SIAM J. CONTROL AND OPTIMIZATION
Vol 28 No 6 pp 1432-1451 November 1990

© 1990 Society for Industrial and Applied Mathematics
009

ON ADAPTIVE STABILIZATION OF TIME-VARYING STOCHASTIC SYSTEMS*

LEI GUO†

Abstract The basic stability issue of time-varying stochastic systems under adaptive control is studied A difficulty arising from treating the stochastic case as compared to the deterministic case is the lack of an a priori upper bound on the sample paths of the random noise sequence A projected gradient algorithm with small stepsize is used, avoiding possible large deviations of the estimates. It is shown that if the unknown parameters vary slowly in some sense, then an adaptive control law can be designed so that the closed-loop system is stable Issues of performance and robustness are also discussed

Key words Stochastic systems, adaptive control, random parameters stability, short memory projection

AMS(MOS) subject classifications primary 93C40, secondary 93E15, 62M10

1. Introduction

1. Introduction. The main objective in adaptive control theory is to design controllers that perform satisfactorily for systems which possess time-varying structure. However, the primary issue is to maintain closed-loop stability

Over the past two decades, the area of adaptive control can roughly be divided into two directions: deterministic and stochastic. In deterministic adaptive control, the system under study is normally assumed to be subjected to no noise, or at most a uniformly bounded disturbance. When the adaptive controller is designed based on deterministic methods, optimality of the performance cannot be guaranteed for time-invariant plants with a nice uniformly bounded white noise disturbance. This is due to the fact that the algorithms used have the so-called *short memory property*, i.e., the adaptation gain is not vanishing Nevertheless, algorithms of this kind have the merit that they may stabilize a time-varying system, as has been shown recently in, e.g., Tsakalis and Ioannou (1986) and Middleton and Goodwin (1988) in the deterministic framework.

In stochastic adaptive control, noise is an essential feature of the system, and it is not necessarily bounded; a standard example is the Gaussian white noise sequence In this case, especially for the constant parameter case, it is of interest not only to guarantee stability of the closed-loop system, but to reject the noise optimally, or at least close to optimally. This is possible, because the algorithms normally used have the so-called *long memory property*, i.e., the adaptation gain tends to zero This guarantees that no large deviations of the estimates can occur, at least for the constant parameter case. Indeed, it has been shown that in the constant parameter case, the parameter estimates in a closed-loop adaptive system can be either nearly consistent (Becker, Kumar, and Wei (1985)) or strongly consistent (Chen and Guo (1987)) However, it is this long memory property that prevents the adaptive law from being effective for general time-varying systems Indeed, with long memory algorithms, it has been found that it is difficult to deal with time-variations which are more complicated than, for instance, those treated in Chen and Caines (1985) and Chen and Guo (1988) For these reasons it is believed that short memory algorithms may be more effective than the long memory ones in the control of more realistically modeled time-varying systems

* Received by the editors February 15, 1989; accepted for publication (in revised form) January 9, 1990

† Institute of Systems Science, Academia Sinica, Beijing 100080, People s Republic of China This work was completed while the author was with the Department of Systems Engineering, Research School of Physical Sciences, The Australian National University, Canberra Australia

Let us illustrate the difference between stability studies of deterministic systems and that of stochastic systems by the following example:

$$y_{k+1} = \theta_k y_k + v_{k+1}, \qquad y_0 \neq 0, \tag{1.1}$$

$$\theta_{k+1} = \alpha\theta_k + \varepsilon_k, \qquad |\alpha| < 1. \tag{1.2}$$

Assume first that $\{v_k\}$ and $\{\varepsilon_k\}$ are deterministic sequences. It is then easy to verify the following assertion: $\{y_k\}$ is bounded for any bounded sequence $\{v_k\}$ and any sequence $\{\varepsilon_k\}$ satisfying $\sup|\varepsilon_k| \leqq \sigma$ if and only if $\sigma(1-\alpha)^{-1} < 1$. Next, let us assume that $\{v_k\}$ and $\{\varepsilon_k\}$ are independent white noise sequences. In this case, necessary conditions for the boundedness of $E\|y_k\|^2$ are discussed in, e.g., Granger and Andersen (1978) and Pourahmadi (1986). However, general sufficient conditions are hard to find even for this seemingly simple problem (see Pourahmadi (1986) for related discussions). One of the difficulties is due to the possible unboundedness of the process noise. Thus, stabilizing the first-order stochastic model (1.1)–(1.2) seems to be a nontrivial task.

By injecting an adaptive control signal to the right-hand side of model (1.1), Meyn and Caines (1987) showed that (1.1) is stabilizable if α is known, $|\sigma| < 1$, and if $\{v_k\}$ and $\{\varepsilon_k\}$ are independent Gaussian white noise sequence with known variances. The noise assumptions on $\{v_k, \varepsilon_k\}$ were subsequently relaxed in Guo and Meyn (1989) by imposing only moment conditions.

Let us now consider (1.1) again but with the unknown parameter $\{\theta_k\}$ a constant plus a first-order moving average process:

$$\theta_k = \theta + \varepsilon_k + d_1\varepsilon_{k-1}, \qquad k \geqq 0. \tag{1.3}$$

Assume that $\{v_k\}$ and $\{\varepsilon_k\}$ are independent Gaussian white noise sequences with $E|\varepsilon_k|^2 = \sigma^2 > 0$ and $d_1 > 0$. Then for second-order stability of (1.1), it is necessary that (see Tjostheim (1986, p. 60))

$$\theta^2 + [1+(d_1)^2]\sigma^2 + 2(d_1)^2\sigma^4 < 1,$$

which implies that $|\theta| < 1$ and that σ should be suitably small. In practice, it is acceptable to assume that the noise variance σ^2 is small. However, assuming the undisturbed parameter θ to be small or less than one is generally not applaudable. Again, to make the unstable open-loop time-varying stochastic system (1.1) and (1.3) stable, the use of stochastic adaptive control techniques seems to be necessary and appealing. This problem is solved as a simple example of Theorem 1 stated later in § 3.

In this paper, we consider the basic stability issue of general time-varying stochastic systems under adaptive control. The assumptions on the random noise include two important cases: bounded sequences and Gaussian sequences. We will study two classes of SISO stochastic models, although generalizations to MIMO and some other classes are straightforward. In the first class (Model 1), the parameters are assumed to be random, and only parameters in the autoregressive part are estimated, while in the second class (Model 2) the parameters are assumed to be deterministic, and parameters in both the autoregressive and exogenous parts are estimated. The remainder of the paper is organized as follows. In § 2 we describe the stochastic models that will be studied in the paper. The main stability results are stated in § 3. Section 4 establishes some inequalities and stability results for general stochastic sequences. In § 5 we present the proofs for theorems. Further discussions on performance and robustness are given in § 6. Section 7 concludes the paper.

2. Stochastic models. In this paper we will mainly consider the following two classes of time-varying stochastic models.

Model 1 (random parameter model).

$$(2.1)\qquad \begin{aligned} &y_{k+1}=a_1(k)y_k+\cdots+a_p(k)y_{k-p+1}+u_k+v_{k+1}, \qquad k\geqq 0,\\ &y_k=u_k=v_k=0 \quad \forall k<0, \end{aligned}$$

where y_k, u_k, and v_k are the scalar output, input, and random noise processes, respectively, and $a_i(k)$, $1\leqq i\leqq p$, are the unknown random time-varying parameters.

Model 2 (deterministic parameter model).

$$(2.2)\qquad \begin{aligned} &y_{k+1}=a_1(k)y_k+\cdots+a_s(k)y_{k-s+1}+b_1(k)u_k+\cdots+b_t(k)u_{k-t+1}+v_{k+1}, \qquad k\geqq 0,\\ &y_k=u_k=v_k=0 \quad \forall k<0, \end{aligned}$$

where $a_i(k)$, $b_j(k)$, $1\leqq i\leqq s$, $1\leqq j\leqq t$, are the unknown deterministic time-varying parameters.

Note that both Models 1 and 2 can be rewritten in the following regression form:

$$(2.3)\qquad z_{k+1}=\varphi_k^\tau\theta_k+v_{k+1},$$

where for Model 1, $z_{k+1}=y_{k+1}-u_k$,

$$(2.4)\qquad \varphi_k=[y_k\cdots y_{k-p+1}]^\tau, \qquad \theta_k=[a_1(k)\cdots a_p(k)]^\tau,$$

while for Model 2, $z_{k+1}=y_{k+1}$, and

$$(2.5)\qquad \varphi_k=[y_k\cdots y_{k-s+1}, u_k\cdots u_{k-t+1}]^\tau, \qquad \theta_k=[a_1(k)\cdots a_s(k), b_1(k)\cdots b_t(k)]^\tau.$$

Let us now introduce the assumptions on the random noise sequence $\{v_k\}$.

Noise assumption. $\{v_k, F_k\}$ is an adapted sequence where $\{F_k\}$ is a nondecreasing family of σ-algebras, and for some integer $r\geqq 0$ and deterministic positive constants ε and M_v:

$$(2.6)\qquad E\{\exp[\varepsilon\|v_{k+1}\|^2]|F_{k-r}\}\leqq \exp\{M_v\} \quad \text{a.s.} \quad \forall k\geqq 0$$

Obviously, any sequence $\{v_k\}$ which is uniformly bounded in sample path satisfies this assumption. We note also that if $\{v_k\}$ is an r-dependent sequence (i.e., for any k, $\{v_i, i\leqq k\}$ and $\{v_{i+r}, i>k\}$ are independent), then the above assumption (2.6) reduces to

$$(2.7)\qquad E\{\exp[\varepsilon|v_{k+1}|^2]\}\leqq \exp\{M_v\} \quad \text{a.s.} \quad \forall k\geqq 0.$$

Let us now give an example where the noise sequence $\{v_k\}$ is unbounded almost surely

Example 1. Let $\{v_k\}$ be the following time-varying moving average process:

$$(2.8)\qquad v_k=e_k+c_1(k)e_{k-1}+\cdots+c_r(k)e_{k-r}, \qquad k\geqq 0,$$

with deterministic coefficients $\{c_i(k)\}$ satisfying

$$(2.9)\qquad \sum_{i=0}^{r}|c_i(k)|^2\leqq c<\infty \quad \forall k\geqq 0, \quad (c_0(k)=1),$$

assuming that $\{e_k\}$ is a Gaussian white noise sequence with variance $\sigma^2>0$. Then

$$(2.10)\qquad \limsup_{k\to\infty}\frac{|v_k|}{(2\log k)^{1/2}}\geqq\sigma \quad \text{a.s.}$$

and the noise assumption (2.6) holds for any

$$(2.11)\qquad \varepsilon<\frac{1}{2c\sigma^2}, \qquad M_v\geqq\frac{\varepsilon c\sigma^2(r+1)}{1-2\varepsilon c\sigma^2}$$

Proof. Property (2.10) follows from the conditional Borel–Cantelli lemma and the Gaussian assumption; details of the proof are omitted (see also Chow and Teicher (1978, p. 64) for a related result). Here, we will only prove that (2.6) is true for any constants ε and M_v satisfying (2.11).

Apparently, $\{v_k\}$ is an r-dependent sequence, so we need only to verify (2.7). By elementary calculations, it is easy to verify that

$$E \exp\{\varepsilon|v_k|^2\} \leqq \{E \exp[\varepsilon c(e_1)^2]\}^{r+1} \leqq \exp\left\{\frac{\varepsilon c\sigma^2}{1-2\varepsilon c\sigma^2}(r+1)\right\}. \tag{2.12}$$

Hence by (2.11) and (2.12), we see that (2.7) is true.

We remark that in the above example, the constants ε and M_v depend only on the *upper bounds* of σ, c, and r.

3. Main results. Since the conditions imposed on the time-varying parameters of Models 1 and 2 are quite different, we will consider these two models separately

3.1. Random parameter case. The assumptions on the parameters of Model 1 are as follows.

Parameter assumption (random case) $\{\theta_k, F_k\}$ defined in (2.4) is an adapted sequence which satisfies

$$E\{\exp[M\|\theta_{k+1}\|^2] | F_{k-m}\} \leqq \exp\{M_\theta\} \quad \text{a.s.} \quad \forall k \geqq 0, \tag{3.1}$$

$$E\{\exp[M\|w_{k+1}\|^2] | F_{k-m}\} \leqq \exp\{\delta_\theta\} \quad \text{a.s.} \quad \forall k \geqq 0, \tag{3.2}$$

where w_{k+1} is the parameter variation process:

$$w_{k+1} = \theta_{k+1} - \theta_k, \qquad k \geqq 0, \tag{3.3}$$

and where $m \geqq 0$ is an integer and M, M_θ, and $\delta_\theta < 1$ are positive deterministic constants.

We now discuss this condition. Condition (3.1) means that the random process $\{\theta_k\}$ is bounded in an average sense and not necessarily bounded in sample path. In the main theorems to follow, we will actually need that the constant M is suitably large and that δ_θ is suitably small (see Remark 3.1), which means that the parameters are slowly varying in an average sense, and again, the variation is not necessarily small in sample path. In particular, these conditions do not rule out occasional but possibly large jumps of the parameter process. Let us give a concrete example.

Example 2. Let the unknown parameter θ_k be a constant vector plus a p-dimensional moving average process:

$$\theta_k = \theta + \varepsilon_k + D_1\varepsilon_{k-1} + \cdots + D_{m-1}\varepsilon_{k-m+1}, \qquad k \geqq 0, \tag{3.4}$$

where D_i, $1 \leqq i \leqq m-1$, are deterministic matrices, and $\{\varepsilon_k\}$ is a Gaussian white noise sequence with covariance matrix $(\sigma_\varepsilon)^2 I$. Then for any $\sigma_\varepsilon > 0$,

$$\limsup_{k\to\infty} \|\theta_k\| = \infty \quad \text{a.s.}, \qquad \limsup_{k\to\infty} \|\theta_k - \theta_{k-1}\| = \infty, \quad \text{a.s.} \tag{3.5}$$

Furthermore, the above parameter assumption holds for all small σ_ε.

Proof. We need only to verify (3.1) and (3.2) here. Note that both the process $\{\theta_k\}$ and its variation process

$$w_k = \varepsilon_k + (D_1 - I)\varepsilon_{k-1} + \cdots + (D_{m-1} - D_{m-2})\varepsilon_{k-m+1} - D_{m-1}\varepsilon_{k-m} \tag{3.6}$$

are m-dependent sequences, so it suffices to verify (3.1) and (3.2) with conditional expectation replaced by expectation.

Similar to the proof of Example 1, we have for any constant $M>0$,

$$E\{\exp[M\|\theta_k\|^2]\}\leqq \exp\left\{2M\left[\|\theta\|^2+\frac{pmd_0(\sigma_\varepsilon)^2}{1-4Md_0(\sigma_\varepsilon)^2}\right]\right\},$$

$$E\exp\{M\|w_k\|^2\}\leqq \exp\left\{\frac{p(m+1)Md_1(\sigma_\varepsilon)^2}{1-2Md_1(\sigma_\varepsilon)^2}\right\},$$

where

$$d_0=\sum_{i=0}^{m-1}\|D_i\|^2,\quad d_1=1+\sum_{i=1}^{m}\|D_i-D_{i-1}\|^2,\quad (D_0=I,\ D_m=0).$$

Hence (3.1) and (3.2) hold. □

We now describe the estimation algorithm. Let $L>0$ and $d>0$ be two constants (which will be specified later). We define D as the following bounded domain:

(3.7) $$D=\{x=(x_1,\cdots,x_p)\in R^p: |x_i|\leqq L,\ 1\leqq i\leqq p\}$$

and $\pi_D\{x\}$ as the nearest point from x to D (under the Euclidean norm).

The estimate for the unknown process $\{\theta_k\}$ is generated by the following projected version of the gradient algorithm:

(3.8) $$\hat{\theta}_{k+1}=\pi_D\left\{\hat{\theta}_k+\frac{\varphi_k}{d+\|\varphi_k\|^2}(y_{k+1}-u_k-\varphi_k^\tau\hat{\theta}_k)\right\}$$

with arbitrary initial condition $\hat{\theta}_0\in D$, where φ_k is defined as in (2.4).

We remark that the use of a projection in estimation algorithms is common in the literature (e.g., Ljung and Soderstrom (1983), Goodwin and Sin (1984)). However, in estimating the parameters of stochastic systems by short memory algorithms, this procedure seems to be particularly important, since otherwise large deviations of the estimates are inevitable even if the system is persistently excited (see, e.g., Guo, Moore, and Xia (1988)). We also note that due to the special form of the domain D, the calculation of the projection in (3.8) is straightforward.

The certainty equivalent minimum variance adaptive control law is

(3.9) $$u_k=-\varphi_k^\tau\hat{\theta}_k.$$

Our first stability result is the following theorem.

THEOREM 1. *For the random parameter model* (2.1), *if the noise assumption* (2.6) *and the parameter assumptions* (3.1)–(3.2) *hold for suitably large* M *and small* δ_θ, *and if in the estimation algorithm* (3.7)–(3.8), L *and* d *are taken appropriately large, then under the adaptive control law* (3.9), *the closed-loop system is stable in the sense that*

(3.10a) $$\limsup_{n\to\infty} E\{|y_n|^\beta+|u_n|^\beta\}<\infty,$$

(3.10b) $$\limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N}\{|y_n|^2+|u_n|^2\}<\infty\quad a.s.,$$

where $\beta>2$ *is a constant depending on* M, δ_θ, L, *and* d.

Remark 3.1. We may ask how large (small) the constant M (δ_θ) is required to be in the above theorem. In § 5, we will prove that Theorem 1 is true when M and δ_θ satisfy the following inequality:

$$M\geqq 3(m+1)2^7\beta p^{3/2}\lambda^{-p}$$

and

(3.11)
$$\delta_\theta < \min\left\{1, \left[\frac{\beta(m+1)}{2}(\log\lambda^{-1})\right]^2\left[\frac{f(2^7(m+1)\beta p^{3/2}L_0\lambda^{-p}M^{-1/2})}{2} + \frac{2^5\beta p\lambda^{-p}(m+1)}{M}\right]^{-2}\right\}$$

for some $\lambda\in(0,1)$ and $\beta>2$, where the function $f(\cdot)$ is defined as

(3.12)
$$f(x)=\tfrac{1}{2}+x+4x^2[1+\exp(8x^2)]\quad\forall x$$

and L_0 denotes

(3.13)
$$L_0=\left\{\frac{4M_\theta}{M}+\lambda^p[96(m+1)\beta p]^{-1}\left|\log\left[\frac{(m+1)\beta}{2}\log(\lambda^{-1})\right]\right|\right\}^{1/2}$$

Moreover, in practical implementations of the algorithm, it is desirable to know the values of L and d. It will also be proved in § 5 that one way to choose L and d is

(3.14)
$$L=L_0,$$
$$d>16p(\varepsilon\lambda^p)^{-1}\max\{8M_v(\log\lambda^{-1})^{-1}, 4\beta(r+1)\}$$

3.2. **Deterministic parameter case.** The assumptions on the parameters of Model 2 are as follows.

Parameter assumption (deterministic case). (i) There is a positive constant $b_1>0$, such that

(3.15)
$$b_1(k)\geqq b_1\quad\forall k\geqq 0,$$

and the model (2.2) is uniformly stably invertible in the sense that there are two constants $A>0$, $\rho\in(0,1)$ such that

(3.16)
$$|u_k|^2\leqq A\sum_{i=0}^{k+1}\rho^{k+1-i}\{|y_i|^2+|v_i|^2\}\quad\forall k.$$

(ii) The parameter is slowly varying in the sense that

(3.17)
$$\|\theta_k\|\leqq M_1,\quad \|\theta_{k+1}-\theta_k\|\leqq\delta_1\quad\forall k\geqq 0,$$

where

(3.18)
$$M_1<\infty,\qquad \delta_1<\min\{1, \log(\lambda^{-1})[24K_1\lambda^{-1}(4(s+t)^{1/2}M_1+1)]^{-1}\},$$
$$K_1=s\lambda^{-s+1}[1+s(M_1/b_1)^2]+(t-1)\lambda^{-t+1}[1+(t-1)(M_1/b_1)^2]A\lambda^2/(\lambda-\rho),$$

and $\lambda\in(\rho,1)$ is some constant.

We remark that since $\{\theta_k\}$ is bounded, the assumption (3.16) is implied by uniform asymptotic stability of the following time-varying polynomial:

(3.19)
$$B_k(z)=b_1(k)+b_2(k)z+\cdots+b_t(k)z^{t-1},$$

which in the constant parameter case is the standard minimum phase condition.

Let us introduce the following bounded domain:

(3.20)
$$D=\{x=(x_1,\cdots,x_{s+t})\in R^{s+t}: |x_i|\leqq L, 1\leqq i\leqq s+t, x_{s+1}\geqq b_1\}.$$

The estimation algorithm is also a projected gradient one:

(3.21)
$$\hat\theta_{k+1}=\pi_D\left\{\hat\theta_k+\frac{\varphi_k}{d+\|\varphi_k\|^2}(y_{k+1}-\varphi_k^\tau\hat\theta_k)\right\},$$

where the initial condition $\hat\theta_0\in D$, and φ_k is defined as in (2.5).

The certainty equivalent minimum variance adaptive control u_k at any time k is solved from the following simple equation:

$$\varphi_k^\tau \hat{\theta}_k = 0. \tag{3.22}$$

Similar to Theorem 1, we have the following result.

THEOREM 2. *For the deterministic parameter model* (2.2), *suppose that the noise assumption* (2.6) *and the parameter assumptions* (3.15)–(3.18) *hold, and that in the estimation algorithm* (3.20)–(3.21), L *is taken as* M_1 *appearing in* (3.17) *and*

$$d > 36K_1(\lambda\varepsilon_0)^{-1} \max\{\beta(r+1), 2M_v[\log(\lambda^{-1})]^{-1}\} \tag{3.23}$$

for some $\beta > 2$. *Then under the adaptive control law* (3.22), *the closed-loop system is stable in the sense that*

$$\limsup_{n\to\infty} E\{|y_n|^\beta + |u_n|^\beta\} < \infty, \tag{3.24a}$$

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \{|y_n|^2 + |u_n|^2\} < \infty \quad a.s. \tag{3.24b}$$

We remark that a precise upper bound for the left-hand side of (3.24a), (3.24b) may be found in the proof—see § 5.

4. General lemmas. For the proof of theorems, we need some inequalities and stability results for stochastic sequences, which we will present in this section.

LEMMA 4.1 (i) (Bellman–Gronwall inequality). *Let* $\{x_k\}$, $\{f_k\}$, *and* $\{h_k\}$ *be three nonnegative sequences, and*

$$x_k \leqq f_k + \sum_{i=0}^{k-1} h_i x_i, \qquad k \geqq 0;$$

then

$$x_k \leqq f_k + \sum_{i=0}^{k-1} \prod_{j=i}^{k-1} (1+h_j) f_i, \qquad k \geqq 0. \tag{4.1}$$

(ii) *Let* $\{x_n, F_n\}$ *be an adapted sequence, and for some integer* $r \geqq 0$ *and some* $\alpha > 1$,

$$\sup_n E\{|x_{n+1}|^\alpha \mid F_{n-r}\} < \infty \quad a.s.; \tag{4.2}$$

then

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} |x_n| < \infty \quad a.s.$$

Proof. The first result is well known and can be easily proved by induction (see, e.g., Desoer and Vidyasagar (1975, p. 254)). As for the second result, we first note that for any fixed k, $0 \leqq k \leqq r$, the sequence

$$M_n = |x_{k+n(r+1)}| - E\{|x_{k+n(r+1)}| \mid F_{k+(n-1)(r+1)}\}$$

is a martingale difference sequence with respect to $\{F_{k+n(r+1)}\}$. Hence by (4.2) and Chow's martingale convergence theorem (see Stout (1974, p. 137)), we know that

$$\frac{1}{N} \sum_{n=0}^{N} M_n \to 0 \quad \text{a.s.} \quad \text{as } N \to \infty$$

Consequently, by (4.2) again,

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} |x_{k+n(r+1)}| < \infty \quad \text{a.s.} \quad \forall k \in [0, r].$$

Finally, the desired result follows by observing

$$\frac{1}{N} \sum_{n=1}^{N} |x_n| \leqq \frac{1}{N} \sum_{k=0}^{r} \sum_{n=0}^{[(N+1)/(r+1)]} |x_{k+n(r+1)}|,$$

where $[(N+1)/(r+1)]$ is the integer part of $(N+1)/(r+1)$. □

We also need the following lemma.

LEMMA 4.2. *Let* $\{x_n, F_n\}$, $\{f_n, F_n\}$, *and* $\{g_n, F_n\}$ *be three adapted nonnegative sequences satisfying*

$$x_{n+1} \leqq f_{n+1} x_n + g_{n+1} \quad \forall n \geqq 0. \tag{4.3}$$

Assume that for some constants $\varepsilon_\alpha < 1$, $\alpha > 1$, *and* $C < \infty$,

$$\sup_n E\{(f_{n+1})^\alpha | F_n\} \leqq \varepsilon_\alpha \quad \text{a.s.} \quad \sup_n E[(g_{n+1})^\alpha | F_n] \leqq C. \tag{4.4}$$

Then

$$\sum_{n=0}^{N} x_n = O(N), \quad \text{a.s} \quad \text{as } N \to \infty \tag{4.5}$$

Proof. Applying the Minkowski inequality to (4.3) and noting (4.4), we see that

$$\begin{aligned}\{E(x_{n+1})^\alpha\}^{1/\alpha} &\leqq \{E(f_{n+1}x_n)^\alpha\}^{1/\alpha} + \{E(g_{n+1})^\alpha\}^{1/\alpha}\\ &= \{E[E[(f_{n+1})^\alpha | F_n](x_n)^\alpha]\}^{1/\alpha} + \{E(g_{n+1})^\alpha\}^{1/\alpha}\\ &\leqq (\varepsilon_\alpha)^{1/\alpha}\{E(x_n)^\alpha\}^{1/\alpha} + \sup_n \{E(g_{n+1})^\alpha\}^{1/\alpha},\end{aligned}$$

from this and the fact that $(\varepsilon_\alpha)^{1/\alpha} < 1$, it is easy to conclude that

$$\sup_n E(x_n)^\alpha < \infty \tag{4.6}$$

Let us denote $M_n = x_n - E[x_n | F_{n-1}]$; then by (4.6) and the martingale stability results (Stout (1974, p. 137)), it is evident that

$$\sum_{n=0}^{N} M_n = o(N) \quad \text{a.s}$$

Thus by (4.4) and the recursion (4.3) we have (where ε_1 is defined as $(\varepsilon_\alpha)^{1/\alpha}$),

$$\begin{aligned}\sum_{n=0}^{N} x_{n+1} &= \sum_{n=0}^{N} E[x_{n+1} | F_n] + \sum_{n=0}^{N} M_{n+1}\\ &\leqq \sum_{n=0}^{N} E[f_{n+1} | F_n] x_n + \sum_{n=0}^{N} E[g_{n+1} | F_n] + o(N)\\ &\leqq \varepsilon_1 \sum_{n=0}^{N} x_n + O(N)\\ &\leqq \varepsilon_1 \sum_{n=0}^{N} x_{n+1} + \varepsilon_1 x_0 + O(N),\end{aligned}$$

consequently the assertion (4.5) holds since $\varepsilon_1 < 1$. □

LEMMA 4.3 *Let $\{f_n\}$ be a sequence of nonnegative random variables defined by*

$$f_n = \sum_{i=0}^{n-1} \lambda^{n-i} \prod_{j=i}^{n-1} x_j, \qquad f_0 = 0,$$

where $\lambda \in (0,1)$ and $\{x_k, F_k\}$ is a nonnegative adapted sequence satisfying $x_k \geqq 1$, and

$$\{E[(x_{k+1})^{\alpha(r+1)} | F_{k-r}]\}^{1/[\alpha(r+1)]} \leqq C \quad \text{a.s.} \quad \lambda C < 1, \tag{4.7}$$

for some integer $r \geqq 0$ and some constants $C > 0$ and $\alpha \geqq 1$, then,

$$\sup_n \{E[f_n]^\alpha\}^{1/\alpha} \leqq \lambda C^{r+2}(1-\lambda C)^{-1} \tag{4.8}$$

Moreover, if in (4.7) $\alpha > 1$, then as $N \to \infty$,

$$\sum_{n=0}^{N} f_n = O(N) \quad \text{a.s.} \tag{4.9}$$

Proof. By the Holder inequality, we have

$$\begin{aligned} E\left\{\prod_{j=i}^{n-1} x_j\right\}^\alpha &\leqq E\left\{\prod_{j=i}^{i+r} \prod_{k=0}^{[(n-i)/(r+1)]} [x_{j+k(r+1)}]^\alpha\right\} \\ &\leqq \prod_{j=i}^{i+r} \left\{E \prod_{k=0}^{[(n-i)/(r+1)]} [x_{j+k(r+1)}]^{\alpha(r+1)}\right\}^{1/(r+1)}, \end{aligned} \tag{4.10}$$

where $[(n-i)/(r+1)]$ is the integer part of $(n-i)/(r+1)$.

Note that for each i and j,

$$\begin{aligned} &E \prod_{k=0}^{[(n-i)/(r+1)]} [x_{j+k(r+1)}]^{\alpha(r+1)} \\ &= E \prod_{k=0}^{[(n-i)/(r+1)]-1} [x_{j+k(r+1)}]^{\alpha(r+1)} E\{[x_{j+[(n-i)/(r+1)](r+1)}]^{\alpha(r+1)} | F_{j+\{[(n-i)/(r+1)]-1\}(r+1)}\} \\ &\leqq C^{\alpha(r+1)} E \prod_{k=0}^{[(n-i)/(r+1)]-1} [x_{j+k(r+1)}]^{\alpha(r+1)} \leqq \cdots \\ &\leqq C^{\alpha(r+1)\{[(n-i)/(r+1)]+1\}} \leqq C^{\alpha(n-i+r+1)}. \end{aligned}$$

Substituting this into (4.10) we see that

$$E\left\{\prod_{j=i}^{n-1} x_j\right\}^\alpha \leqq C^{\alpha(n-i+r+1)}.$$

Consequently by the definition of f_n and the Minkowski inequality,

$$\begin{aligned} \{E[f_n]^\alpha\}^{1/\alpha} &\leqq \sum_{i=0}^{n-1} \lambda^{n-i} \left\{E \prod_{j=i}^{n-1} (x_j)^\alpha\right\}^{1/\alpha} \\ &\leqq C^{(r+1)} \sum_{i=0}^{n-1} (\lambda C)^{n-i} \leqq \lambda C^{r+2}(1-\lambda C)^{-1} \end{aligned}$$

We now prove (4.9). By the Holder inequality,

$$\begin{aligned} \sum_{n=0}^{N} f_n &= \sum_{n=0}^{N} \sum_{i=0}^{n-1} \lambda^{n-i} \prod_{j=i}^{n-1} x_j \\ &\leqq \sum_{n=0}^{N} \sum_{i=0}^{n-1} \lambda^{n-i} \prod_{j=0}^{r} \prod_{k=0}^{[(n-i)/(r+1)]} [x_{i+j+k(r+1)}] \\ &\leqq \prod_{j=0}^{r} \left\{\sum_{n=0}^{N} \sum_{i=0}^{n-1} \lambda^{n-i} \prod_{k=0}^{[(n-i)/(r+1)]} [x_{i+j+k(r+1)}]^{r+1}\right\}^{1/(r+1)}. \end{aligned} \tag{4.11}$$

Note that for each j

$$\sum_{i=0}^{n-1} \lambda^{n-i} \prod_{k=0}^{[(n-i)/(r+1)]} [x_{i+j+k(r+1)}]^{r+1} \le \sum_{s=0}^{r} \sum_{i=0}^{[n/(r+1)]} \lambda^{n-s-i(r+1)} \prod_{k=0}^{[(n-s)/(r+1)]-i} [x_{s+j+(i+k)(r+1)}]^{r+1}, \qquad \left(\prod_{k=0}^{-1} = 1\right). \tag{4.12}$$

Let us denote

$$g_n = \sum_{i=0}^{[n/(r+1)]} \lambda^{n-s-i(r+1)} \prod_{k=0}^{[(n-s)/(r+1)]-i} [x_{s+j+(i+k)(r+1)}]^{r+1}, \qquad 0 \le s, j \le r. \tag{4.13}$$

Similar to the proof of Lemma 4.1(ii), we consider the following subsequence of $\{g_n\}$ for any fixed $t \in [0, r]$, $0 \le s, j \le r$:

$$\begin{aligned} g_{t+n(r+1)} &= \sum_{i=0}^{[t/(r+1)]+n} \lambda^{t-s+(n-i)(r+1)} \prod_{k=0}^{[(t-s)/(r+1)]-i+n} [x_{s+j+(i+k)(r+1)}]^{r+1} \\ &\le \sum_{i=0}^{n} \lambda^{t-s+(n-i)(r+1)} \prod_{k=0}^{n-i} [x_{s+j+(i+k)(r+1)}]^{r+1} \\ &\triangleq M_n. \end{aligned} \tag{4.14}$$

It is obvious that $\{M_n, G_n\}$ is an adapted sequence, where $G_n = F_{s+j+n(r+1)}$. Note also that

$$M_n = [\lambda x_{s+j+n(r+1)}]^{r+1} M_{n-1} + \lambda^{t-s} [x_{s+j+n(r+1)}]^{r+1}$$

and that by the assumption (4.7),

$$\sup_n E\{[\lambda x_{s+j+n(r+1)}]^{(r+1)\alpha} \mid G_{n-1}\} \le (\lambda C)^{\alpha(r+1)} < 1 \quad \text{a.s.}$$

Hence applying Lemma 4.2, we have

$$\begin{aligned} &\sum_{n=0}^{N} M_n = O(N) \quad \text{a.s.} \\ \Rightarrow &\sum_{n=0}^{N} g_n = O(N) \quad \text{a.s.} \quad (\text{since in (4.14) } t \in [0, r] \text{ is arbitrary}) \\ \Rightarrow &\sum_{n=0}^{N} \sum_{i=0}^{n-1} \lambda^{n-i} \prod_{k=0}^{[(n-i)/(r+1)]} [x_{i+j+k(r+1)}]^{r+1} = O(N) \quad \text{a.s.} \quad (\text{by (4.12)}) \\ \Rightarrow &\sum_{n=0}^{N} f_n = O(N) \quad \text{a.s.} \quad (\text{by (4.11)}). \end{aligned}$$

This completes the proof. □

LEMMA 4.4. *Let w and F be any random variable and σ-algebra, respectively. If*

$$E\{\exp(w^2) \mid F\} \le \exp(\delta) \quad \textit{a.s.} \quad \textit{for some } \delta > 0,$$

then for any real number $a > 0$,

$$E\{\exp(a|w|) \mid F\} \le \exp\{a\delta^{1/2} + [\tfrac{1}{2} + 4a^2(1 + \exp(8a^2))]\delta\} \quad \textit{a.s.}$$

We remark that the key point in the above upper bound is the dependence on δ. If $\delta < 1$, then the above result implies that $E\{\exp(a|w|) \mid F\} \le \exp\{f(a)\delta^{1/2}\}$, where $f(\cdot)$ is the function defined by (3.12).

Proof. We first note that by the Jensen's inequality,

$$E\{\exp(w^2)|F\} \geqq \exp\{E[w^2|F]\} \quad \text{a.s.}$$

so it follows from the assumption that

$$E[w^2|F] \leqq \delta \quad \text{a.s.}$$

Next, we will use the following fact that can be proven in exactly the same way as that for Lemma 4.1.1 of Stout (1974, p. 226): For any random variable Y, if $0 \leqq Y \leqq 1$, almost surely, then

$$E\{\exp(Y)|F\} \leqq \exp\{E[Y|F] + E[Y^2|F]\}.$$

Applying this we have

$$E\{\exp[4a|w|I(4a|w| \leqq 1)|F]\} \leqq \exp\{4a\delta^{1/2} + (4a)^2\delta\}.$$

Hence, by this inequality, the Schwarz inequality and the Markov inequality, we have (where $E^F(\cdot)$ denotes $E(\cdot|F)$, for simplicity)

$$\begin{aligned}
&E^F \exp(a|w|) \\
&\quad = E^F \exp\{a|w|[I(|w| \geqq 2a) + I(|w| < 2a)]\} \\
&\quad \leqq E^F \exp\left(\frac{w^2}{2}\right) \exp\{a|w|I(|w| < 2a)\} \\
&\quad \leqq \exp\left(\frac{\delta}{2}\right)\left\{E^F \exp\left\{2a|w|\left[I\left(|w| \leqq \frac{1}{4a}\right) + I\left(\frac{1}{4a} < |w| < 2a\right)\right]\right\}\right\}^{1/2} \\
&\quad \leqq \exp\left(\frac{\delta}{2}\right)\left\{E^F \exp[4a|w|I(4a|w| \leqq 1)]E^F \exp\left[4a|w|I\left(\frac{1}{4a} < |w| < 2a\right)\right]\right\}^{1/4} \\
&\quad \leqq \exp\left(\frac{\delta}{2}\right)\{\exp[4a\delta^{1/2} + (4a)^2\delta]\}^{1/4}\left\{E^F \exp\left[4a|w|I\left(\frac{1}{4a} < |w| < 2a\right)\right]\right\}^{1/4} \\
&\quad \leqq \exp\left[\frac{\delta}{2} + a\delta^{1/2} + 4a^2\delta\right]\left\{E^F I\left(|w| \leqq \frac{1}{4a}\right)\right. \\
&\qquad \left. + E^F \exp[4a|w|I(|w| < 2a)]I\left(|w| > \frac{1}{4a}\right)\right\}^{1/4} \\
&\quad \leqq \exp\left[a\delta^{1/2} + \left(\frac{1}{2} + 4a^2\right)\delta\right]\left\{1 + \exp(8a^2)P\left(|w| > \frac{1}{4a}|F\right)\right\}^{1/4} \\
&\quad \leqq \exp\left[a\delta^{1/2} + \left(\frac{1}{2} + 4a^2\right)\delta\right]\{1 + (4a)^2\delta \exp(8a^2)\}^{1/4} \\
&\quad \leqq \exp\left\{a\delta^{1/2} + \left[\frac{1}{2} + 4a^2 + 4a^2 \exp(8a^2)\right]\delta\right\}.
\end{aligned}$$

This completes the proof. □

5. Proof of the theorems. The proofs of Theorems 1 and 2 are divided into several lemmas.

Let us denote

$$\alpha_k = \frac{\|\varphi_k^\tau \tilde{\theta}_k\|^2}{d + \|\varphi_k\|^2}, \qquad \tilde{\theta}_k = \theta_k - \hat{\theta}_k. \tag{5.1}$$

We have Lemma 5.1.

LEMMA 5.1. *Under conditions of Theorem* 1, *the following inequality holds for any* $k \geqq 0$:

$$\alpha_k \leqq 2(\|\tilde{\theta}_k\|^2 - \|\tilde{\theta}_{k+1}\|^2) + (8/d)\|v_{k+1}\|^2 + 4\{2p^{1/2}L + \|w_{k+1}\|\}\|w_{k+1}\| + 12\{p^{1/2}L + \|\theta_k\|\}\|\theta_k\| I(\theta_k \notin D),$$

where $I(A)$ *is the indicator function of a set* A.

Proof. Let us denote $\bar{\theta}_k = \theta_k I(\theta_k \in D)$. We have

$$\|\hat{\theta}_{k+1} - \bar{\theta}_k\|^2 \leqq 4pL^2$$

and

$$\|\theta_{k+1} - \bar{\theta}_k\| \leqq \|\theta_{k+1} - \theta_k\| + \|\theta_k I(\theta_k \notin D)\| \leqq \|w_{k+1}\| + \|\theta_k I(\theta_k \notin D)\|.$$

So we have

$$\begin{aligned} \|\hat{\theta}_{k+1} - \theta_{k+1}\|^2 &= \|\hat{\theta}_{k+1} - \bar{\theta}_k + \bar{\theta}_k - \theta_{k+1}\|^2 \\ &\leqq \|\hat{\theta}_{k+1} - \bar{\theta}_k\|^2 + 4p^{1/2}L\{\|w_{k+1}\| + \|\theta_k I(\theta_k \notin D)\|\} \\ &\quad + 2\|w_{k+1}\|^2 + 2\|\theta_k I(\theta_k \notin D)\|^2 \\ &\leqq \|\hat{\theta}_{k+1} - \bar{\theta}_k\|^2 + 2\{2p^{1/2}L + \|w_{k+1}\|\}\|w_{k+1}\| \\ &\quad + 2\{2p^{1/2}L + \|\theta_k\|\}\|\theta_k\| I(\theta_k \notin D) \end{aligned} \tag{5.2}$$

But by (3.8) and the properties of the projection we know that

$$\begin{aligned} \|\bar{\theta}_k - \hat{\theta}_{k+1}\|^2 &\leqq \left\| \bar{\theta}_k - \hat{\theta}_k - \frac{\varphi_k}{d + \|\varphi_k\|^2}[\varphi_k^\tau(\theta_k - \hat{\theta}_k) + v_{k+1}] \right\|^2 \\ &= \left\| \left(I - \frac{\varphi_k \varphi_k^\tau}{d + \|\varphi_k\|^2}\right)\tilde{\theta}_k - \left\{\theta_k I(\theta_k \notin D) - \frac{\varphi_k v_{k+1}}{d + \|\varphi_k\|^2}\right\} \right\|^2 \\ &\leqq \|\tilde{\theta}_k\|^2 - \frac{\|\varphi_k^\tau \tilde{\theta}_k\|^2}{d + \|\varphi_k\|^2} + 2\|\theta_k\|^2 I(\theta_k \notin D) + \frac{2}{d}\|v_{k+1}\|^2 \\ &\quad + 2\|\tilde{\theta}_k\|\,\|\theta_k\| I(\theta_k \notin D) + 2\frac{\|\varphi_k^\tau \tilde{\theta}_k\|\,|v_{k+1}|}{d + \|\varphi_k\|^2}. \end{aligned}$$

Applying the following elementary inequality

$$2xy \leqq \tfrac{1}{2}x^2 + 2y^2 \quad \forall x, y$$

with

$$x = \frac{\|\varphi_k^\tau \tilde{\theta}_k\|}{(d + \|\varphi_k\|^2)^{1/2}}, \qquad y = \frac{|v_{k+1}|}{(d + \|\varphi_k\|^2)^{1/2}}$$

to the last term, we then see that

$$\begin{aligned} \|\bar{\theta}_k - \hat{\theta}_{k+1}\|^2 &\leqq \|\tilde{\theta}_k\|^2 - \frac{1}{2}\frac{\|\varphi_k^\tau \tilde{\theta}_k\|^2}{d + \|\varphi_k\|^2} + \frac{4}{d}\|v_{k+1}\|^2 + 2\{(p^{1/2}L + \|\theta_k\|)\|\theta_k\| + \|\theta_k\|^2\} I(\theta_k \notin D) \\ &\leqq \|\tilde{\theta}_k\|^2 - \frac{1}{2}\frac{\|\varphi_k^\tau \tilde{\theta}_k\|^2}{d + \|\varphi_k\|^2} + \frac{4}{d}\|v_{k+1}\|^2 + 2\{p^{1/2}L + 2\|\theta_k\|\}\|\theta_k\| I(\theta \notin D). \end{aligned}$$

Substituting this into (5.2) we have

$$\|\tilde{\theta}_{k+1}\|^2 \leqq \|\tilde{\theta}_k\|^2 - \frac{1}{2}\frac{\|\varphi_k^\tau \tilde{\theta}_k\|^2}{d+\|\varphi_k\|^2} + \frac{4}{d}\|v_{k+1}\|^2 + 2\{2p^{1/2}L + \|w_{k+1}\|\}\|w_{k+1}\|$$
$$+6\{p^{1/2}L + \|\theta_k\|\}\|\theta_k\| I(\theta_k \notin D),$$

which is tantamount to the desired result. □

In a similar way, the following lemma can also be proved.

LEMMA 5.1′. *Under the conditions of Theorem* 2,

$$\alpha_k \leqq 2(\|\tilde{\theta}_k\|^2 - \|\tilde{\theta}_{k+1}\|^2) + (6/d)\|v_{k+1}\|^2 + 2\delta_1(4(s+t)^{1/2}L + \delta_1).$$

LEMMA 5.2. *Let the closed-loop system be expressed by*

$$y_{k+1} = \varphi_k^\tau \tilde{\theta}_k + v_{k+1}$$

Assume that there are constants $\lambda \in (0,1)$, $K_1 > 0$, $K_2 \geqq 0$ *such that*

$$\sum_{i=0}^{n} \lambda^{n-i}\|\varphi_i\|^2 \leqq \sum_{i=0}^{n} \lambda^{n-i}\{K_1(y_i)^2 + K_2(v_i)^2\} \quad \forall n \geqq 0 \tag{5.3}$$

Then for any $\beta \geqq 2$,

$$\{E\|\varphi_n\|^\beta\}^{1/\beta} \leqq K_0\left\{1 + (1-\lambda)^{-1/2}\left\{E\left[\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)^2\right]^{\beta/2}\right\}^{1/\beta}\right\}^{1/2} \quad \forall n \geqq 1,$$

$$\frac{1}{N}\sum_{n=0}^{N}\|\varphi_n\|^2 \leqq O\left(\left\{\frac{1}{N}\sum_{n=0}^{N}\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)^2\right\}^{1/2}\right) + O(1),$$

where α_j *is defined in* (5.1), *and*

$$K_0 = (1-\lambda)^{-1/2}\{[2K_1+K_2][\sigma_{2\beta}(v)]^2 + 2dK_1[2pL^2 + 2(\sigma_{2\beta}(\theta))^2]\}^{1/2},$$
$$\sigma_{2\beta}(v) = \sup_k\{E|v_k|^{2\beta}\}^{1/(2\beta)}, \qquad \sigma_{2\beta}(\theta) = \sup_k\{E|\theta_k|^{2\beta}\}^{1/(2\beta)}.$$

Proof. By the assumption it follows that

$$\|\varphi_n\|^2 \leqq \sum_{i=0}^{n}\lambda^{n-i}\|\varphi_i\|^2$$
$$\leqq \sum_{i=0}^{n}\lambda^{n-i}\{K_1[2\|\varphi_{i-1}^\tau\tilde{\theta}_{i-1}\|^2 + 2(v_i)^2] + K_2(v_i)^2\}$$
$$= 2K_1\sum_{i=0}^{n-1}\lambda^{n-i-1}\alpha_i(\|\varphi_i\|^2 + d) + (2K_1+K_2)\sum_{i=0}^{n}\lambda^{n-i}(v_i)^2$$
$$\leqq 2K_1\sum_{i=0}^{n-1}\lambda^{n-i-1}\alpha_i\|\varphi_i\|^2 + (2K_1+K_2)\sum_{i=0}^{n}\lambda^{n-i}(v_i)^2 + 2dK_1\sum_{i=0}^{n-1}\lambda^{n-i-1}\|\tilde{\theta}_i\|^2$$

So by Lemma 4.1(i) with $x_i = \lambda^{-i}\|\varphi_i\|^2$, it is seen that

$$\|\varphi_n\|^2 \leqq \xi_n + \sum_{i=0}^{n-1}\lambda^{n-i}\left[\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)\right]\xi_i \tag{5.4}$$

where

$$\xi_i = (2K_1+K_2)\sum_{k=0}^{i}\lambda^{i-k}(v_k)^2 + 2dK_1\sum_{k=0}^{i-1}\lambda^{i-k-1}\|\tilde{\theta}_k\|^2.$$

Applying the Minkowski inequality and the Schwarz inequality to (5.4), we get

$$\begin{aligned}\{E\|\varphi_n\|^\beta\}^{2/\beta} &\leqq \{E|\xi_n|^{\beta/2}\}^{2/\beta}+\left\{E\left[\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)\xi_i\right]^{\beta/2}\right\}^{2/\beta}\\
&\leqq \{E|\xi_n|^{\beta/2}\}^{2/\beta}+\left\{E\left[\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)^2\right]^{\beta/4}\right.\\
&\qquad\left.\cdot\sum_{i=0}^{n-1}\lambda^{n-i}\|\xi_i\|^2\right]^{\beta/4}\Bigg\}^{2/\beta}\\
&\leqq \{E|\xi_n|^{\beta/2}\}^{2/\beta}+\left\{E\left[\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)^2\right]^{\beta/2}\right.\\
&\qquad\left.\cdot E\left[\sum_{i=0}^{n-1}\lambda^{n-i}\|\xi_i\|^2\right]^{\beta/2}\right\}^{1/\beta}\\
&\leqq \{E|\xi_n|^{\beta/2}\}^{2/\beta}+\left\{E\left[\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)^2\right]^{\beta/2}\right\}^{1/\beta}\\
&\qquad\cdot\left\{\sum_{i=0}^{n-1}\lambda^{n-i}[E\|\xi_i\|^\beta]^{2/\beta}\right\}^{1/2}\\
&\leqq\left\{1+(1-\lambda)^{-1/2}\left\{E\left[\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)^2\right]^{\beta/2}\right\}^{1/\beta}\right\}\\
&\qquad\cdot\sup_{0\leqq i\leqq n}\{E[\xi_i]^\beta\}^{1/\beta}.\end{aligned}$$

Again by the Minkowski inequality,

$$\begin{aligned}\{E[\xi_i]^\beta\}^{1/\beta}&\leqq(2K_1+K_2)\sum_{k=0}^{i}\lambda^{i-k}\{E(v_k)^{2\beta}\}^{1/\beta}+2dK_1\sum_{k=0}^{i-1}\lambda^{i-k-1}\{E\|\tilde{\theta}_k\|^{2\beta}\}^{1/\beta}\\
&\leqq(1-\lambda)^{-1}\{[2K_1+K_2][\sigma_{2\beta}(v)]^2+2dK_1[2pL^2+2(\sigma_{2\beta}(\theta))^2]\}.\end{aligned}$$

Hence the first assertion of the lemma is true, while the second assertion can easily be proved by following the similar argument and by using (5.4), Lemma 4.1(ii), and the Schwarz inequality. □

LEMMA 5.3. *Under conditions of Theorem* 1, *the property* (5.3) *holds with* $K_1=p\lambda^{-(p-1)}$, $K_2=0$. *Furthermore,*

$$\{E\|\varphi_n\|^\beta\}^{1/\beta}\leqq K_0\left\{1+(1-\lambda)^{-1/2}\prod_{k=1}^{4}\{E[I_k(n)]^{\beta/2}\}^{1/(4\beta)}\right\}^{1/2}\quad\forall n\geqq1,\tag{5.5}$$

$$\frac{1}{N}\sum_{n=0}^{N}\|\varphi_n\|^2\leqq O\left(\prod_{k=1}^{4}\left\{\frac{1}{N}\sum_{n=0}^{N}I_k(n)\right\}^{1/8}\right)+O(1)\quad as\ N\to\infty,\tag{5.6}$$

where $I_k(n)$, $k=1,\cdots,4$, *are defined as*

$$I_1(n)=\sum_{i=0}^{n-1}\lambda^{n-i}\exp\{8\beta_1\|\tilde{\theta}_i\|^2\},\qquad \beta_1=4K_1\lambda^{-1},\tag{5.7}$$

$$I_2(n)=\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}\exp\left\{\frac{32\beta_1}{d}\|v_{j+1}\|^2\right\},\tag{5.8}$$

$$I_3(n)=\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}\exp\{16\beta_1(2p^{1/2}L+\|w_{j+1}\|)\|w_{j+1}\|\},\tag{5.9}$$

$$I_4(n)=\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}\exp\{48\beta_1(p^{1/2}L+\|\theta_j\|)\|\theta_j\|I(\theta_j\notin D)\}.\tag{5.10}$$

Proof. By the definition of φ_k in (2.4) and the fact that $y_k=0$ for $k<0$, it is easy to see that (5.3) holds with $K_1=p\lambda^{-(p-1)}$, $K_2=0$.

By Lemma 5.1 and the inequality $\log(1+x)\leqq x$, for all $x\geqq 0$, we have

$$\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)^2=\exp\left\{\sum_{j=i}^{n-1}2\log(1+2K_1\lambda^{-1}\alpha_j)\right\}$$

$$\leqq \exp\left\{\beta_1\sum_{j=i}^{n-1}\alpha_j\right\},\qquad (\beta_1=4K_1\lambda^{-1}),$$

$$\leqq \exp\{2\beta_1\|\tilde{\theta}_i\|^2\}\exp\left\{\beta_1\sum_{j=i}^{n-1}\left[\frac{8}{d}\|v_{j+1}\|^2+4(2p^{1/2}L+\|w_{j+1}\|)\|w_{j+1}\|\right.\right.$$

$$\left.\left.+12(p^{1/2}L+\|\theta_j\|)\|\theta_j\|I(\theta_j\notin D)\right]\right\}$$

$$\leqq \exp\{2\beta_1\|\tilde{\theta}_i\|^2\}\prod_{j=i}^{n-1}\exp\left\{\frac{8\beta_1}{d}\|v_{j+1}\|^2\right\}$$

$$\cdot\prod_{j=i}^{n-1}\exp\{4\beta_1(2p^{1/2}L+\|w_{j+1}\|)\|w_{j+1}\|\}$$

$$\cdot\prod_{j=i}^{n-1}\exp\{12\beta_1(p^{1/2}L+\|\theta_j\|)\|\theta_j\|I(\theta_j\notin D)\}. \tag{5.11}$$

Consequently, by the Holder inequality and (5.7)–(5.10),

$$E\left[\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}(1+2K_1\lambda^{-1}\alpha_j)^2\right]^{\beta/2}\leqq E\left\{\prod_{k=1}^{4}I_k(n)\right\}^{\beta/8}$$

$$\leqq\left\{\prod_{k=1}^{4}E[I_k(n)]^{\beta/2}\right\}^{1/4}.$$

Substituting this into Lemma 5.2, we see that (5.5) is true, while (5.6) can be proved in a similar way by using (5.11) and the Hölder inequality. The details will not be repeated. □

LEMMA 5.3′. *Under conditions of Theorem* 2, *property* (5.3) *holds with*

$$K_1=s\lambda^{-s+1}[1+s(M_1/b_1)^2]+(t-1)\lambda^{-t+1}[1+(t-1)(M_1/b_1)^2]A\lambda^2/(\lambda-p),$$

$$K_2=(t-1)\lambda^{-t+1}[1+(t-1)(M_1/b_1)^2]A\lambda^2/(\lambda-p).$$

Furthermore,

$$\{E\|\varphi_n\|^\beta\}^{1/\beta}\leqq K_0\left\{1+(1-\lambda)^{-1/2}\prod_{k=1}^{3}\{E[J_k(n)]^{\beta/2}\}^{1/(3\beta)}\right\}^{1/2}\quad\forall n\geqq 1, \tag{5.12}$$

$$\frac{1}{N}\sum_{n=0}^{N}\|\varphi_n\|^2\leqq O\left(\prod_{k=1}^{3}\left\{\frac{1}{N}\sum_{n=0}^{N}J_k(n)\right\}^{1/6}\right)+O(1)\quad as\ N\to\infty, \tag{5.13}$$

where $J_k(n)$, $k=1,2,3$, *are defined as*

$$J_1(n)=\sum_{i=0}^{n-1}\lambda^{n-i}\exp\{6\delta_1\beta_1(n-i)[4(s+t)^{1/2}L+\delta_1]\},\qquad \beta_1=4K_1\lambda^{-1}, \tag{5.14}$$

$$J_2(n)=\sum_{i=0}^{n-1}\lambda^{n-i}\exp\{6\beta_1\|\tilde{\theta}_i\|^2\}, \tag{5.15}$$

$$J_3(n)=\sum_{i=0}^{n-1}\lambda^{n-i}\prod_{j=i}^{n-1}\exp\left\{\frac{18\beta_1}{d}\|v_{j+1}\|^2\right\}. \tag{5.16}$$

Proof. Let us write $\hat{a}_i(k)$, $\hat{b}_j(k)$, as the estimates for $a_i(k)$, $b_i(k)$ given by $\hat{\theta}_k$. Then from (3.22),

$$u_k = \frac{-1}{\hat{b}_1(k)}\{\hat{a}_1(k)y_k + \cdots + \hat{a}_s(k)y_{k-s+1} + \hat{b}_2(k)u_{k-1} + \cdots + \hat{b}_t(k)u_{k-t+1}\}.$$

Since $\hat{\theta}_k$ belongs to the domain D defined by (3.20), it follows by the Schwarz inequality that

$$|u_k|^2 \leqq \left(\frac{M_1}{b_1}\right)^2 \left\{ s \sum_{j=0}^{s-1} |y_{k-j}|^2 + (t-1) \sum_{j=1}^{t-1} |u_{k-j}|^2 \right\}.$$

Therefore, by the definition of φ_i in (2.5),

$$\begin{aligned}
\sum_{i=0}^{n} \lambda^{n-i}\|\varphi_i\|^2 &= \sum_{i=0}^{n} \lambda^{n-i} \sum_{j=0}^{s-1} |y_{i-j}|^2 + \sum_{i=0}^{n} \lambda^{n-i} \sum_{j=1}^{t-1} |u_{i-j}|^2 + \sum_{i=0}^{n} \lambda^{n-i}|u_i|^2 \\
&\leqq \left[1 + s\left(\frac{M_1}{b_1}\right)^2\right] \sum_{i=0}^{n} \lambda^{n-i} \sum_{j=0}^{s-1} |y_{i-j}|^2 \\
&\quad + \left[1 + (t-1)\left(\frac{M_1}{b_1}\right)^2\right] \sum_{i=0}^{n} \lambda^{n-i} \sum_{j=1}^{t-1} |u_{i-j}|^2 \\
&\leqq s\lambda^{-s+1}\left[1 + s\left(\frac{M_1}{b_1}\right)^2\right] \sum_{i=0}^{n} \lambda^{n-i}|y_i|^2 \\
&\quad + (t-1)\lambda^{-t+1}\left[1 + (t-1)\left(\frac{M_1}{b_1}\right)^2\right] \sum_{i=0}^{n-1} \lambda^{n-i}|u_i|^2. \qquad (5.17)
\end{aligned}$$

Note that by the assumption (3.16),

$$\begin{aligned}
\sum_{i=0}^{n-1} \lambda^{n-i}|u_i|^2 &\leqq A \sum_{i=0}^{n-1} \lambda^{n-i} \sum_{j=0}^{i+1} \rho^{i+1-j}\{|y_j|^2 + |v_j|^2\} \\
&\leqq A \sum_{j=0}^{n} \sum_{i=j-1}^{n-1} \lambda^{n-i}\rho^{i+1-j}\{|y_j|^2 + |v_j|^2\} \\
&= A\lambda \sum_{j=0}^{n} \lambda^{n-j} \sum_{i=j-1}^{n-1} \left(\frac{\rho}{\lambda}\right)^{i+1-j} \{|y_j|^2 + |v_j|^2\} \\
&\leqq \frac{A\lambda^2}{\lambda - \rho} \sum_{j=0}^{n} \lambda^{n-j}\{|y_j|^2 + |v_j|^2\}.
\end{aligned}$$

Combining this with (5.17) we see that (5.3) is true. The second assertion can easily be proved by using techniques similar to those used in Lemma 5.3. We need only to note that under the present conditions the inequality (5.11) is changed to (via Lemma 5.1′),

$$\begin{aligned}
\prod_{j=i}^{n-1} (1 + 2K_1\lambda^{-1}\alpha_j)^2 &\leqq \exp\{2\delta_1\beta_1(n-i)[4(s+t)^{1/2}L + \delta_1]\} \\
&\quad \cdot \exp\{2\beta_1\|\tilde{\theta}_i\|^2\} \prod_{j=i}^{n-1} \exp\left\{\frac{6\beta_1}{d}\|v_{j+1}\|^2\right\}.
\end{aligned}$$

The details will not be repeated here. □

We now proceed to analyze the quantities $I_k(n)$, $k = 1, \cdots, 4$, appearing in (5.7)–(5.10) by using Lemma 4.3. For this we need the following lemma.

LEMMA 5.4. *Under conditions of Theorem* 1, *the following inequalities hold* (*where* $\beta_1 = 4p\lambda^{-p}$):

(5.18) (i) $\sup_{(j,\omega)} E\{\exp[(16\beta\beta_1(r+1)/d)\|v_{j+1}\|^2] \mid F_{j-r}\} < \lambda^{-\beta(r+1)/2}$,

(5.19) (ii) $\sup_{(j,\omega)} E\{\exp\{8\beta\beta_1(m+1)(2p^{1/2}L + \|w_{j+1}\|)\|w_{j+1}\|\} \mid F_{j-m}\} < \lambda^{-\beta(m+1)/2}$,

(5.20) (iii) $\sup_{(j,\omega)} E\{\exp[24\beta\beta_1(m+1)(p^{1/2}L + \|\theta_{j+1}\|)\|\theta_{j+1}\| I(\theta_{j+1} \notin D)] \mid F_{j-m}\}$
$$< \lambda^{-\beta(m+1)/2},$$

where j takes nonnegative integer values and ω *is the sampling point*

Proof. (i) By the noise assumption (2.6), the choice of d in (3.14), and the Hölder inequality, we have (note that $\beta_1 = 4p\lambda^{-p}$)

$$E\{\exp[(16\beta\beta_1(r+1)/d)\|v_{j+1}\|^2] \mid F_{j-r}\} \leqq \exp\{2^6 p\lambda^{-p}\beta(r+1)M_v/(\varepsilon_0 d)\} < \lambda^{-\beta(r+1)/2}$$

(ii) By Lemma 4.4 and the parameter assumption (3.2),

$$E\{\exp\{2^5\beta\beta_1(m+1)p^{1/2}L\|w_{j+1}\|\} \mid F_{j-m}\}$$
(5.21)
$$= E\{\exp\{[2^7(m+1)\beta p^{3/2}\lambda^{-p}LM^{-1/2}][(M)^{1/2}\|w_{k+1}\|]\} \mid F_{k-m}\}$$
$$\leqq \exp\{f(2^7(m+1)\beta p^{3/2}L\lambda^{-p}M^{-1/2})\delta_\theta^{1/2}\},$$

where the function $f(\cdot)$ is defined by (3.12)

Again, by the parameter assumption (3.2) and the Hölder inequality,

$$E\{\exp\{2^4\beta\beta_1(m+1)\|w_{j+1}\|^2\} \mid F_{j-m}\} \leqq \exp\{2^6\beta p\lambda^{-p}(m+1)\delta_\theta/M\};$$

combining this with (5.21) we have via the Schwarz inequality,

$$E\{\exp\{8\beta\beta_1(m+1)(2p^{1/2}L + \|w_{j+1}\|)\|w_{j+1}\|\} \mid F_{j-m}\}$$
$$\leqq \exp\{[f(2^7(m+1)\beta p^{3/2}L\lambda^{-p}M^{-1/2})/2 + 2^5\beta p\lambda^{-p}(m+1)/M]\delta_\theta^{1/2}\}$$
$$< \lambda^{-\beta(m+1)/2},$$

where the last inequality is derived from (3.11).

(iii) We now proceed to prove (5.20) Let us denote $b = 192\beta(m+1)p^{3/2}\lambda^{-p}$; then by the parameter assumption (3.1) and the Markov inequality,

$$E\{\exp[48\beta\beta_1(m+1)p^{1/2}L\|\theta_{j+1}\|I(\theta_{j+1} \notin D)] \mid F_{j-m}\}$$
$$= E\{\exp[bL\|\theta_{j+1}\|I(\theta_{j+1} \notin D)] \mid F_{j-m}\}$$
$$= E\{I(\theta_{j+1} \in D) \mid F_{j-m}\} + E\{\exp[bL\|\theta_{j+1}\|]I(\theta_{j+1} \notin D) \mid F_{j-m}\}$$
$$\leqq 1 + \{E[\exp(2bL\|\theta_{j+1}\|) \mid F_{j-m}]\}^{1/2}\{P(\|\theta_{j+1}\| > L \mid F_{j-m})\}^{1/2}$$
$$= 1 + \{E[\exp(2bL\|\theta_{j+1}\|) \mid F_{j-m}]\}^{1/2}\{P[\exp(2bL\|\theta_{j+1}\|) > \exp(2bL^2) \mid F_{j-m}]\}^{1/2}$$
$$\leqq 1 + E[\exp(2bL\|\theta_{j+1}\|) \mid F_{j-m}]/\exp(bL^2)$$
$$\leqq 1 + \exp(-bL^2/2)E[\exp(2b\|\theta_{j+1}\|^2) \mid F_{j-m}]$$
$$\leqq 1 + \exp(-bL^2/2)\exp(2bM_\theta/M)$$
$$\leqq \exp\{\exp[192(m+1)\beta p^{3/2}\lambda^{-p}(2M_\theta/M - L^2/2)]\}$$
(5.22)
$$\leqq \exp\{\exp[192(m+1)\beta p\lambda^{-p}(2M_\theta/M - L^2/2)]\},$$

where for the last inequality we have used the fact that $2M_\theta/M - L^2/2 \leqq 0$, which is seen from (3.13) and the choice $L = L_0$.

Similarly, we have ($c = 192\beta(m+1)p\lambda^{-p}$),

$$E\{\exp[48\beta\beta_1(m+1)\|\theta_{j+1}\|^2 I(\theta_{j+1} \notin D)] | F_{j-m}\}$$
$$= E\{\exp[c\|\theta_{j+1}\|^2 I(\theta_{j+1} \notin D)] | F_{j-m}\}$$
$$\leqq 1 + \{E\{\exp[2c\|\theta_{j+1}\|^2] | F_{j-m}\}\}^{1/2}\{P(\|\theta_{j+1}\|^2 > L^2 | F_{j-m})\}^{1/2}$$
$$\leqq 1 + E\{\exp[2c\|\theta_{j+1}\|^2] | F_{j-m}\}/\exp\{cL^2\}$$
$$\leqq 1 + \exp\{2cM_\theta/M - cL^2\}$$
$$\leqq \exp\{\exp[192\beta(m+1)p\lambda^{-p}(2M_\theta/M - L^2)]\}. \tag{5.23}$$

Combining (5.22) and (5.23), we obtain via the Schwarz inequality,

$$E\{\exp[24(m+1)\beta\beta_1(p^{1/2}L + \|\theta_{j+1}\|)\|\theta_{j+1}\| I(\theta_{j+1} \notin D)] | F_{j-m}\}$$
$$\leqq \exp\{\exp[192(m+1)\beta p\lambda^{-p}(2M_\theta/M - L^2/2)]\} < \lambda^{-\beta(m+1)/2},$$

where the last inequality is obtained from (3.13). This completes the proof □

Proofs of theorems. By Lemma 4.3 (with $\alpha = \beta/2 > 1$) and Lemma 5.4, we know that the quantities $I_k(n)$, $k = 2, \cdots, 4$, defined in Lemma 5.3 satisfy

$$\sup_n E[I_k(n)]^{\beta/2} < \infty \quad \text{and} \quad \sum_{n=0}^{N} I_k(n) = O(N) \quad \text{a.s.} \quad k = 2, 3, 4, \tag{5.24}$$

while for $I_1(n)$, we note that

$$\exp\{8\beta_1\|\tilde{\theta}_i\|^2\} \leqq \exp\{16\beta_1 pL^2\}\exp\{16\beta_1\|\theta_i\|^2\}$$

Then by the parameter assumption (3.1) and Lemma 4.1(ii), it is easy to see that (5.24) is also true for $k = 1$. Hence by Lemma 5.3 we get

$$\sup_n E\|\varphi_n\|^\beta < \infty \quad \sum_{n=1}^{N} \|\varphi_n\|^2 = O(N) \quad \text{a.s.} \quad \text{as } N \to \infty;$$

combining this with (3.9) we immediately conclude that Theorem 1 holds.

In a similar way, Theorem 2 can be proved. The details will not be repeated here. □

6. Further discussions. In this section we will give some brief discussions on the issues of performance and robustness

6.1. Performance. Since our control objective is to minimize the output process, it is natural to ask if the output "approaches zero" when both the noise and the parameter variation processes are "small." Mathematically, this needs the study of, e.g., for Model 1, the asymptotic properties of $\{y_k\}$ when $(\varepsilon)^{-1} \to 0$ (M_v fixed), and $\delta_\theta \to 0$. Note that by (3.14), d is allowed to be chosen as $d \to 0$ and $(\varepsilon d)^{-1} \to 0$.

Let us denote $\bar{\varepsilon} = (\delta_\theta, d, (\varepsilon d)^{-1})$ and parameterize the output process as $\{y_k^{\bar{\varepsilon}}\}$; then from the proof of Theorem 1, it is easy to see that

$$\lim_{\bar{\varepsilon} \to 0} \limsup_{k \to \infty} E\|y_k^{\bar{\varepsilon}}\|^\beta = 0 \tag{6.1}$$

For any small but fixed $\bar{\varepsilon}$, the Markov chain ergodic theory may be applied to prove the existence of the limit $\lim_{k\to\infty} E\|y_k\|^2$ if we strengthen the assumptions. For example, if $\{v_k\}$ is a Gaussian white noise sequence, and the parameter is modeled as in Example 2, then under the assumptions of Theorem 1, the closed-loop system equations will give rise to a Markov state process $\{\Phi_k\}$, which is, in particular, (i) weakly stochastically controllable in the sense of Meyn and Caines (1988), and (ii) bounded in probability due to Theorem 1. Thus, applying Theorem 1 (for $\beta > 2$) and the important results developed in Meyn and Caines (1988) and Meyn (1988), we know that

$$\lim_{k\to\infty} P(|y_k| > x) = \pi(|y| > x) \quad \forall x,$$
$$(6.2)$$
$$\lim_{k\to\infty} E|y_k|^2 = \lim_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k} |y_i|^2 = \int y^2\, d\pi < \infty,$$

where y denotes the function $y(\cdot)$ such that $y_k = y(\Phi_k)$, and π is the invariant probability of $\{\Phi_k\}$. Detailed and further results are currently under investigation. We mention that establishing the existence of the limits in (6.2) *without* using Markov chain theory appears to be a challenging problem.

6.2. Robustness. Let us assume that in addition to the random noise $\{v_k\}$ there are unmodeled dynamics $\{\eta_k\}$ acting on the system (2.3):

$$z_{k+1} = \varphi_k^\tau \theta_k + v_{k+1} + \eta_k. \tag{6.3}$$

We assume that the unmodeled dynamics $\{\eta_k\}$ depend on the previous input-output data, and have the following time-varying upper bound (see, e.g., Ioannou and Tsakalis (1985), Chen and Guo (1988)):

$$|\eta_k| \leqq \varepsilon^* m_k, \quad m_k = \gamma m_{k-1} + \|\varphi_k\|, \quad m_0 > 0, \quad k \geqq 0, \tag{6.4}$$

where $\varepsilon^* > 0$, $\gamma \in (0, 1)$.

Similar to the normalization idea used in Ioannou and Tsakalis (1985), we replace the quantity $d + \|\varphi_k\|^2$ in (3.8) or (3.21) by $d + (m_k)^2$, and consider the following algorithm:

$$\hat\theta_{k+1} = \pi_D\left\{\hat\theta_k + \frac{\varphi_k}{d + (m_k)^2}(z_{k+1} - \varphi_k^\tau \hat\theta_k)\right\} \tag{6.5}$$

Then stability of the closed-loop system under the certainty equivalent minimum variance adaptive control law can also be established, provided that ε^* is appropriately small. The proof is essentially the same as that for Theorems 1 and 2.

7. Conclusion. In this paper, stabilizing adaptive controllers are presented for possible open-loop unstable time-varying stochastic systems described by Models 1 and 2. The closed-loop stability is proved based on an analysis of products of random variables and truncation techniques. We have seen that the use of projection in the estimation algorithm plays a crucial role in getting useful estimates in the stochastic case, especially when the noise is unbounded in sample path. Further asymptotic results are currently being explored by applying the weak convergence theory and the Markov chain ergodic theory.

Acknowledgments. The author thanks Dr. S. P. Meyn for his comments and valuable discussions on Markov chain theory. Thanks are also due to Professor P. A. Ioannou of the University of Southern California for his valuable discussions during the author's visit to the Australian National University.

REFERENCES

A. BECKER, P. R. KUMAR, AND C. Z. WEI (1985), *Adaptive control with the stochastic approximation algorithm, geometry and convergence*, IEEE Trans. Automat. Control, 30, pp. 330–338.

H. F. CHEN AND P. E. CAINES (1985), *On the adaptive control of a class of systems with random parameters and disturbances*, Automatica, 21, pp. 737–741.

H. F. CHEN AND L. GUO (1987), *Asymptotically optimal adaptive control with consistent parameter estimates*, SIAM J. Control Optim., 25, pp. 558–575.

——— (1988), *A robust stochastic adaptive controller*, IEEE Trans. Automat. Control, 33, pp. 1035–1043.

Y. S. CHOW AND H. TEICHER (1978), *Probability Theory: Independency, Interchangeability and Martingales*, Springer-Verlag, New York.

C. A. DESOER AND M. VIDYASAGAR (1975), *Feedback Systems: Input-Output Properties*, Academic Press, New York.

G. C. GOODWIN AND K. S. SIN (1984), *Adaptive Filtering, Predication and Control*, Prentice-Hall, Englewood Cliffs, NJ.

C. W. J. GRANGER AND A. P. ANDERSEN (1978), *An Introduction to Bilinear Time Series Models*, Vandenhoeck & Ruprecht, Gottingen.

L. GUO AND S. P. MEYN (1989), *Adaptive control for time-varying systems: a combination of martingale and Markov chain techniques*, Internat. J. Adaptive Control Signal Process., 3, pp. 1–14.

L. GUO, J. B. MOORE, AND L. G. XIA (1988), *Tracking randomly varying parameters—analysis of a standard algorithm*, Proc. 27th IEEE Conf. on Decision and Control, Austin, Texas, pp. 1514–1519.

P. IOANNOU AND K. TSAKALIS (1985), *Robust discrete-time adaptive control*, in Adaptive and Learning Systems; Theory and Applications, K. S. Narendra, ed., Plenum Press, New York, pp. 73–85.

L. LJUNG AND T. SODERSTROM (1983), *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA.

R. H. MIDDLETON AND G. C. GOODWIN (1988), *Adaptive control of time-varying linear systems*, IEEE Trans. Automat. Control, 33, pp. 150–157.

S. P. MEYN AND P. E. CAINES (1987), *A new approach to stochastic adaptive control*, IEEE Trans. Automat. Control, 32, pp. 220–226.

S. P. MEYN (1989), *Ergodic theorems for discrete time stochastic systems using a stochastic Lyapunov function*, SIAM J. Control Optim., 27, pp. 1409–1439.

S. P. MEYN AND P. E. CAINES (1991), *Asymptotic behavior of stochastic systems possessing Markovian realizations*, SIAM J. Control Optim., to appear.

M. POURAHMADI (1986), *On stationary of the solution of a double stochastic model*, J. Time Series Anal., 7, pp. 123–131.

W. F. STOUT (1974), *Almost Sure Convergence*, Academic Press, New York.

D. TJOSTHEIM (1986), *Some double stochastic time series models*, J. Time Series Anal., 7, pp. 51–72.

K. TSAKALIS AND P. A. IOANNOU (1986), *Adaptive control of linear time-varying plants*, in Proc. IFAC Workshop on Adaptive Systems Control and Signal Processing, Lund, Sweden.