha_1^\tau(N_k) z + \cdots + \alpha_p^\tau(N_k) z^p + \alpha_{p+1}^\tau(N_k) H_1(z) + \cdots \\ &\qquad\qquad + \alpha_{p+q}^\tau(N_k) H_1(z) z^{q-1}) v_n]^2 \\ &= \lim_{k\to\infty} \frac{\log^{1/4} N_k}{N_k} \sum_{n=1}^{N_k} (H_{N_k}(z) v_n)(H_{N_k}(z) v_n)^\tau \\ &= \lim_{k\to\infty} \frac{\log^{1/4} N_k}{N_k} \sum_{i=0}^{N_k} h_i^\tau(N_k) \sum_{n=2}^{N_k} \frac{1}{\log^{1/4} n} h_i(N_k), \end{aligned}$$

which implies that

$$\lim_{k\to\infty} \sum_{i=0}^{N_k} \|h_i(N_k)\|^2 = 0$$

and hence $\sum_{i=0}^{\infty} \|h_i\|^2 = 0$ by the dominated convergence theorem; therefore $H(z) = 0$. Setting $z = 0$ and paying attention to the fact that $H_1(0)(= B_1^+)$ is of full row rank we see $\alpha_{p+1}^\tau B_1^+ = 0$ and so $\alpha_{p+1} = 0$. Then it follows directly from (5.40) that

$$(\alpha_1^\tau + \alpha_2^\tau z + \cdots + \alpha_p^\tau z^{p-1}) = -(\alpha_{p+2}^\tau + \cdots + \alpha_{p+q}^\tau z^{q-2}) H_1(z). \tag{5.41}$$

In view of condition A_3, applying Lemma 6.6-1 of Kailath [16] to (5.41) we know that there exists a polynomial with vector coefficients $f(z)$:

$$f(z)=f_1+f_2z+\cdots+f_sz^{s-1}, \qquad s\geqq 1,$$

such that

$$(5.42)\qquad \begin{aligned}(\alpha_{p+2}^\tau+\cdots+\alpha_{p+q}^\tau z^{q-2})&=f^\tau(z)B_1^+B(z)\\ &=(f_1^\tau+\cdots+f_s^\tau z^{s-1})(B_1^+B_1+\cdots+B_1^+Bqz^{q-1}).\end{aligned}$$

From here it is easy to conclude that $f_i=0$ $(1\leqq i\leqq s)$ since $B_1^+B_q$ is of full rank by condition A_3, then $\alpha_{p+j}=0$ by (5.42), and then $\alpha_j=0$ by (5.41) $(1\leqq j\leqq q,\ 1\leqq i\leqq p)$. Thus $\alpha=0$, and $\alpha_{N_k}\xrightarrow[k\to\infty]{}0$, thus contradicting (5.37). Hence (5.36) holds. □

6. Concluding discussion. In order to get optimality in both tracking and estimating we have added to $\{y_n^*\}$ a random disturbance with covariance matrix tending to zero, but, intuitively, the disturbance may harm the tracking if time is bounded. However, all assertions of Theorem 1 can remain valid for u_n defined from (3.1) with v_n deleted (i.e. from (2.10)) if the reference signal y_n^* itself is "complicated" enough in the sense that

$$(6.1)\qquad \varliminf_{N\to\infty}\lambda_{\min}\left(\frac{\log^{1/4}N}{N}\sum_{n=1}^N Y_n^*Y_n^{*\tau}\right)\neq 0.$$

This remark can easily be seen from Lemma 7.

For the single-input and single-output system it is easy to show that for (6.1) it suffices to require condition A_3 and

$$(6.2)\qquad \varliminf_{N\to\infty}\lambda_{\min}\left(\frac{\log^{1/4}N}{N}\sum_{n=1}^N [y_n^*\cdots y_{n-p-q+1}^*]^\tau[y_n^*\cdots y_{n-p-q+1}^*]\right)\neq 0.$$

Recently, for multidimensional and random $\{y_n^*\}$ we have obtained conditions similar to (6.2) in order that all conclusions of Theorem 1 hold by applying u_n defined from (2.10). It will be published elsewhere.

Appendix 1. Existence of adaptive control.

LEMMA. (1) *Let A and B be two matrices of dimensions $m\times n$ and $n\times m$, respectively. Then the following equality takes place*

$$\det(I_m+AB)=\det(I_n+BA)$$

where I_n means the $n\times n$ identity matrix.

(2) *Provided x_1 and x_2 are independent random variables, then*

$$\sup_{a\in R^1}P(x_1+x_2=a)\leqq\min\left\{\sup_{a\in R^1}P(x_1=a),\ \sup_{a\in R^1}P(x_2=a)\right\}.$$

Proof. (1) By taking determinants for both sides of the following matrix identity:

$$\begin{bmatrix}I_m & -A\\ 0 & BA+I_n\end{bmatrix}=\begin{bmatrix}I_m & 0\\ -B & I_n\end{bmatrix}\cdot\begin{bmatrix}I_m+AB & -A\\ 0 & I_n\end{bmatrix}\cdot\begin{bmatrix}I_m & 0\\ B & I_n\end{bmatrix},$$

the desired equality is immediately verified.

(2) Denote by $F_1(x)$, $F_2(x)$, $F_{12}(x)$ the distributions of x_1, x_2, x_1+x_2, respectively. Clearly we have

$$F_{12}(x)=\int_{-\infty}^{\infty}F_1(x-y)\,dF_2(y)$$

and

$$F_{12}(x+)=\int_{-\infty}^{\infty} F_1((x-y)+)\, dF_2(y)$$

by the dominated convergence theorem.

Then for any $a \in R^1$

$$P(x_1+x_2=a)=F_{12}(a+)-F_{12}(a)=\int_{-\infty}^{\infty}[F_1((a-y)+)-F_1(a-y)]\, dF_2(y)$$

$$=\int_{-\infty}^{\infty} P(x_1=a-y)\, dF_2(y) \leqq \sup_{a\in R^1} P(x_1=a)\int_{-\infty}^{\infty} dF_2(y)$$

$$=\sup_{a\in R^1} P(x_1=a).$$

Similarly, we have

$$P(x_1+x_2=a)\leqq \sup_{a\in R^1} P(x_2=a)$$

and thus the desired result follows. □

THEOREM. *Assume $m \leqq l$ and $\{w_n\}$ and $\{v_n\}$ are two sequences of mutually independent random vectors and the components of w_n are independent and with continuous distribution functions. Then for any $n \geqq 1$ there exists u_n satisfying* (3.1) *if the initial values are appropriately chosen. Further, this u_n is unique if and only if $m=l$.*

Proof. Let A_{in}, B_{jn}, C_{kn}, $i=1\cdots p$, $j=1\cdots q$, $k=1\cdots r$ be the matrix components of θ_n, i.e.,

$$\theta_n^\tau=[-A_{1n}\cdots -A_{pn}B_{1n}\cdots B_{qn}C_{1n}\cdots C_{rn}].$$

Set

$$\bar{\theta}_n^\tau=[-A_{1n}\cdots -A_{pn}0B_{2n}\cdots B_{qn}C_{1n}\cdots C_{rn}],$$

and

$$\bar{\varphi}_n=[y_n^\tau\cdots y_{n-p+1}^\tau, 0, u_{n-1}^\tau\cdots u_{n-q+1}^\tau, y_n^\tau-\varphi_{n-1}^\tau\theta_{n-1}, \cdots, y_{n-r+1}^\tau-\varphi_{n-r}^\tau\theta_{n-r}]^\tau.$$

Equation (3.1) is equivalent to

$$B_{1n}u_n=y_{n+1}^*+v_n-\bar{\theta}_n^\tau\bar{\varphi}_n. \tag{A.1}$$

First let $m=l$. For this case we only need to prove that B_{1n} is invertible a.s. In fact, if this is true, then from (A.1) u_n is uniquely defined by $u_n=B_{1n}^{-1}(y_{n+1}^*+v_n-\bar{\theta}_n^\tau\bar{\varphi}_n)$, which obviously is $\mathscr{F}_n$-measurable. (In adaptive tracking cases we take $\mathscr{F}_n \triangleq \sigma\{w_i, v_i, i\leqq n\}$.)

From (2.7) and (4.2) we obtain

$$B_{1n+1}=B_{1n}+\frac{1}{r_n}(\xi_{n+1}+w_{n+1})u_n^\tau. \tag{A.2}$$

It is easy to take initial values φ_0, θ_0 such that B_{11} is invertible; for example, take $u_0=0$ and B_{10} invertible.

We now inductively prove that B_{1n} is nondegenerate for any $n \geqq 0$. Assuming B_{1n} is nonsingular a.s., we show that B_{1n+1} is also. In other words, we need to prove that $P(N)=0$ implies $P(DN^c)=0$, where

$$N \triangleq \{\omega | \det B_{1n}=0\}, \qquad D \triangleq \{\omega | \det B_{1n+1}=0\}.$$

Suppose that the opposite were true, i.e., $P(N)=0$, but $P(DN^c)>0$.

From (A.2) we have

$$\det\left(B_{1n}+\frac{1}{r_n}(\xi_{n+1}+w_{n+1})u_n^\tau\right)=0 \quad \forall \omega \in DN^c$$

but $\det B_{1n} \neq 0$ for $\omega \in DN^c$; hence

$$\det\left(I+\frac{1}{r_n}B_{1n}^{-1}(\xi_{n+1}+w_{n+1})u_n^\tau\right)=0 \quad \forall \omega \in DN^c$$

or

$$\det\left(1+\frac{1}{r_n}u_n^\tau B_{1n}^{-1}(\xi_{n+1}+w_{n+1})\right)=0 \quad \forall \omega \in DN^c$$

by part (1) of the lemma.

Then we have

$$u_n^\tau B_{1n}^{-1}(\xi_{n+1}+w_{n+1})=-r_n \quad \forall \omega \in DN^c \tag{A.3}$$

and consequently,

$$u_n^\tau B_{1n} \neq 0 \quad \forall \omega \in DN^c \tag{A.4}$$

since $r_n \geqq 1$.

We denote by $\alpha_i(\omega)$ and $w_{n+1,i}$ the components of $u_n^\tau B_{1n}^{-1}$ and w_{n+1} respectively, i.e.,

$$u_n^\tau B_{1n}^{-1}=[\alpha_1(\omega), \cdots, \alpha_m(\omega)], \tag{A.5}$$

$$w_{n+1}=[w_{n+1,1}, \cdots, w_{n+1,m}]^\tau. \tag{A.6}$$

Then from (A.3), (A.5) and (A.6) we have

$$\sum_{i=1}^m \alpha_i(\omega)w_{n+1,i}+r_n+u_n^\tau B_{1n}\xi_{n+1}=0 \quad \forall \omega \in DN^c. \tag{A.7}$$

From (A.4) and the assumption $P(DN^c)>0$ we would have some $\alpha_i(\omega)$ and a subset $D_1 \subset DN^c$ such that

$$\alpha_i(\omega) \neq 0 \quad \forall \omega \in D_1, \qquad P(D_1)>0. \tag{A.8}$$

Without loss of generality, we assume $i=1$, and define the random variable $z(\omega)$:

$$z(\omega)=\begin{cases} \dfrac{1}{\alpha_1(\omega)}\left[\sum_{i=2}^m \alpha_i(\omega)w_{n+1,i}+r_n+u_n^{-1}B_{1n}^{-1}\xi_{n+1}\right], & \omega \in D_1, \\ 0, & \omega \in D_1^c, \end{cases}$$

which is clearly independent of $w_{n+1,1}$. By part (2) of the lemma, it follows that

$$P(w_{n+1,1}+z(\omega)=0)=0. \tag{A.9}$$

However, (A.7) and (A.8) would yield

$$P(w_{n+1,1}+z(\omega)=0) \geqq P(D_1)>0. \tag{A.10}$$

The contradiction obtained proves $P(DN^c)=0$, and hence the nonsingularity of $B_{1,n+1}$ a.s.

Now assume $m<l$.

Let

$$B_{1n} \triangleq [\overbrace{B_{1n}^1}^{m}, \overbrace{B_{1n}^2}^{l-m}]\}m, \qquad u_n^\tau \triangleq [\overbrace{u_n^{1\tau}}^{m}, \overbrace{u_n^{2\tau}}^{l-m}].$$

From (A.2) we see

$$B_{1n+1}^1 = B_{1n}^1 + \frac{1}{r_n}(\xi_{n+1}+w_{n+1})u_n^{1\tau}.$$

By an argument similar to that given for the $m=l$ case we can prove that B_{1n}^1 is invertible a.s. for any $n \geqq 1$ if φ_0, θ_0 are adequately chosen. Then (A.1) is equivalent to

$$\text{(A.11)} \qquad [I, (B_{1n}^1)^{-1}B_{1n}^2]u_n = (B_{1n}^1)^{-1}(y_{n+1}^* + v_n - \bar\theta_n^\tau \bar\varphi_n)$$

or

$$u_n^1 + (B_{1n}^1)^{-1}B_{1n}^2 u_n^2 = (B_{1n}^1)^{-1}(y_{n+1}^* + v_n - \bar\theta_n^\tau \bar\varphi_n).$$

Obviously, the solution of (A.11) can be expressed by

$$u_n = \begin{bmatrix} (B_{1n}^1)^{-1}(y_{n+1}^* + v_n - \bar\theta_n^\tau\bar\varphi_n - B_{1n}^2 u_n^2) \\ u_n^2 \end{bmatrix} \quad \text{a.s.}$$

with any $(l-m)$-dimensional and $\mathscr{F}_n$-measurable u_n^2. This means that for the case $m<l$ the control u_n satisfying (3.1) exists but it is not unique. □

Remark. Recently Caines and Meyn [4] also have shown the existence of u_n satisfying (2.10) for a one-dimensional case but under conditions different from those imposed here.

Appendix 2. Proof of lemmas.

Proof of Lemma 5. Due to the assumption $v_n=0$ for $n<0$, we have

$$\begin{aligned}\sum_{h=1}^N (H_N(z)v_n)(H_N(z)v_n)^\tau &= \sum_{i,j=0}^\infty H_i(N)\left(\sum_{n=1}^N v_{n-i}v_{n-j}^\tau\right)H_j^\tau(N) \\ &= \sum_{i,j=0}^\infty H_i(N)\left(\sum_{n=\max(i,j,1)}^N v_{n-i}v_{n-j}^\tau\right)H_j^\tau(N).\end{aligned}$$

Set

$$S_N(i,j) = \sum_{n=\max(i,j,1)}^N [v_{n-i}v_{n-j}^\tau - \delta_{ij}R_{n-i}],$$

$$R_n = Ev_nv_n^\tau = \begin{cases} \dfrac{1}{\log^{1/4} n} I, & n>1, \\ 0, & n \leqq 1. \end{cases}$$

Clearly, $S_N(i,j)$ is a martingale and by Burkholder inequality (Chow and Teicher [14]), C_r—inequality and Schwarz inequality we have

$$\begin{aligned} E\|S_N(i,j)\|^{2+\delta/2} &\leqq c_1 E\left(\sum_{n=\max(i,j,1)}^N \|v_{n-i}v_{n-j}^\tau - \delta_{ij}R_{n-i}\|^2\right)^{1+(\delta/4)} \\ &\leqq c_1 N^{\delta/4} E \sum_{n=\max(i,j,1)}^N \|v_{n-i}v_{n-j}^\tau - \delta_{ij}R_{n-i}\|^{2+(\delta/2)} \\ &\leqq c_2 N^{1+\delta/4} \quad \text{for any } i\geqq 0,\ j\geqq 0 \text{ and some } c_1>0,\ c_2>0. \end{aligned}$$

From here and the Hölder inequality it follows that for any $\varepsilon > 0$ and

$$\gamma \in \left(\frac{2+(\delta/4)}{2+(\delta/2)}, 1\right),$$

$$P\left\{\sum_{i,j=0}^{\infty} e^{-k_2(i+j)}\|S_N(i,j)\| > N^\gamma \cdot \varepsilon\right\}$$

$$\leqq \frac{1}{\varepsilon^{2+(\delta/2)} N^{\gamma(2+(\delta/2))}} E\left(\sum_{i,j=0}^{\infty} e^{-k_2(i+j)}\|S_N(i,j)\|\right)^{2+(\delta/2)}$$

$$\leqq c_3 \frac{1}{N^{\gamma(2+(\delta/2))}} E \sum_{i,j=0}^{\infty} (e^{-k_2(i+j)})^{1+(\delta/4)} \|S_N(i,j)\|^{2+(\delta/2)}$$

$$\leqq c_4 \cdot \frac{1}{N^{\gamma(2+(\delta/2))-(1+(\delta/4))}} \quad \text{for any } N \geqq 1 \text{ and some constants } c_3 > 0,\ c_4 > 0.$$

Then by the Borel-Cantelli lemma we see

$$\text{(A.12)} \qquad \overline{\lim_{N\to\infty}} \frac{1}{N^\gamma} \left\| \sum_{i,j=0}^{\infty} H_i(N) S_N(i,j) H_j^\tau(N) \right\| \leqq \lim_{N\to\infty} \frac{k_1^2}{N^\gamma} \sum_{i,j=0}^{\infty} e^{-k_2(i+j)} \|S_N(i,j)\| \xrightarrow[N\to\infty]{} 0.$$

Finally, we obtain the desired result

$$\lim_{N\to\infty} \frac{\log^{1/4} N}{N} \sum_{n=1}^{N} (H_N(z)v_n)(H_N(z)v_n)^\tau$$

$$= \lim_{N\to\infty} \frac{\log^{1/4} N}{N} \sum_{i,j=0}^{\infty} (H_i(N) S_N(i,j) H_j^\tau(N)$$

$$+ \lim_{N\to\infty} \frac{\log^{1/4} N}{N} \sum_{i,j=0}^{\infty} H_i(N) \sum_{n=\max(i,j,1)}^{N} \delta_{ij} R_{n-i} H_j^\tau(N)$$

$$= \lim_{n\to\infty} \frac{\log^{1/4} N}{N} \sum_{i=0}^{N} H_i(N) \sum_{n=\max(i,1)}^{N} R_{n-i} H_i^\tau(N)$$

$$= \lim_{N\to\infty} \frac{\log^{1/4} N}{N} \left[H_0(N) \sum_{n=1}^{N} R_n H_0^\tau(N) + \sum_{i=1}^{N} H_i(N) \sum_{n=0}^{N-i} R_n H_i^\tau(N) \right]$$

$$= \lim_{N\to\infty} \frac{\log^{1/4} N}{N} \left[H_0(N) \sum_{n=1}^{N} R_n H_0^\tau(N) \right.$$

$$\left. + \sum_{i=1}^{N} H_i(N) \left(\sum_{n=1}^{N} R_n + R_0 - \sum_{n=N-i+1}^{N} R_n \right) H_i^\tau(N) \right]$$

$$= \lim_{N\to\infty} \frac{\log^{1/4} N}{N} \sum_{i=0}^{N} H_i(N) \sum_{n=1}^{N} R_n H_i^\tau(N). \qquad \square$$

Proof of Lemma 6. Set

$$S_N(i,j) \triangleq \sum_{n=1}^{N} w_{n+1-l-i} v_{n-m-j}^\tau.$$

Similar to the proof of (A.12), one can easily be convinced that

$$\lim_{N\to\infty} \frac{1}{N^\gamma} \sum_{i,j=0}^{\infty} H_i S_N(i,j) G_j^\tau = 0,$$

which is tantamount to (5.1).

Clearly, (5.2) can be verified in similar fashion.

By setting $H_N(z)\equiv H(z)$ and $v_n\equiv w_{n+1-l}$ in (A.12) we have

$$\lim_{N\to\infty}\frac{1}{N^\gamma}\sum_{i,j=0}^{\infty}H_iS_N(i,j)H_j^\tau=0$$

where

$$S_N(i,j)=\sum_{n=\max(i,j,1)}^{N}[w_{n-l+1-i}w_{n-l+1-j}^\tau-\delta_{ij}R_{n-l+1-i}]$$

and

$$R_{n-l+1-i}\triangleq Ew_{n-l+1-i}w_{n-l+1-i}^\tau.$$

Hence by the uniform boundedness of R_n we have

$$\begin{aligned}&\overline{\lim_{N\to\infty}}\frac{1}{N}\sum_{n=1}^{N}\|H(z)w_{n+1-l}\|^2\\&=\overline{\lim_{N\to\infty}}\operatorname{tr}\left[\frac{1}{N}\sum_{n=1}^{N}(H(z)w_{n+1-l})(H(z)w_{n+1-l})^\tau\right]\\&=\overline{\lim_{N\to\infty}}\operatorname{tr}\left[\frac{1}{N}\sum_{i,j=0}^{\infty}H_iS_N(i,j)H_j^\tau+\sum_{i,j=0}^{\infty}H_i\frac{1}{N}\sum_{n=\max(i,j,1)}^{N}\delta_{ij}R_{n-l+1-i}H_j^\tau\right]\\&=\overline{\lim_{N\to\infty}}\operatorname{tr}\sum_{i=0}^{\infty}H_i\frac{1}{N}\sum_{n=\max(i,1)}^{N}R_{n-l+1-i}H_i^\tau<\infty.\end{aligned}$$

This completes the proof of the lemma. □

REFERENCES

[1] K. J. ÅSTRÖM AND B. WITTENMARK, *On self-tuning regulators*, Automatica—J. IFAC, 9 (1973), pp. 185–195.

[2] A. BECKER, P. R. KUMAR AND C. Z. WEI, *Adaptive control with the stochastic approximation algorithm—Geometry and convergence*, IEEE Trans. Automat. Control, AC-30 (1985), pp. 330-338.

[3] P. E. CAINES AND S. LAFORTUNE, *Adaptive control with recursive identification for stochastic linear systems*, IEEE Trans. Automat. Control, AC-29 (1984), pp. 312-321.

[4] P. E. CAINES AND S. MEYN, *On the zero divisor problem and singularities occurring in the recursive schemes of stochastic adaptive control*, Systems Control Lett., 6 (1985), pp. 235-238.

[5] H. F. CHEN, *Quasi-least-squares identification and its strong consistency*, Internat. J. Control, 34 (1981), pp. 921-936.

[6] ———, *Strong consistency of recursive identification under correlated noise*, J. Systems Sci. Math. Sci., 1 (1981), pp. 34-52.

[7] ———, *Recursive system identification and adaptive control by use of the modified least squares algorithm*, this Journal, 22 (1984), pp. 758-776.

[8] ———, *Recursive Estimation and Control for Stochastic Systems*, John Wiley, New York, 1985.

[9] H. F. CHEN AND P. E. CAINES, *Strong consistency of the stochastic gradient algorithm of adaptive control*, IEEE Trans. Automat. Control, AC-30 (1985), pp. 189-192.

[10] H. F. CHEN AND L. GUO (1984), *Adaptive control with recursive identification for stochastic linear systems*, in Advances in Control and Dynamic Systems, C. T. Leondes, ed., Vol. 24, Academic Press, New York, to appear.

[11] ———, *Strong consistency of recursive identification by no use of persistent excitation condition*, Acta Math. Appl. Sinica, to appear.

[12] ———, *Strong consistency of parameter estimates for discrete-time stochastic systems*, J. Systems Sci. Math. Sci., 5 (1985), pp. 81-93.

[13] H. F. CHEN AND L. GUO, *The limit of stochastic gradient algorithm for identifying systems excited not persistently*, Kexue Tongbao (Science Bulletin), to appear.

[14] Y. S. CHOW AND H. TEICHER (1978), *Probability Theory*, Springer, New York.

[15] G. C. GOODWIN, P. T. RAMADGE AND P. E. CAINES, *Discrete-time stochastic adaptive control*, this Journal, 19 (1981), pp. 829-853.

[16] T. KAILATH, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[17] H. J. KUSHNER AND D. S. CLARK (1978), *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer, New York.

[18] L. LJUNG, *Consistency of the least squares identification method*, IEEE Trans. Automat. Control, AC-22 (1976), pp. 551-575.

[19] ———, *Analysis of recursive sotchastic algorithms*, IEEE Trans. Automat. Control, AC-22 (1977), pp. 551-575.

[20] J. B. MOORE, *On strong consistency of least squares identification algorithm*, Automatica—J. IFAC, 14 (1978), pp. 505-509.

[21] K. S. SIN AND G. C. GOODWIN, *Stochastic adaptive control using a modified least squares algorithm*, Automatica—J. IFAC, 18 (1982), pp. 315-321.

[22] V. SOLO, *The convergence of* AML, IEEE Trans. Automat. Control, AC-24 (1979), pp. 958-962.