Self-Convergence of Weighted Least-Squares with Applications to Stochastic Adaptive Control

Lei Guo, *Member, IEEE*

***Abstract*— A recursive least-squares algorithm with slowly decreasing weights for linear stochastic systems is found to have self-convergence property, i.e., it converges to a certain random vector almost surely irrespective of the control law design. Such algorithms enjoy almost the same nice asymptotic properties as the standard least-squares. This "universal convergence" result combined with a method of random regularization then easily can be applied to construct a self-convergent and uniformly controllable estimated model and thus may enable us to form a general framework for adaptive control of possibly nonminimum phase autoregressive-moving average with exogenous input (ARMAX) systems. As an application, we give a simple solution to the well-known stochastic adaptive pole-placement and linear-quadratic-Gaussian (LQG) control problems in the paper.**

I. Introduction

CONSIDER the following linear regression model

$$y_{t+1} = \theta^\tau \phi_t + w_{t+1} \tag{1}$$

where $\theta \in \mathbb{R}^n$ is an unknown parameter vector, y_t, ϕ_t, and w_t are the observation, regressor, and noise processes, respectively.

A common and natural way to estimate the parameter θ is the least-squares (LS) method. That is, the estimate is the minimizer of the following criterion

$$J_t(\theta) = \frac{1}{2}\sum_{i=0}^{t} \alpha_i (y_{i+1} - \theta^\tau \phi_i)^2 \tag{2}$$

where $\alpha_i \geq 0$ is a weighting sequence; it allows us to give different weights to different measurements of interest. Various weighted least-squares (WLS) in the literature differ only in the choice of weights. The standard least-squares correspond to $\alpha_i \equiv 1$. It is well known that the optimal choice of weights (in the sense of minimum variance) is that $\{\alpha_i\}$ is taken as the inverse of the noise variance (cf. [1] p. 36). Also, when θ is a time-varying parameter, less (or decaying) weights should be given to the old measurements so that the estimate has good tracking performance (cf. Guo *et al.* [2]).

In this paper, we shall study the WLS from an adaptive control point of view. In contrast to the insights from pure estimation, we shall use slowly decreasing weights in (2) which will give fewer weights to the current measurements. The motivation of doing so may be explained as follows: In analyzing stochastic adaptive control systems, nothing can be assumed *a priori* about the closed-loop signals $\{y_t, \phi_t\}$. Thus, if $\{y_t, \phi_t\}$ has a tendency of growing up unboundedly, decreasing weights in (2) will depress the undesirable effect of instability on the estimation, and hence (2) may still give useful parameter estimates in the case of instability and lack of excitation. Of course, so that the WLS enjoys the similar nice asymptotic properties as the standard LS, the decreasing rate of $\{\alpha_i\}$ should be chosen as slow as possible.

The first paper using WLS in stochastic adaptive control seems to have been Kumar and Moore [3] where the weights are chosen according to some stability/excitation measure of $\{\phi_t\}$. More recently, Bercu and Duflo [4] and Bercu [5] obtained various interesting and useful results on WLS which are parallel and based on those of the standard least-squares. The weights in [4] and [5] are simpler than those in [3] and are called "ponderations." All of these papers, however, concern exclusively with adaptive control of minimum phase stochastic systems.

This paper is motivated by, and aims at, the study of general stochastic adaptive control systems. The analysis of such systems has long been recognized as difficult, due to the inherent nonlinearity of the closed-loop equations and the complexity of the stochastic process involved. As a matter of fact, if the standard LS estimate is used in adaptive control design, we are, at present, only able to prove stability and optimality of the certainty equivalence minimum variance adaptive control (cf. Guo and Chen [6] and Guo [7], [8]) or its generalizations under a certain minimum phase condition (cf. Meyn and Brown [9] and Ren and Kumar [10]). For more complicated control problems such as pole-placement and linear-quadratic-Gaussian (LQG) control, the stability analysis has been hampered by the following facts: i) the standard LS estimates may not converge (or even may not be bounded) and ii) the estimated models may not be uniformly controllable.

In this paper, our first contribution is to establish that the WLS has self-convergence property, i.e., the WLS with slowly decreasing weights converges to a certain random vector almost surely irrespective of the control law design.

This "universal convergence" result may considerably ease the painful task of analyzing stochastic adaptive control systems and may also enable us to form a general framework for adaptive control of possibly nonminimum phase autoregressive-moving average with exogenous input (ARMAX) models. It is worth noting that the standard LS algorithm does not have the above mentioned self-convergence property, in general (see [16]).

Manuscript received August 5, 1994; revised July 3, 1995. Recommended by Associate Editor at Large, B. Pasik-Duncan. This work was supported in part by the National Natural Science Foundation of China.

The author is with Institute of Systems Science, Chinese Academy of Sciences, Beijing 100080, China.

Publisher Item Identifier S 0018-9286(96)00384-4.

Based on this self-convergence result, the estimates modification procedure of Lozano and Zhao [12] and the idea of random search in global optimization (cf. e.g., [13]), a WLS-based parameter estimate, can then easily be constructed so that the corresponding estimated model is almost surely self-convergent and uniformly controllable. This finally enables us to give a simple and complete solution to the longstanding adaptive pole-placement and LQG control problems for ARMAX models without resorting to any projection mechanisms and conditions other than controllability and passivity, and that constitutes the other contributions of the paper.

The remainder of the paper is organized as follows: Section II proves the self-convergence of WLS and gives a comparison with the standard LS. Section III describes how to get uniformly controllable estimated models by random search method. Sections IV and V are devoted to adaptive pole-placement and LQG control problems, respectively. Some concluding remarks are made in Section VI.

II. Self-Convergence of WLS

Consider the following ARMAX model

$$A(z)y_t = B(z)u_t + C(z)w_t, \quad t \geq 0 \tag{3}$$

$$A(z) = 1 + a_1 z + \cdots + a_p z^p, \quad p \geq 0 \tag{4}$$

$$B(z) = b_1 z + \cdots + b_q z^q, \quad q \geq 1 \tag{5}$$

$$C(z) = 1 + c_1 z + \cdots + c_r z^r, \quad r \geq 0 \tag{6}$$

where y_t, u_t, and w_t are the system output, input, and noise sequences, respectively, and $A(z)$, $B(z)$, and $C(z)$ are polynomials in backward-shift operator z with unknown coefficients and known upper bounds p, q, and r for orders.

To describe the WLS algorithm for estimating the unknown parameter vector

$$\theta = [-a_1 \cdots - a_p \, b_1 \cdots b_q \, c_1 \cdots c_r]^\tau \tag{7}$$

we need to introduce a set of functions as follows

$$F \triangleq \{f(\cdot) : f(x) \text{ is slowly increasing and } \int_M^\infty \frac{dx}{xf(x)} < \infty \text{ for some } M > 0\}. \tag{8}$$

Here, a function $f(\cdot)$ is called slowly increasing if it is positive, nondecreasing, and satisfies $f(x^2) = O(f(x))$ for all large $x > 0$.

The recursive WLS algorithm has the following form

$$\theta_{t+1} = \theta_t + L_t(y_{t+1} - \theta_t^\tau \phi_t) \tag{9}$$

$$L_t = \frac{P_t \phi_t}{\alpha_t^{-1} + \phi_t^\tau P_t \phi_t} \tag{10}$$

$$P_{t+1} = P_t - \frac{P_t \phi_t \phi_t^\tau P_t}{\alpha_t^{-1} + \phi_t^\tau P_t \phi_t} \tag{11}$$

$$\phi_t = [y_t \cdots y_{t-p+1} \, u_t \cdots u_{t-p+1} \, \widehat{w}_t \cdots \widehat{w}_{t-r+1}]^\tau \tag{12}$$

$$\widehat{w}_t = y_t - \theta_t^\tau \phi_{t-1}, \quad t \geq 0 \tag{13}$$

where the initial values θ_0 and $P_0 = \alpha I$, $(0 < \alpha < 1)$ are chosen arbitrarily, and where $\{\alpha_t\}$ is the weighting sequence defined by

$$\alpha_t = \frac{1}{f(r_t)}, \qquad r_t = \|P_0^{-1}\| + \sum_{i=0}^{t} \|\phi_i\|^2 \tag{14}$$

with $f(\cdot)$ being any measurable function in the set F defined by (8).

Remark 1: It can easily be shown (see Appendix A) that the necessary condition for a function $f \in F$ is that

$$\log x = o(f(x)), \quad \text{as} \quad x \longrightarrow \infty. \tag{15}$$

Typical functions in F are, for example

$$f(x) = \log^{1+\delta} x, \quad (\log x)(\log\log x)^{1+\delta}, \cdots \quad (\delta > 0).$$

In general, let $f(x)$ be any slowly increasing function as defined above, and it can be shown (see Appendix A) that for any $M > 0$

$$f(x^M) = O(f(x)), \quad x > 0 \tag{16}$$

and that there exists a constant $L > 0$ such that

$$f(x) = O(\log^L x), \qquad \text{for all large } x. \tag{17}$$

The property (17) together with (14) implies that

$$\alpha_t^{-1} = O(\log^L r_t), \quad t \geq 0. \tag{18}$$

The choice of weights in the WLS here is different from that in [3] and is also somewhat different from that in [4] and [5]. From the general definition (14) for α_t, it is clear that the key adaptation property $\alpha_t \in \sigma\{\phi_i, i \leq t\}$ can be guaranteed automatically. Also, (14) together with (8) defines the class of weights of interest in an explicit way.

To analyze the WLS, we need the following standard assumptions:

A1) $\{w_t, \mathcal{F}_t\}$ is a martingale difference sequence defined on the basic probability space $(\Omega, \mathcal{F}, P)$ with

$$\sup_{t \geq 0} E[w_{t+1}^2 \,|\, \mathcal{F}_t] < \infty \quad \text{a.s.} \tag{19}$$

A2) The input sequence $\{u_t\}$ is adapted to $\{\mathcal{F}_t\}$.

A3) $C^{-1}(z) - \frac{1}{2}$ is strictly positive real, i.e.,

$$\max_{|z|=1} |C(z) - 1| < 1.$$

Under these conditions, the following facts parallel to those of the standard LS hold (cf. Bercu [5]).

Lemma 1: Let the ARMAX model (3) satisfy the conditions A1)–A3). Then the WLS described by (9)–(14) has the following properties:

i) $\|P_{t+1}^{-1/2}\widetilde{\theta}_{t+1}\|^2 = O(1)$ a.s.

ii) $\sum_{t=1}^{\infty} \alpha_t[(\phi_t^\tau \widetilde{\theta}_{t+1})^2 + (\widehat{w}_t - w_t)^2] < \infty$ a.s.

iii) $\sum_{t=1}^{\infty} \frac{(\phi_t^\tau \widetilde{\theta}_t)^2}{\alpha_t^{-1} + \phi_t^\tau P_t \phi_t} < \infty$ a.s.

where $\widetilde{\theta}_t \triangleq \theta - \theta_t$.

Remark 2: In the white noise case ($r = 0$), the WLS is precisely the standard LS for the following regularized linear regression

$$\overline{y}_{t+1} = \theta^\tau \overline{\phi}_t + \overline{w}_{t+1}$$

where $[\overline{y}_{t+1}, \overline{\phi}_t, \overline{w}_{t+1}] \triangleq \sqrt{\alpha_t}[y_{t+1}, \phi_t, w_{t+1}]$. Hence, by the standard analysis for LS (see, e.g., [6, (A.3)]) we have

$$\begin{aligned}\widetilde{\theta}^\tau_{k+1} P^{-1}_{k+1} \widetilde{\theta}_{k+1} = \widetilde{\theta}^\tau_k P^{-1}_k \widetilde{\theta}_k &- \frac{(\phi^\tau_k \widetilde{\theta}_k)^2}{\alpha^{-1}_k + \phi^\tau_k P_k \phi_k} \\ &- \frac{2\phi^\tau_k \widetilde{\theta}_k w_{k+1}}{\alpha^{-1}_k + \phi^\tau_k P_k \phi_k} \\ &+ \alpha^2_k \phi^\tau_k P_{k+1} \phi_k w^2_{k+1}. \end{aligned} \tag{20}$$

Now, by (11) and (14) we have ($c_0 \triangleq (\max\{1, \alpha_0\})^{p+q+r}$)

$$\begin{aligned}\det(P^{-1}_{t+1}) &= \det(\sum_{i=0}^{t} \alpha_i \phi_i \phi^\tau_i + P^{-1}_0) \\ &\leq (\sum_{i=0}^{t} \alpha_i \|\phi_i\|^2 + \|P^{-1}_0\|)^{p+q+r} \\ &\leq c_0 r^{p+q+r}_t.\end{aligned}$$

Since $f(\cdot)$ is slowly increasing, by (16) in Remark 1 there exists $c > 0$ such that

$$f(c^{-1}_0 \det(P^{-1}_{t+1})) \leq f(r^{p+q+r}_t) \leq c f(r_t)$$

and consequently, by (14) we have $\alpha_t \leq c / f(c^{-1}_0 \det(P^{-1}_{t+1}))$. From this, a similar treatment as that for the proof of (30)–(37) in [15] gives

$$\sum_{t=1}^{\infty} \alpha^2_t \phi^\tau_t P_{t+1} \phi_t = \sum_{t=1}^{\infty} \alpha_t \cdot \overline{\phi}^\tau_t P_{t+1} \overline{\phi}_t < \infty.$$

Now let us write

$$\begin{aligned}\sum_{k=0}^{\infty} \alpha^2_k \phi^\tau_k P_{k+1} \phi_k w^2_{k+1} &= \sum_{k=0}^{\infty} \alpha^2_k \phi^\tau_k P_{k+1} \phi_k \\ &\quad \times (w^2_{k+1} - E[w^2_{k+1} | \mathcal{F}_k]) \\ &\quad + \sum_{k=0}^{\infty} \alpha^2_k \phi^\tau_k P_{k+1} \phi_k E[w^2_{k+1} | \mathcal{F}_k].\end{aligned}$$

The first term on the right-hand side is convergent a.s. by Chow's martingale convergence theorem (cf. [14, p. 36]), and the second term is convergent since condition (19) holds. Hence, the summation of the last term in (20) is bounded a.s. (thanks to the choice of the weights).

As for the second last term in (20) ([14, Theorem 2.8, p. 41]), we have for any $\delta > 0$

$$\left| \sum_{k=0}^{t} \frac{\phi^\tau_k \widetilde{\theta}^\tau_k w_{k+1}}{\alpha^{-1}_k + \phi^\tau_k P_k \phi_k} \right| = O\left(\left\{ \sum_{k=0}^{t} \frac{(\phi^\tau_k \widetilde{\theta})^2}{\alpha^{-1}_k + \phi^\tau_k P_k \phi_k} \right\}^{\frac{1}{2}+\delta} \right) \text{ a.s.}$$

Finally, summing up both sides of (20) from $k = 0$ to t, it is easy to see that Lemma 1 holds. (Note that since both the noise condition A1) and the definition of weights $\{\alpha_t\}$ here are somewhat different from those used in [5], there are some necessary differences between the proofs here and there).

The general $r \geq 0$ case can be proved similarly by using the standard treatment for ELS together with A3) and the properties of $\{\alpha_t\}$ (see Theorem 1 of Bercu [5] for related results and analysis). □

Based on Lemma 1, we may now prove the following main result of this section.

Theorem 1: Let the ARMAX model (3) satisfy A1)–A3). Then the WLS described by (9)–(14) has self-convergence property, i.e., θ_t converges almost surely to a finite random vector $\overline{\theta}$ (not necessarily equal to θ).

Proof: Set

$$\phi^0_t = [y_t \cdots y_{t-p+1} u_t \cdots u_{t-q+1} w_t \cdots w_{t-r+1}]^\tau. \tag{21}$$

Then (3) can be rewritten as

$$y_{t+1} = \theta^\tau \phi^0_t + w_{t+1}$$

and substituting this into (9) we get

$$\begin{aligned}\theta_{t+1} &= \theta_t + L_t[\widetilde{\theta}^\tau_t \phi_t + \theta^\tau(\phi^0_t - \phi_t) + w_{t+1}] \\ &= \theta_0 + \sum_{i=0}^{t} L_i[\widetilde{\theta}^\tau_i \phi_i + \theta^\tau(\phi^0_i - \phi_i) + w_{i+1}].\end{aligned} \tag{22}$$

Now, by taking trace on both sides of (11) and summing up, we have

$$\sum_{i=0}^{\infty} \frac{\|P_i \phi_i\|^2}{\alpha^{-1}_i + \phi^\tau_i P_i \phi_i} = \sum_{i=0}^{\infty} [\text{tr}(P_i) - \text{tr}(P_{i+1})] \leq \text{tr}(P_0) < \infty. \tag{23}$$

From this, Lemma 1-iii), and the Schwarz inequality, it follows that

$$\begin{aligned}\sum_{i=0}^{\infty} \|L_i \widetilde{\theta}^\tau_i \phi_i\| &\leq \left\{ \sum_{i=0}^{\infty} \frac{\|P_i \phi_i\|^2}{\alpha^{-1}_i + \phi^\tau_i P_i \phi_i} \times \sum_{i=0}^{\infty} \frac{(\widetilde{\theta}^\tau_i \phi_i)^2}{\alpha^{-1}_i + \phi^\tau_i P_i \phi_i} \right\}^{1/2} \\ &< \infty, \quad \text{a.s.}\end{aligned}$$

Hence $\sum_{i=0}^{t} L_i \phi^\tau_i \widetilde{\theta}_i$ converges almost surely. Similarly, $\sum_{i=0}^{t} L_i \theta^\tau (\phi^0_i - \phi_i)$ also converges a.s., since by Lemma 1-ii) we have

$$\begin{aligned}\sum_{i=0}^{\infty} \|L_i \theta^\tau (\phi^0_i - \phi_i)\| &\leq \|\theta\| \Big\{ \sum_{i=0}^{\infty} \frac{\|P_i \phi_i\|^2}{\alpha^{-1}_i + \phi^\tau_i P_i \phi_i} \sum_{i=0}^{\infty} \alpha_i \|\phi^0_i - \phi_i\|^2 \Big\}^{1/2} \\ &< \infty.\end{aligned}$$

As for the last term in (22), by A1), and using (23) again, we know that

$$\sum_{t=1}^{\infty} E[\|L_t w_{t+1}\|^2 | \mathcal{F}_t] < \infty \qquad \text{a.s.}$$

So by Chow's martingale convergence theorem we know that $\sum_{i=0}^{\infty} L_i w_{i+1}$ also converges a.s. Finally, combining all the above proved facts, we find from (22) that θ_t converges a.s. as desired. □

We are now in a position to give a detailed comparison between WLS and LS.

Remark 3: Theorem 1 can readily be extended to general linear regression models. Under only the measurability conditions on the regressors as used here, neither the familiar stochastic gradient (SG) nor the standard LS algorithms are known to converge. In fact, for the SG we only know that the norm $\|\theta_t - \theta\|$ converges (e.g., [14, p. 108]), while for LS even the boundedness of $\{\theta_t\}$ cannot be guaranteed (see [16, p. 367]), nor convergence. The only exceptions are the results derived in a Bayesian framework with Gaussian white noise, where it was shown that the LS estimate converges outside an exceptional set of true parameter vectors of Lebesgue measure zero (see Sternby [17] and Kumar [18]). This property has been used by Kumar [18] to analyze a variety of adaptive control schemes for minimum phase systems and is discussed in more detail by Nassiri-Toussi and Ren [16], where it is shown that the exceptional set for convergence of LS can indeed be nonempty.

Remark 4: By Lemma 1-i) we know that

$$\|\widetilde{\theta}_{t+1}\|^2 = O(\lambda_{\max}(P_{t+1})) \quad \text{a.s.} \tag{24}$$

Consequently, a sufficient condition for strong consistency of WLS is that

$$P_t \underset{t\to\infty}{\longrightarrow} 0 \quad \text{a.s.} \tag{25}$$

If $r = 0$ and the regressor ϕ_t is free of the initial condition θ_0, then $\theta_t \longrightarrow \theta$ for any initial condition if and only if (25) holds. The necessity can be proven as follows: By (9)–(11) we have

$$\widetilde{\theta}_{t+1} = P_{t+1}P_0^{-1}\widetilde{\theta}_0 - P_{t+1}\sum_{i=0}^{t}\alpha_i\phi_i w_{i+1}$$

and since $\widetilde{\theta}_t \longrightarrow 0$, $\forall \theta_0$, the last term, free of θ_0, must converge to zero, and thus $P_{t+1}P_0^{-1}\widetilde{\theta}_0 \longrightarrow 0$, $\forall\theta_0$ which necessarily implies (25).

Remark 5: Let $\lambda_{\min}(t)$ be the minimum eigenvalue of $\sum_{i=0}^{t}\phi_i\phi_i^\tau$ and the weighting sequence be taken as $\alpha_k = \dfrac{1}{(\log r_k)(\log\log r_k)^\delta}$ for some $\delta > 1$. Then we have

$$\lambda_{\min}(P_{t+1}^{-1}) = \lambda_{\min}\Big(\sum_{i=1}^{t}\alpha_i\phi_i\phi_i^\tau + P_0^{-1}\Big) \geq \alpha_t\lambda_{\min}(t)$$

consequently, by (24), we have for WLS

$$\|\widetilde{\theta}_{t+1}\|^2 = O\left(\frac{(\log r_t)(\log\log r_t)^\delta}{\lambda_{\min}(t)}\right) \quad \text{a.s.} \quad \delta > 1. \tag{26}$$

This is precisely the convergence rate established for the standard LS (see [14, p. 96] and [15]). Of course, if the moment condition in A1) is strengthened to a condition of order greater than two, then for the standard LS, (26) holds with $\delta = 0$ (see, Lai and Wei [19]). Thus, the WLS may have a mild compromise on convergence rate (in the case of convergence to θ) compared with LS. Nevertheless, due to its self-convergence, the WLS is more convenient than LS in applications to general adaptive control systems as can be seen from the following sections.

III. Uniform Controllability of Estimated Models

Let us denote

$$A(\theta) = \begin{bmatrix} -a_1 & 1 & \cdots & 0 \\ \vdots & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ -a_n & 0 & \cdots & 0 \end{bmatrix}, \quad b(\theta) = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \tag{27}$$

$$C(\theta) = [1 c_1 \cdots c_{n-1}]^\tau, \quad H = [10\cdots 0] \tag{28}$$

where $n \triangleq \max(p, q, r+1)$, and $a_i = b_j = c_k = 0$ for $i > p$, $j > q$, $k > r$. Then (3) can be written in the state-space form

$$\begin{aligned} x_{t+1} &= A(\theta)x_t + b(\theta)u_t + C(\theta)w_{t+1} \\ y_t &= Hx_t, \quad x_0 = [y_0, 0\cdots 0]^\tau. \end{aligned} \tag{29}$$

Many standard control law designs depend on the controllability of the system (29), and hence we make the following assumption:

A4) The pair $[A(\theta), b(\theta)]$ is controllable where θ is the true system parameter.

It is a standard result from linear system theory that controllability of $[A(\theta), b(\theta)]$ is equivalent to the coprimeness of the polynomials $A(z)$ and $B(z)$ defined by (4) and (5) and which in turn is equivalent to the nonsingularity of the Sylvester resultant (eliminant) matrix associated with θ. [Note that A4) implies that either a_n or b_n in (27) is nonzero, and hence $n = \max(p, q)$.]

Now, let θ_t be the WLS estimate discussed in the last section and $\overline{\theta}$ be its limit. If for the initial condition θ_0 the pair $[A(\theta_0), b(\theta_0)]$ is controllable, u_t is a rational function of $\{y_0, \cdots, y_t\}$, and all finite dimensional distributions of the noise process $\{w_t\}$ are absolutely continuous with respect to Lebesgue measure, then, following Meyn and Caines [15] or Caines [16], we may prove that $[A(\theta_t), b(\theta_t)]$ is controllable a.s. for each $t \geq 1$ (see Appendix B).

This, however, does not guarantee that $[A(\theta_t), b(\theta_t)]$ is uniformly controllable, since the limiting model $[A(\overline{\theta}), b(\overline{\theta})]$ may be uncontrollable. Thus, modifications on $\{\theta_t\}$ seem to be necessary.

Here, we first follow an idea similar to that used by Lozano [11] and Lozano and Zhao [12] for such a modification. Let us denote $\beta_t^* = P_t^{-1/2}(\theta - \theta_t)$. Then, by Lemma 1-i) we know that the sequence $\{\beta_t^*\}$ is bounded almost surely, and that

$$\theta = \theta_t + P_t^{1/2}\beta_t^*. \tag{30}$$

As observed in [11] and [12], although $\{\beta_t^*\}$ is not available here, given A4), this formula suggests the possibility of finding a bounded adapted sequence $\{\beta_t, \mathcal{F}_t\}$ such that the modified estimate

$$\widehat{\theta}_t \triangleq \theta_t + P_t^{1/2}\beta_t \tag{31}$$

corresponds to a uniformly controllable model. We may call $\widehat{\theta}_t$ the WLS-based estimate.

The key point of the modification in (31) is that $\{\widehat{\theta}_t\}$ possesses almost the same nice properties as those of the WLS estimate θ_t, as demonstrated in the following lemma (see Appendix C for the proof).

Lemma 2: Let $\widehat{\theta}_t$ be defined by (31) with θ_t being the WLS estimate defined by (9)–(14) and $\{\beta_t\}$ being any bounded sequence. Then under conditions of Lemma 1 we have:

i) $\|P_t^{-1/2}\overline{\theta}_t\| = O(1)$ a.s.

ii) $\sum_{i=1}^{t}[(\phi_i^\tau\overline{\theta}_{i+1})^2 + (\widehat{w}_i - w_i)^2] = O(\alpha_t^{-1}\log r_t)$ a.s.

iii) $\sum_{i=1}^{t}(\phi_i^\tau\overline{\theta}_i)^2 = o(r_t) + O(1)$ a.s.

where $\overline{\theta}_t = \theta - \widehat{\theta}_t$, and all other quantities are defined in (9)–(14).

Next, we proceed to show how $\{\beta_t\}$ can be constructed such that the pair $[A(\widehat{\theta}_t), b(\widehat{\theta}_t)]$ with $\widehat{\theta}_t$ defined by (31) is a.s. uniformly controllable.

For $\theta = [-a_1 \cdots a_p\, b_1 \cdots b_q\, c_1 \cdots c_r]^\tau \in I\!R^d$, $(d = p+q+r)$, let the Sylvester resultant matrix $M(\theta)$ be defined by

$$M(\theta) = \begin{bmatrix} 1 & & 0 & 0 & & 0 \\ a_1 & \ddots & & b_1 & \ddots & \\ \vdots & \ddots & 1 & \vdots & \ddots & 0 \\ a_n & & a_1 & b_n & & b_1 \\ & \ddots & \vdots & & \ddots & \vdots \\ 0 & & a_n & 0 & & b_n \end{bmatrix} \tag{32}$$

where $n = \max(p, q)$. (Note that only the first $p+q$ components of θ are used in the above definition).

Now we introduce the following function

$$f_t(x) = |\det M(\theta_t + P_t^{1/2}x)|, \quad x \in I\!R^d,\ t \geq 0. \tag{33}$$

It is clear that $\{\beta_t\}$ should be chosen such that $f_t(\beta_t)$ is bounded from below. If we look at this problem from an optimization point of view, then an intuitive way for doing so is to let β_t be as close to the maxima of the function $f_t(x)$ as possible for all $t \geq 0$.

Instead of using the deterministic search method for choosing β_t as in Lozano and Zhao [12], here we use an optimization-based random search method which has the advantage that only two matrix determinants are needed to be calculated and compared at each step t. To be specific, let D be any compact subset of $I\!R^d$ which coincides with the closure of its interior. For example, D may be taken as simple domains like the unit ball $\{x \in I\!R^d : \|x\| \leq 1\}$ or the unit cube

$$\{x = (x_1 \cdots x_d)^\tau \in I\!R^d : 0 \leq x_i \leq 1, \quad 1 \leq i \leq d\}. \tag{34}$$

Let $\{\eta_t\}$ be an independent sequence of d-dimensional random vectors which are uniformly distributed on D. Also let $\{\eta_t\}$ and $\{w_t\}$ be independent. Take a number $\gamma > 0$ small enough so that

$$1 \geq 2\gamma + \gamma^2. \tag{35}$$

Finally, the sequence $\{\beta_t\}$ can be recursively defined as follows

$$\beta_t = \begin{cases} \eta_t, & \text{if } f_t(\eta_t) \geq (1+\gamma)f_t(\beta_{t-1}) \\ \beta_{t-1}, & \text{if } f_t(\eta_t) < (1+\gamma)f_t(\beta_{t-1}) \end{cases} \tag{36}$$

for all $t \geq 1$, where the initial condition is $\beta_0 = \eta_0$.

Obviously, for any adaptive control sequence $\{u_t\}$, $\{\beta_t, \mathcal{F}_t'\}$ is an adapted sequence where $\mathcal{F}_t' = \sigma\{w_i, \eta_i, i \leq t\}$. Note that, similar to the deterministic case (e.g., [12] and [24]), the introduction of the hysteresis constant γ in (36) plays a role of ensuring the convergence of $\{\beta_t\}$ in sample path. The following key theorem states that $\{\beta_t\}$ defined by (36) does indeed meet our requirements (see Appendix D for the proof).

Theorem 2: Let A1)–A4) hold for the ARMAX model (3). Then, the WLS-based parameter estimate $\widehat{\theta}_t$ defined by (31) with β_t chosen as in (36) has the following properties:

i) $\{\widehat{\theta}_t\}$ converges almost surely.

ii) $[A(\widehat{\theta}_t), b(\widehat{\theta}_t)]$ is uniformly controllable a.s.

iii) All properties of Lemma 2 hold.

IV. Adaptive Pole-Placement Control

Let $A^*(z)$ be an arbitrary stable polynomial of degree $2n-1$ with $n = \max(p, q)$. Then by A4) we know that there exist unique polynomials $L(z)$ and $R(z)$, both of order $(n-1)$ with $L(0) = 1$, such that

$$A(z)L(z) + B(z)R(z) = A^*(z). \tag{37}$$

Now, if the feedback control law is generated by

$$L(z)u_t = R(z)\{y_t^* - y_t\} \tag{38}$$

where $\{y_t^*\}$ is an arbitrary but bounded deterministic reference sequence, then the resulting closed-loop system has characteristic polynomial $A^*(z)$ and (cf. Goodwin and Sin [22])

$$A^*(z)y_t = L(z)C(z)w_t + B(z)R(z)y_t^*, \quad \forall t. \tag{39}$$

To ensure boundedness of the long-run average of the squared output process $\{y_t\}$, it is natural to require that

$$\sigma_w^2 \triangleq \limsup_{t\longrightarrow\infty} \frac{1}{t}\sum_{i=1}^{t} w_i^2 < \infty \qquad \text{a.s.} \tag{40}$$

and this condition will be assumed throughout the sequel.

Let $\{\widehat{\theta}_t\}$ be defined as in Theorem 2, from which we form the following estimated polynomials in a standard way

$$\begin{aligned} A_t(z) &= 1 + a_1(t)z + \cdots + a_n(t)z^n \\ B_t(z) &= b_1(t)z + \cdots + b_n(t)z^n \\ C_t(z) &= 1 + c_1(t)z + \cdots + c_n(t)z^n. \end{aligned} \tag{41}$$

Then by Theorem 2, the following Diophantine equation

$$A_t(z)L_t(z) + B_t(z)R_t(z) = A^*(z) \tag{42}$$

will uniquely determine polynomials $L_t(z)$ and $R_t(z)$, both of order $(n-1)$ with coefficients bounded and convergent.

From this we are able to prove that the following certainty equivalent pole-placement adaptive control law

$$L_t(z)u_t = R_t(z)\{y_t^* - y_t\} \tag{43}$$

is stabilizing (see Appendix E for the proof).

Theorem 3: Consider the ARMAX system (3) and the pole-placement adaptive control law (41)–(43) with $\{\widehat{\theta}_t\}$ defined as in Theorem 2. Then the closed-loop system is stable in the sense that

$$\limsup_{T\longrightarrow\infty} \frac{1}{T}\sum_{t=1}^{T}(y_t^2 + u_t^2) < \infty \qquad \text{a.s.}$$

Since in Theorem 3 the parameter estimate $\{\widehat{\theta}_t\}$ may not be strongly consistent, the closed-loop equation under the adaptive law (43) may not approach to the ideal one (39), in general. We now use the "attenuating excitation technique" developed in [15] and [14] to design an optimal controller.

Let $\{\epsilon_i\}$ be a bounded i.i.d. sequence of random variables independent of $\{w_i, \eta_i\}$ with zero mean and unit variance, and let u_t^0 be defined by

$$L_t(z)u_t^0 = R_t(z)\{y_t^* - y_t\}. \tag{44}$$

The actual system input is taken as

$$u_t = u_t^0 + \frac{\epsilon_t}{t^{\epsilon/2}}, \qquad \epsilon \in (0, \frac{1}{4n}). \tag{45}$$

Theorem 4: Consider the ARMAX system (3) and the adaptive control law defined by (42), (44), and (45), with $\{\widehat{\theta}_t\}$ defined as in Theorem 2. Then the closed-loop system is asymptotically optimal in the sense that

$$\frac{1}{T}\sum_{t=1}^{T}\left[A^*(z)y_t - L(z)C(z)w_t - B(z)R(z)y_t^*\right]^2 \underset{T\to\infty}{\longrightarrow} 0 \qquad \text{a.s.}$$

The proof is also given in Appendix E.

V. Adaptive LQG Control

Consider the following quadratic cost function

$$J(u) = \limsup_{T\longrightarrow\infty} \frac{1}{T}\sum_{i=0}^{T-1}[(y_i - y_i^*)^2 + \lambda u_i^2] \tag{46}$$

where $\lambda > 0$, and $\{y_i^*\}$ is a known bounded deterministic reference signal.

Define the set U of admissible controls

$$U = \{u : \sum_{i=0}^{t-1}(u_i^2 + \|x_i\|^2) = O(t),\ \|x_t\|^2 = o(t) \text{ a.s. } u_t \in \mathcal{F}_t\ \forall t\} \tag{47}$$

where $\{x_i\}$ is the state vector defined by (29). Then the optimal control minimizing $J(u)$ in U is (cf. [14, p. 76])

$$u_t = Lx_t + d_t \tag{48}$$

where

$$L = -(\lambda + b^\tau Sb)^{-1}b^\tau SA \tag{49}$$
$$d_t = -(\lambda + b^\tau Sb)^{-1}b^\tau g_{t+1} \tag{50}$$
$$g_t = -\sum_{j=0}^{\infty} F^{j\tau}H^\tau y_{t+j}^* = F^\tau g_{t+1} - H^\tau y_t^* \tag{51}$$
$$S = A^\tau SA - A^\tau Sb(\lambda + b^\tau Sb)^{-1}b^\tau SA + H^\tau H \tag{52}$$
$$F = A - b(\lambda + b^\tau Sb)^{-1}b^\tau SA. \tag{53}$$

Here we have written A and b for $A(\theta)$ and $b(\theta)$ defined by (27), for simplicity. H is defined by (28).

Since (A, H) is observable and (A, b) is controllable by A4), it is known that $S > 0$ is the unique positive solution of (52) and F defined by (53) is a stable matrix (cf., e.g., [14, p. 75]).

Now, let $\widehat{\theta}_t$ be the estimate for θ which is defined as in Theorem 2, and let A_t, b_t, and C_t stand for $A(\widehat{\theta}_t)$, $b(\widehat{\theta}_t)$, and $C(\widehat{\theta}_t)$. The certainty equivalence LQG control takes the form

$$u_t = L_t\widehat{x}_t + \widehat{d}_t \tag{54}$$

where

$$L_t = -(\lambda + b_t^\tau S_t b_t)^{-1}b_t^\tau S_t A_t \tag{55}$$
$$\widehat{d}_t = -(\lambda + b_t^\tau S_t b_t)^{-1}b_t^\tau \widehat{g}_{t+1} \tag{56}$$
$$\widehat{g}_t = -\sum_{j=0}^{\infty} F_t^{j\tau}H^\tau y_{t+j}^* \tag{57}$$
$$F_t = A_t - b_t(\lambda + b_t^\tau S_t b_t)^{-1}b_t^\tau S_t A_t \tag{58}$$

and $\{S_t\}$ is recursively defined by

$$S_t = A_t^\tau S_{t-1}A_t - A_t^\tau S_{t-1}b_t(\lambda + b_t^\tau S_{t-1}b_t)^{-1}b_t^\tau S_{t-1}A_t + H^\tau H, \qquad S_0 \geq 0. \tag{59}$$

Also in (54), $\{\widehat{x}_t\}$ is the estimate for $\{x_t\}$ which is generated by the following adaptive filter

$$\widehat{x}_{t+1} = A_t\widehat{x}_t + b_t u_t + C_t[y_{t+1} - HA_t\widehat{x}_t - Hb_t u_t]$$
$$\widehat{x}_0 = [y_0, 0, \cdots, 0]^\tau. \tag{60}$$

To get an optimal adaptive control law, we use the "attenuating excitation technique" again. Let $\{\epsilon_t\}$ be the same as that in (45) and $\{u_t\}$ be defined by (54)–(59). Then define

$$u_t^* = u_t + \frac{\epsilon_t}{t^{\epsilon/2}}, \qquad \epsilon \in (0, \frac{1}{4n}). \tag{61}$$

The following result is proven in Appendix F.

Theorem 5: Consider the ARMAX model (3) where A1), A3), and A4) hold. Let $\{\widehat{\theta}_t\}$ be defined as in Theorem 2 which is used in defining the adaptive LQG control laws (54) and (61). Then $\{u_t\}$ is stabilizing and $\{u_t^*\}$ is optimal, i.e.,

$$J(u) < \infty \qquad \text{and } J(u^*) = J_{\min} \qquad \text{a.s.}$$

where $J_{\min}$ is the minimum of the cost function (46) (see [14, p. 250]).

VI. Concluding Remarks

The self-convergence property of WLS together with the method of random regularization has enabled us to give a complete solution to both the stochastic adaptive pole-placement and LQG control problems for ARMAX systems, requiring only controllability of the true system together with the standard passivity condition on the noise model. The stability of the closed-loop system is achieved without introducing any excitation probing signals into the system, and the optimality is established by only incorporating with decaying excitations (which seems to be necessary).

The applications of the self-convergence property are not limited to these, and a universal procedure for the design of stochastic adaptive control laws for ARMAX models may be formed as follows: i) To get a self-convergent parameter estimate via WLS, ii) To modify the WLS estimate via (31), so that the estimated model is uniformly controllable (or satisfy other requirements, depending on specific applications), and iii) To form the adaptive control law via the certainty equivalence principle (incorporating with the "attenuating excitation technique," if necessary). Finally, we remark that generalizations of the results of the paper to multidimensional ARMAX models as studied, in e.g., [14], are straightforward, and the proofs are only notationally more complex.

Appendix A
Proofs of (15)–(17)

First, for any $f(x) \in F$ and for large $x > 0$

$$\frac{\log x}{f(x)} = \frac{2}{f(x)} \int_{\sqrt{x}}^{x} \frac{dt}{t} \leq 2 \int_{\sqrt{x}}^{x} \frac{dt}{t f(t)} = o(1), \quad \text{as} \quad x \longrightarrow \infty$$

which proves (15). Next, for any $M > 0$, there exists an integer $m \geq 1$ such that $M \leq 2^m$. Hence, following the proof ideas in [25, p. 276], we have for $x \geq 1$

$$\frac{f(x^M)}{f(x)} \leq \frac{f(x^{2^m})}{f(x)} = \prod_{i=0}^{m-1} \frac{f(x^{2^{i+1}})}{f(x^{2^i})} = O(1)$$

which is (16). Furthermore, for $x \in [e^{2^k}, e^{2^{k+1}})$, $k \geq 1$

$$\begin{aligned} \log f(x) \leq \log f(e^{2^{k+1}}) &= \log \prod_{i=1}^{k} \frac{f(e^{2^{i+1}})}{f(e^{2^i})} + \log f(e^2) \\ &= O(k) = O(\log \log e^{2^k}) \\ &= O(\log \log x) \end{aligned}$$

which implies (17). Hence the proof is completed.

Appendix B
Controllability of $[A(\theta_t), b(\theta_t)]$

Let $M(\theta)$ be the Sylvester resultant matrix defined by (32). We only need to prove that $\det M(\theta_t) \neq 0$ a.s. $\forall t$. For this, let $N(\theta)$ be defined in the same way as $M(\theta)$ but with the elements' 1's in the diagonal replaced by 0. Then, by (9), we have

$$\begin{aligned} M(\theta_{t+1}) &= M(\theta_t) + N(L_t(y_{t+1} - \phi_t^\tau \theta_t)) \\ &= M(\theta_t) + (y_{t+1} - \phi_t^\tau \theta_t) N(L_t). \end{aligned}$$

Let us assume that $\det M(\theta_t) \neq 0$ a.s. for some $t \geq 0$. Then

$$\begin{aligned} &\det M(\theta_{t+1}) \\ &\quad = \det M(\theta_t) \cdot \det[I + (y_{t+1} - \phi_t^\tau \theta_t) M^{-1}(\theta_t) N(L_t)]. \end{aligned}$$

Obviously

$$f(w_0, \cdots, w_{t+1}) \triangleq \det[I + (y_{t+1} - \phi_t^\tau \theta_t) M^{-1}(\theta_t) N(L_t)]$$

is a rational function of $\{w_0, \cdots, w_{t+1}\}$. If it is a constant, then set $w_{t+1} = \phi_t^\tau \theta_t - \theta^\tau \phi_t^0$, where ϕ_t^0 is defined by (21), and we get $(y_{t+1} - \phi_t^\tau \theta_t) = 0$; hence this constant must be 1 and $M(\theta_{t+1})$ is nonsingular a.s. by our assumption. If $f(w_0, \cdots, w_{t+1})$ is not a constant, then by the absolute continuity of the distribution of $\{w_i\}$, we know from Meyn and Caines [20] or Caines [21] that $f(w_0, \cdots, w_{t+1}) \neq 0$ a.s., and hence $M(\theta_{t+1})$ is again nonsingular a.s. This completes the induction proof.

Appendix C
Proof of Lemma 2

i) By Lemma 1-i) and (31) we have

$$\begin{aligned} \|P_t^{-1/2} \overline{\theta}_t\| &= \|P_t^{-1/2}(\theta - \theta_t - P_t^{1/2} \beta_t)\| \\ &\leq \|P_t^{-1/2} \widetilde{\theta}_t\| + \|\beta_t\| = O(1). \end{aligned}$$

ii) By Lemma 1-ii) and the Kronecker Lemma

$$\sum_{i=1}^{t} [(\phi_i^\tau \widetilde{\theta}_{i+1})^2 + (\widehat{w}_i - w_i)^2] = O(\alpha_t^{-1}) \qquad \text{a.s.}$$

but

$$\begin{aligned} \sum_{i=1}^{t} (\phi_i^\tau \overline{\theta}_{i+1})^2 &\leq 2 \sum_{i=1}^{t} [(\phi_i^\tau \widetilde{\theta}_{i+1})^2 + (\phi_i^\tau P_{i+1}^{1/2} \beta_{i+1})^2] \\ &= O(\alpha_t^{-1}) + O(\sum_{i=1}^{t} \phi_i^\tau P_{i+1} \phi_i) \\ &= O(\alpha_t^{-1}) + O(\alpha_t^{-1} \sum_{i=1}^{t} \alpha_i \phi_i^\tau P_{i+1} \phi_i) \\ &= O(\alpha_t^{-1} \log r_t) \end{aligned}$$

and hence the assertion ii) holds.

iii) First, Lemma 1-iii) together with (18) in Remark 1 implies that $\sum_{t=1}^{\infty} (\phi_t^\tau \widetilde{\theta}_t)^2 / r_t < \infty$ a.s. So, by the Kronecker Lemma, $\sum_{i=1}^{t} (\phi_i^\tau \widetilde{\theta}_i)^2 = o(r_t) + O(1)$ a.s. Now, since $P_i - P_{i+1} \longrightarrow 0$ as $i \longrightarrow \infty$

$$\begin{aligned} \sum_{i=1}^{t} (\phi_i^\tau \overline{\theta}_i)^2 &\leq 2 \sum_{i=1}^{t} [(\phi_i^\tau \widetilde{\theta}_i)^2 + (\phi_i^\tau P_i^{1/2} \beta_i)^2] \\ &= o(r_t) + O(\sum_{i=1}^{t} \phi_i^\tau P_i \phi_i) \end{aligned}$$

$$\begin{aligned}&=o(r_t)+O(\sum_{i=1}^{t}\phi_i^\tau P_{i+1}\phi_i)\\&\quad+O(\sum_{i=1}^{t}\phi_i^\tau(P_i-P_{i+1})\phi_i)\\&=o(r_t)+O(\alpha_t^{-1}\log r_t)+o(\sum_{i=1}^{t}\|\phi_i\|^2)\\&=o(r_t)+O(1).\end{aligned}$$

This completes the proof.

Appendix D
Proof of Theorem 2

The proof is divided into the following four steps.

Step 1: We first prove that for any $t \geq 0$, the pair $[A(\widehat{\theta}_t), b(\widehat{\theta}_t)]$ is controllable, a.s., or $f_t(\beta_t) \neq 0$, a.s., where $f_t(x)$ is defined by (33).

For this, we only need to show that $f_t(\eta_t) \neq 0$ a.s., $\forall t \geq 0$; since by the definition (36) we have

$$f_t(\beta_t) \geq \frac{f_t(\eta_t)}{1+\gamma}, \quad \forall t \geq 0. \tag{62}$$

We need the following fact which follows from the proof in (e.g. [21, pp. 778–780]): Let $P(x_1 \cdots x_d)$ be a real-valued nonzero polynomial of d real variables. Then $m((x_1 \cdots x_d): P(x_1 \cdots x_d) = 0) = 0$, where $m(\cdot)$ is the Lebesgue measure on $I\!R^d$.

Now, by (30) and A4) we know that $\det M(\theta_t + P_t^{1/2}x)$ is an a.s. nonzero polynomial for any $t \geq 0$. Consequently, by the above fact we know that

$$m(x : f_t(x) = 0) = 0 \quad \text{a.s.} \quad \forall t \geq 0$$

which implies that the uniform probability measure $\mu(\cdot)$ defined on D satisfies

$$\mu(x \in D : f_t(x) = 0) = 0, \quad \text{a.s.} \quad \forall t \geq 0$$

since $\mu(\cdot)$ is absolutely continuous with respect to $m(\cdot)$.

Note that for any adaptive input $\{u_t\}$, the random process $f_t(\cdot)$ is measurable with respect to the σ-algebra $\sigma\{w_i, \eta_{i-1}, i \leq t\} \triangleq \mathcal{G}_{t-1}$. Let $I(\cdot)$ denote the indicator function of a set. By the independence of η_t and $\mathcal{G}_{t-1}$ and [14, Theorem 1.8], we have

$$\begin{aligned}EI(f_t(\eta_t)=0) &= E\{E[I(f_t(\eta_t)=0)|\mathcal{G}_{t-1}]\}\\&= E\{\int_{x\in D} I(f_t(x)=0)\mu(dx)\}\\&= E\{\mu(x\in D : f_t(x)=0)\} = 0\end{aligned}$$

which means that $P(f_t(\eta_t) = 0) = 0$ or $f_t(\eta_t) \neq 0$ a.s., $\forall t \geq 0$. Hence the desired controllability is proven.

Step 2: Next, we prove that there exists a positive random variable $\delta_\infty > 0$ such that

$$\limsup_{t\longrightarrow\infty} f_t(\eta_t) \geq \delta_\infty \qquad \text{a.s.} \tag{63}$$

Denote

$$\begin{aligned}\delta_t &\triangleq \max_{x\in D} f_t(x)\\ D_t &\triangleq \{x \in D : f_t(x) \geq \tfrac{\delta_t}{2}\}.\end{aligned}$$

Note that θ_t, $P_t^{1/2}$ and δ_t are all $\mathcal{G}_{t-1}$-measurable, and that η_t is independent of $\mathcal{G}_{t-1}$, where $\mathcal{G}_{t-1}$ is defined as in Step 1. Then again by properties of conditional expectation (cf. [14, Theorem 1.8]), we have

$$\begin{aligned}P\Big(f_t(\eta_t) \geq \frac{\delta_t}{2}|\mathcal{G}_{t-1}\Big) &= \int_{x\in D} I(f_t(x) \geq \frac{\delta_t}{2})\mu(dx)\\&= \int_{x\in D_t}\mu(dx)\\&=\mu(D_t).\end{aligned} \tag{64}$$

We now proceed to show that $\mu(D_t) \not\longrightarrow 0$, a.s. as $t \longrightarrow \infty$.

Note that both $\{\theta_t\}$ and $\{P_t^{1/2}\}$ are convergent a.s.; we may then define a function $f(x)$ as

$$f(x) = \lim_{t\longrightarrow\infty} f_t(x), \quad \text{a.s.} \quad x \in I\!R^d. \tag{65}$$

Now, let β^* be a convergence point of $\{\beta_t^*\}$ defined in (30), then by condition A4) we know that $f(\beta^*) = |\det M(\theta)| \neq 0$. Therefore, $f(x) \not\equiv 0$, a.s., which necessarily implies that $\max_{x\in D} f(x) \neq 0$, a.s., since $f(x)$ is the absolute value of a real polynomial (with variables being the components of x), and since $m(D) > 0$. Furthermore, it is easy to see that $f_t(x)$ converges to $f(x)$ uniformly on D. Consequently, $\delta_t \longrightarrow \delta_\infty$ with $\delta_\infty = \max_{x\in D} f(x) > 0$ a.s.

Since $f(x)$ is a continuous function, we have $m(D_\infty) > 0$ a.s., where D_∞ is defined by

$$D_\infty = \{x \in D : f(x) \geq \lambda\delta_\infty\}, \quad \frac{1}{2} < \lambda < 1.$$

Hence it is easy to see from the convergence of $\{f_t(x), \delta_t\}$ to $\{f(x), \delta_\infty\}$ that for sufficiently large t, $m(D_t) \geq m(D_\infty)$, which implies $\mu(D_t) \not\longrightarrow 0$ a.s. since $\mu(D_t) = \dfrac{m(D_t)}{m(D)}$.

Hence, by (64) we know that

$$\sum_{t=1}^{\infty} P\Big(f_t(\eta_t) \geq \frac{\delta_t}{2}|\mathcal{G}_{t-1}\Big) = \infty, \quad \text{a.s.}$$

Consequently, by the Borel–Cantelli–Lévy Lemma (cf. Theorem 2.5 in [14]) we have

$$\sum_{t=1}^{\infty} I\Big(f_t(\eta_t) \geq \frac{\delta_t}{2}\Big) = \infty \quad \text{a.s.}$$

which implies that

$$\limsup_{t\longrightarrow\infty} f_t(\eta_t) \geq \frac{1}{2}\lim_{t\longrightarrow\infty}\delta_t = \frac{\delta_\infty}{2} > 0 \quad \text{a.s.}$$

hence (63) is proved.

Step 3: We now prove that there exist positive random variables δ and t_0 such that

$$f(\beta_t) \geq \delta \quad \text{a.s.}, \quad \forall t \geq t_0 \tag{66}$$

where $f(x)$ is defined by (65).

By (62) and (63) it is obvious that

$$\limsup_{t\longrightarrow\infty} f_t(\beta_t) \geq \frac{\delta_\infty}{1+\gamma} > 0 \quad \text{a.s.} \tag{67}$$

From this and the uniform convergence of $f_t(x)$ to $f(x)$ on D, it is easy to see that there exists a positive random variable $\delta > 0$ and a random time $t_0 > 0$ such that

$$f_{t_0}(\beta_{t_0}) \geq 2\delta > 0 \tag{68}$$

and

$$|f(\beta_s) - f_t(\beta_s)| \leq \gamma^2\delta, \quad \forall t \geq t_0, \forall s \geq 0 \tag{69}$$

where γ is given as in (36).

We now prove (66) by induction. First, for $t = t_0$, by (68) and (69)

$$f(\beta_{t_0}) \geq f_{t_0}(\beta_{t_0}) - \gamma^2\delta \geq (2-\gamma^2)\delta \geq \delta.$$

Next, assume that (66) holds for $t = k \geq t_0$, i.e., $f(\beta_k) \geq \delta$, we need to consider the case where $t = k+1$. If $\beta_{k+1} = \beta_k$, then (66) is true by the induction assumption. Otherwise, if $\beta_{k+1} \neq \beta_k$ then by the definition (36), we know that

$$f_{k+1}(\beta_{k+1}) \geq (1+\gamma) f_{k+1}(\beta_k) \tag{70}$$

from this, (69), and $f(\beta_k) \geq \delta$, we have

$$\begin{aligned} f(\beta_{k+1}) &\geq f_{k+1}(\beta_{k+1}) - \gamma^2\delta \\ &\geq (1+\gamma) f_{k+1}(\beta_k) - \gamma^2\delta \\ &\geq (1+\gamma)[f(\beta_k) - \gamma^2\delta] - \gamma^2\delta \\ &\geq [(1+\gamma)(1-\gamma^2) - \gamma^2]\delta \geq \delta \end{aligned} \tag{71}$$

where for the last inequality we have used (35). Hence (66) is proven.

Step 4: Finally, we prove that all the results of Theorem 2 hold. By Lemma 2 we need only to prove the properties i) and ii).

Let us first show that the limit $\lim_{t\longrightarrow\infty} f(\beta_t) = f$ exists and $f > 0$ a.s. For this we need only to show that $f(\beta_k)$ is nondecreasing for $k \geq t_0$. This is true since when $\beta_{k+1} \neq \beta_k$ with $k \geq t_0$ we have from (71) and (66) that

$$\begin{aligned} f(\beta_{k+1}) \geq&\ f(\beta_k) + \gamma f(\beta_k) - (1+\gamma)\gamma^2\delta - \gamma^2\delta \\ \geq&\ f(\beta_k) + \delta\gamma[1 - 2\gamma - \gamma^2] \\ \geq&\ f(\beta_k). \end{aligned}$$

Furthermore, by the uniform convergence of $f_t(x)$ to $f(x)$ we know that $\lim_{t\longrightarrow\infty} f_{t+1}(\beta_t) = \lim_{t\longrightarrow\infty} f_t(\beta_t) = f > 0$ a.s. This together with the result in Step 1 implies that property ii) is true.

For proving i) we need only to prove that $\{\beta_t\}$ is convergent a.s. This is true, because, otherwise, by the definition (36), (70) would hold for infinitely many k on a set Γ with $P(\Gamma) > 0$. Now, let such k tend to infinity we would have

$$f \geq (1+\gamma) f \quad \text{on } \Gamma.$$

This is impossible since $\gamma > 0$ and $f > 0$ a.s. Hence the proof of Theorem 2 is completed.

Appendix E
Proof of Theorem 3

Denote

$$v_t = A_t(z) y_t - B_t(z) u_t. \tag{72}$$

Then we have [with ϕ_t^0 defined by (21)]

$$\begin{aligned} v_t &= y_t - \widehat{\theta}_t^\tau \phi_{t-1} + [C_t(z) - 1]\widehat{w}_t \\ &= w_t + \theta^\tau \phi_{t-1}^0 - \widehat{\theta}_t^\tau \phi_{t-1} + [C_t(z) - 1]\widehat{w}_t \\ &= w_t + \overline{\theta}_t^\tau \phi_{t-1} + \theta^\tau(\phi_{t-1}^0 - \phi_{t-1}) + [C_t(z) - 1]\widehat{w}_t \\ &= C(z) w_t + \overline{\theta}_t^\tau \phi_{t-1} + [C_t(z) - C(z)]\widehat{w}_t \\ &= C_t(z) w_t + \overline{\theta}_t^\tau \phi_{t-1} + [C(z) - C_t(z)](w_t - \widehat{w}_t). \end{aligned} \tag{73}$$

Hence, by Lemma 2 ii) and (18) we know that for some $L > 0$

$$\sum_{i=0}^{t} (v_t - C_t(z) w_t)^2 = O(\log^{L+1} r_t). \tag{74}$$

Now, since all the coefficients of the polynomials $A_t(z)$, $B_t(z)$, $L_t(z)$, and $R_t(z)$ are convergent, by (42), (43), and (72) with some standard manipulations, we obtain

$$\begin{aligned} A^*(z) y_t &= [A_t(z) L_t(z) + B_t(z) R_t(z)] y_t \\ &= L_t(z)[A_t(z) y_t] + B_t(z)[R_t(z) y_t] \\ &\quad + o(\max_{0\leq i\leq 2n}\{|y_{t-i}|\}) \\ &= L_t(z)[B_t(z) u_t + v_t] + B_t(z)[R_t(z) y_t^* - L_t(z) u_t] \\ &\quad + o(\max_{0\leq i\leq 2n}\{|y_{t-i}|\}) \\ &= L_t(z) v_t + [B_t(z) R_t(z)] y_t^* \\ &\quad + o(\max_{0\leq i\leq 2n}\{|y_{t-i}| + |u_{t-i}|\}). \end{aligned} \tag{75}$$

Similarly, we have

$$\begin{aligned} A^*(z) u_t = &\ -R_t(z) v_t + [A_t(z) R_t(z)] y_t^* \\ &+ o(\max_{0\leq i\leq 2n}\{|y_{t-i}| + |u_{t-i}|\}). \end{aligned} \tag{76}$$

By (40), (74)–(76), and the stability of $A^*(z)$, it is easy to convince oneself that

$$\sum_{i=1}^{t} (y_i^2 + u_i^2) = O(t) + o(r_t).$$

From this it is easy to see that $r_t = O(t)$, and the proof is complete.

Proof of Theorem 4: First, in the completely similar way as the proof of Theorem 3, it is easy to show that

$$\limsup_{T\longrightarrow\infty} \frac{1}{T} \sum_{t=1}^{T} (y_t^2 + u_t^2) < \infty \quad \text{a.s.}$$

Consequently, applying Lemma 3 of Guo and Chen [6], we know that

$$\lambda_{\min}(\sum_{i=1}^{t} \phi_i \phi_i^\tau) \geq c_0 t^{1-2\epsilon n}, \qquad \forall t$$

for some constant $c_0 > 0$. From this and (18) it follows that $P_t \longrightarrow 0$, and then by Lemma 2 (i), $\widehat{\theta}_t$ is a strongly consistent estimate for the true parameter θ. Consequently, $L_t(z)$, $B_t(z)$, and $R_t(z)$ converge to $L(z)$, $B(z)$, and $R(z)$, respectively, and the desired result follows immediately from (75) [since the decaying term $\{\frac{\epsilon_t}{t^{\epsilon/2}}\}$ in (45) does not influence (75)].

Appendix F
Proof of Theorem 5

We first prove that $\{u_t\}$ is stabilizing. By Theorem 2, $[A_t, b_t]$ converges a.s. to a controllable pair $[\widehat{A}, \widehat{b}]$, and it follows from [14, Theorem 3.4] that S_t defined by (59) converges to the unique positive solution $\widehat{S}$ of the following algebraic Riccati equation

$$\widehat{S} = \widehat{A}^\tau \widehat{S}\widehat{A} - \widehat{A}^\tau \widehat{S}\widehat{b}(\lambda + \widehat{b}^\tau \widehat{S}\widehat{b})^{-1}\widehat{b}^\tau \widehat{S}\widehat{A} + H^\tau H \qquad (77)$$

and that F_t defined by (58) converges to a stable matrix. (Note that the condition $det A \neq 0$ in [14, p. 70] can be removed since it is unnecessarily used in proving boundedness of $\{R_t\}$ there, see e.g., [23, Lemma 5.1]).

Now, let $\widehat{\theta}_t \longrightarrow \widehat{\theta}$, and rewrite the model (3) as

$$\begin{aligned} y_{t+1} &= \theta^\tau \phi_t^0 + w_{t+1} \\ &= \widehat{\theta}^\tau \phi_t^0 + w_{t+1} + (\theta - \widehat{\theta})^\tau \phi_t^0 \\ &= \widehat{\theta}^\tau \phi_t^0 + w_{t+1} + v_t \end{aligned} \qquad (78)$$

where $v_t \triangleq (\theta - \widehat{\theta})^\tau \phi_t^0 = (\theta - \widehat{\theta}_t)\phi_t^0 + (\widehat{\theta}_t - \widehat{\theta})\phi_t^0$. By Lemma 2-iii) it can be shown that

$$\sum_{i=1}^t v_i^2 = o(r_t) + O(1). \qquad (79)$$

By (78), similar to (29) we can get an alternative state space representation as

$$\begin{aligned} x_{t+1} &= \widehat{A}x_t + \widehat{b}u_t + \widehat{C}w_{t+1} + H^\tau v_t \\ y_t &= Hx_t, \qquad x_t = [y_0, 0, \cdots, 0]^\tau \end{aligned} \qquad (80)$$

where $\widehat{A}, \widehat{b}$, and $\widehat{C}$ stand for $A(\widehat{\theta})$, $b(\widehat{\theta})$, and $C(\widehat{\theta})$, and H is defined in (28).

Now, let us write

$$x_t^\tau = [x_t^1, z_t^\tau], \quad \widehat{x}_t^\tau = [\widehat{x}_t^1, \widehat{z}_t^\tau]$$

where z_t and $\widehat{z}_t$ are $(n-1)$-dimensional.

Then, similar to the proof of (8.71) in [14, p. 257] we know that

$$\begin{bmatrix} \widehat{x}_{t+1} \\ z_{t+1} - \widehat{z}_{t+1} \end{bmatrix} = \Phi_t \begin{bmatrix} x_t \\ z_t - \widehat{z}_t \end{bmatrix} + M_t \widehat{d}_t + N_t w_{t+1} + \begin{bmatrix} C_t \\ 0 \end{bmatrix} v_t \qquad (81)$$

where $\Phi_t \longrightarrow \Phi$ and Φ is a stable matrix (see, [14] p. 259), M_t and N_t are bounded vectors, and $\widehat{d}_t$ is defined by (56).

From (79) and (81), it follows that

$$\sum_{i=0}^t [\|\widehat{x}_i\|^2 + \|z_i - \widehat{z}_i\|^2] = O(t) + o(r_t)$$

by this and the fact that $\widehat{x}_t^1 \equiv x_t^1 \equiv y_t$ [this can be seen from (60) and (80)], we have

$$\sum_{i=0}^t (\|\widehat{x}_i\|^2 + \|x_i\|^2) = O(t) + o(r_t).$$

From this and (54) we find that

$$\sum_{i=1}^t (u_i^2 + y_i^2) = O(t) + o(r_t)$$

which implies that $r_t = O(t)$, and hence $\{u_t\}$ is stabilizing.

By a completely similar argument, we know that the control law $\{u_t^*\}$ defined by (61) is also stabilizing. Then similar to the proof of Theorem 4 we know that under this control law $\widehat{\theta}_t$ is a strongly consistent estimate for the true parameter θ. By this, the optimality of $\{u_t^*\}$ can easily be proved along the same lines as those in [14, pp. 264–265]. This completes the proof.

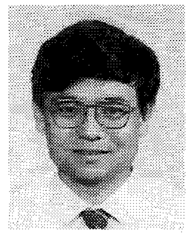
Acknowledgment

The author wishes to thank B. Bercu, H. F. Chen, R. Lozano, and J. F. Zhang for their valuable discussions and comments. The author would also like to thank W. Ren for providing the reference [16].

References

[1] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification.* Cambridge, MA: MIT Press, 1983.
[2] L. Guo, L. Ljung, and P. Priouret, "Performance analysis of the forgetting factor RLS algorithm," *Int. J. Adaptive Contr. Signal Processing*, vol. 7, pp. 525–537, 1993.
[3] R. Kumar and J. B. Moore, "Convergence of adaptive minimum variance algorithm via weighting coefficient selection," *IEEE Trans. Automat. Contr.*, vol. AC-27, no. 1, pp. 146–153, 1982.
[4] B. Bercu and M. Duflo, "Moindres Carrés Pondérés et Poursuite," *Ann. Inst. Henri Poincaré*, vol. 28, pp. 403–430, 1992.
[5] B. Bercu, "Weighted estimation and tracking for ARMAX models," *SIAM J. Contr. Optimization*, vol. 33, no. 1, pp. 89–106, 1995.
[6] L. Guo and H. F. Chen, "The Å ström-Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers," *IEEE Trans. Automat. Contr.*, vol. 36, no. 7, pp. 802–812, 1991.
[7] L. Guo, "Further results on least squares based adaptive minimum variance control," *SIAM J. Contr. Optimization*, vol. 32, no. 1, pp. 187–212, 1994.
[8] ——, "Convergence and logarithm laws of self-tuning regulators," *Automatica*, vol. 31, no. 3, pp. 435–450, 1995.
[9] S. P. Meyn and L. J. Brown, "Model reference adaptive control of time-varying and stochastic systems," *IEEE Trans. Automat. Contr.*, vol. 38, no. 12, pp. 1738–1753, 1993.
[10] W. Ren and P. R. Kumar, "Stochastic adaptive prediction and model reference control," *IEEE Trans. Automat. Contr.*, vol. 39, no. 10, pp. 2047–2060, 1994.
[11] R. Lozano, "Singularity-free adaptive pole-placement without resorting to persistency of excitation: Detailed analysis for first order systems," *Automatica*, vol. 28, pp. 27–33, 1992.
[12] R. Lozano and X. H. Zhao, " Adaptive pole placement without excitation probing signals," *IEEE Trans. Automat. Contr.*, vol. 39, no. 1, pp. 47–58, 1994.
[13] A. A. Zhigljavsky, *Theory of Global Random Search.* Boston, MA: Kluwer, 1991.
[14] H. F. Chen and L. Guo, *Identification and Stochastic Adaptive Control.* Boston, MA: Birkhäuser, 1991.
[15] ——, "Convergence rate of least-squares identification and adaptive control for stochastic systems," *Int. J. Contr.*, vol. 44, no. 5, pp. 1459–1476, 1986.

[16] K. Nassiri-Toussi and W. Ren, "On the convergence of least squares estimation in white noise," *IEEE Trans. Automat. Contr.*, vol. 39, no. 2, pp. 364–368, 1994.
[17] J. Sternby, "On consistency for the method of least squares using martingale theory," *IEEE Trans. Automat. Contr.*, vol. AC–22, pp. 346–352, 1977.
[18] P. R. Kumar, "Convergence of adaptive control schemes using least-squares estimates," *IEEE Trans. Automat. Contr.*, vol. 35, no. 4, pp. 416–424, 1990.
[19] T. L. Lai and C. Z. Wei, "Least-squares estimation in stochastic regression models with application to identification and control of dynamic systems," *Ann. Statist.*, vol. 10, pp. 154–166, 1982.
[20] S. P. Meyn and P. E. Caines, "The zero divisor problem of multivariable stochastic adaptive control," *Syst. Contr. Lett.*, vol. 6, no. 4, pp. 235–238, 1985.
[21] P. E. Caines, *Linear Stochastic Systems*. New York: Wiley, 1988.
[22] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[23] B. D. O. Anderson and J. B. Moore, "Detectability and stabilizability of time-varying discrete-time linear systems," *SIAM J. Contr. Optim.*, vol. 19, no. 1, pp. 20–32, 1981.
[24] R. H. Middleton, G. C. Goodwin, D. J. Hill, and D. Q. Mayne, "Design issues in adaptive control," *IEEE Trans. Automat. Contr.*, vol. AC-33, pp. 50–58, Jan. 1988.
[25] G. Pólya and G. Szegö, *Problems and Theorems in Analysis*, vol. 1. New York: Springer-Verlag, 1972.



Lei Guo (M'88) was born in Shandong, China, in November 1961. He received the B.S. degree in mathematics from Shandong University in 1982, and the M.S. and the Ph.D. degrees in control theory from the Chinese Academy of Sciences in 1984 and 1987, respectively.

He was a Postdoctoral Fellow at the Australian National University during 1987–1989. Since 1992, he has been a Professor with the Institute of Systems Science of the Chinese Academy of Sciences, where he is currently the Deputy Director of the Systems and Control Laboratory. He is the author of the book *Time–Varying Stochastic Systems: Stability, Estimation and Control* (in Chinese, 1993), and coauthor of the book *Identification and Stochastic Adaptive Control* (Birkhäuser, 1991).

Dr. Guo is the recipient of the Young Scientist Award of China (1994), the Natural Science Prize of the Chinese Academy of Sciences (1994), the Young Author Prize of the International Federation of Automatic Control World Congress (1993), and (with his coauthor) the National Natural Science Prize of China (1987). He was an Associate Editor of the *SIAM Journal on Control and Optimization* during 1991–1993, and is currently a standing member of the Council of the Chinese Association of Automation, and Vice Chairman of its Committee on Control Theory.