Identification and Adaptive Control for Systems with Unknown Orders, Delay, and Coefficients

HAN-FU CHEN AND JI-FENG ZHANG

Abstract—**This paper gives recursive estimates for the time-delay, system orders, and coefficients of single-input single-output linear discrete-time deterministic systems and stochastic systems with uncorrelated noise under the assumption that a lower bound for the time delay and upper bounds for system orders are known. The optimal adaptive control is designed for both tracking and linear quadratic regulation when the system parameters, including time-delay, orders, and coefficients, are unknown. The rates of convergence, both of the coefficient estimates to their true values and of the loss functions to their minima, are derived.**

I. Introduction

LET the *a priori* information about the plant be merely that it is single-input, single-output, linear, deterministic or stochastic, and that bounds for its time-delay and orders are available. The question is how to design a control to minimize a tracking error or a quadratic loss function and simultaneously to get consistent estimates for time-delay, orders, and coefficients of the system.

In time series analysis, there is extensive literature devoted to estimating orders and coefficients of a stationary ARMA process from a nonrecursive point of view; see [1]–[5]. Recently, however, Rissanen [6] established results concerning recursive order estimation. But in the above works, some sort of stationarity and ergodicity of the stochastic processes involved are usually assumed. Therefore, the previously mentioned results cannot directly be applied to an ARMAX process when the exogenous input is a feedback control so that the process is neither ergodic nor stationary.

To estimate the orders of a stochastic feedback control system, the first step was made by Chen and Guo [7], [8] who introduced a new information criterion *CIC* for both uncorrelated noise [7] and correlated noise cases [8]. Further effort in this direction was made by Hemerly and Davis [9] for multidimensional ARX systems with uncorrelated noise; by combining the PLS (predictive least squares) criterion for order estimation with an adaptive control strategy minimizing a quadratic cost, they showed that one could estimate, recursively and in a strongly consistent way, both the order and the coefficients of the controlled system, while achieving asymptotically optimal cost. However, all these papers not only need some strong assumptions because of the technical problem, but also need a great deal of computation since they require a set of parallel algorithms (one for each of the possible orders of the system) for estimating system coefficients and system states appearing in the construction of an optimal linear quadratic adaptive control.

This paper is devoted to reducing the computational load and the assumptions required in [7]–[9]. The parameters we want to estimate are not only the system orders, but also the system time-delay which is not estimated in previous works. The knowledge about the time-delay is unnecessary in some cases where adaptive tracking [10] or adaptive control with quadratic cost [11] are used without paying attention to parameter estimation, but it is crucial for some other control problems. For example, minimum variance control is sensitive to time-delay [12]. The recursion is also given for the information criteria depending on time as in [9], but the number of system coefficients we need to estimate here is much less than that estimated in [7]–[9], since we have modified the criterion *CIC* used in [8] and use only one algorithm for estimating system coefficients and system states appearing in the LQ adaptive control problem. In addition, conditions used in this paper have essentially been weakened in comparison to those in [7]–[9]. The main results of the paper can be briefly summarized as follows: for both stochastic and deterministic systems with unknown orders, time-delay, and coefficients, optimal adaptive controls are derived for tracking and quadratic regulation, respectively; rates of convergence, both of the performance index to its minimum and of the parameter estimates to their true values, are also established.

For clarity of the description, this paper deals with single-input single-output systems only. The corresponding results for multidimensional systems can be obtained similarly. The organization of this paper is as follows. Section II presents methods and criteria for estimating system orders, time-delay, and coefficients. Section III discusses sufficient conditions guaranteeing consistency of the estimates. Section IV designs an optimal adaptive tracking control which makes the estimated parameters strongly consistent, while Section V gives an optimal linear quadratic adaptive control which guarantees the strong consistency of the estimated parameters and the asymptotic minimality of the loss function. The convergence rates both of the coefficient estimates to their true values and of the loss functions to their minima are also derived in Sections IV and V. Finally, we conclude this paper in Section VI.

II. Estimation Methods for Time-Delay, Orders, and Coefficients

In this section, we present methods estimating the unknown time-delay, orders, and coefficients of a deterministic system or a stochastic system with uncorrelated noise. We start with stochastic systems.

A. Stochastic Systems

Let the single-input single-output system be described by a linear stochastic equation

$$A(z)y_n = B(z)u_n + w_n, \qquad n > 0;$$

$$y_n = u_n = w_n = 0, \qquad n \le 0, \tag{2.1}$$

where y_n, u_n, and w_n are the output, input, and noise, respectively; $A(z)$ and $B(z)$ are polynomials

$$A(z) = 1 + a_1 z + \cdots + a_{p_0} z^{p_0}, \qquad p_0 \ge 0, \tag{2.2}$$

Manuscript received August 15, 1988; revised April 25, 1989, November 24, 1989, and January 31, 1990. Paper recommended by Past Associate Editor, M. Gevers. This work was supported by the National Natural Science Foundation of China and the TWAS RG MP under Grant 898-117.

The authors are with the Institute of Systems Science, Academia Sinica, Beijing 100080, People's Republic of China.

IEEE Log Number 9036245.

0018-9286/90/0800-0866$01.00

$$B(z) = b_{d_0} z^{d_0} + \cdots + b_{q_0} z^{q_0}, \qquad q_0 \ge d_0 \ge 1 \quad (2.3)$$

in the backward shift operator z.

The coefficients a_i $(i = 1, \cdots, p_0)$, b_j $(j = d_0, \cdots, q_0)$, the time-delay d_0, and the orders (p_0, q_0) are unknown but it is assumed that a lower bound for d_0 and upper bounds for p_0, q_0 are available, i.e., integers p^*, d^* and $q^* \ge d^* \ge 1$ are given such that

$$(p_0, q_0) \in M_0 \triangleq \{(p, q): 0 \le p \le p^*, d^* \le q \le q^*\}, \quad (2.4)$$

$$d_0 \in M_d \triangleq \{d: d^* \le d \le q^*\}. \quad (2.5)$$

We now write methods for estimating d_0, (p_0, q_0) and a_i $(i = 1, \cdots, p_0)$, b_j $(j = d_0, \cdots, q_0)$.

Corresponding to the largest possible orders and the smallest possible time-delay, we take the stochastic regressor

$$\phi_n^* = [y_n \cdots y_{n-p^*+1} u_{n-d^*+1} \cdots u_{n-q^*+1}]^\tau \quad (2.6)$$

and denote unknown coefficients by

$$\theta(p, d, q) = [-a_1 \cdots -a_p b_d \cdots b_q]^\tau, \quad \theta^* = \theta(p^*, d^*, q^*) \quad (2.7)$$

where $a_i = 0$ for $i > p_0$ and $b_j = 0$ for $j < d_0$ or $j > q_0$ by definition.

Given any initial value θ_0^*, the estimate

$$\theta_n^* = [-a_{1n} \cdots -a_{p^*n} b_{d^*n} \cdots b_{q^*n}]^\tau \quad (2.8)$$

for θ^* is given by the least-squares method

$$\theta_n^* = \left(I + \sum_{i=0}^{n-1} \phi_i^* \phi_i^{*\tau} \right)^{-1} \sum_{i=0}^{n-1} \phi_i^* y_{i+1} \quad (2.9)$$

or recursively given by

$$\theta_{n+1}^* = \theta_n^* + b_n^* P_n^* \phi_n^* (y_{n+1} - \phi_n^{*\tau} \theta^*), \quad (2.10)$$

$$P_{n+1}^* = P_n^* - b_n^* P_n^* \phi_n^* \phi_n^{*\tau} P_n^*, \quad (2.11)$$

$$P_0^* = I, \; b_n^* = (1 + \phi_n^{*\tau} P_n^* \phi_n^*)^{-1}. \quad (2.12)$$

For any $(p, q) \in M_0$ and $d \in M_d$ we set

$$\theta_n(p, d, q) = [-a_{1n} \cdots -a_{pn} b_{dn} \cdots b_{qn}]^\tau, \quad (2.13)$$

$$\phi_n(p, d, q) = [y_n \cdots y_{n-p+1} u_{n-d+1} \cdots u_{n-q+1}]^\tau \quad (2.14)$$

and

$$\sigma_n(p, d, q) = \sum_{i=0}^{n-1} (y_{i+1} + a_{1n} y_i + \cdots + a_{pn} y_{i-p+1} - b_{dn} u_{i-d+1} - \cdots - b_{qn} u_{i-q+1})^2$$
$$= \sum_{i=0}^{n-1} (y_{i+1} - \theta_n^\tau(p, d, q) \phi_i(p, d, q))^2. \quad (2.15)$$

Obviously, $\theta_n(p^*, d^*, q^*) = \theta_n^*$ and $\phi_n(p^*, d^*, q^*) = \phi_n^*$.

Introduce the criteria

$$CIC_1(p)_n = \sigma_n(p, d^*, q^*) + p s_n, \quad (2.16)$$

$$CIC_2(q)_n = \sigma_n(p^*, d^*, q) + q s_n \quad (2.17)$$

and

$$CIC_3(d)_n = \sigma_n(p^*, d, q^*) - d s_n \quad (2.18)$$

where $s_n = (\log n)^2$.

Then we can estimate p_0, q_0, and d_0 respectively, as follows:

$$p_n = \arg \min_{0 \le p \le p^*} CIC_1(p)_n, \quad (2.19)$$

$$q_n = \arg \min_{0 \le q \le q^*} CIC_2(q)_n \quad (2.20)$$

and

$$d_n = \arg \min_{d^* \le d \le q_n} CIC_3(d)_n. \quad (2.21)$$

Notice that $\sigma_n(p, d, q)$ can be calculated recursively as follows:

$$\sigma_{n+1}(p, d, q) = \sigma_n(p, d, q) + (y_{n+1} - \theta_n^\tau(p, d, q)\phi_n(p, d, q))^2 + (\theta_{n+1}(p, d, q) - \theta_n(p, d, q))^\tau (N_{n+1}(p, d, q) \times \theta_{n+1}(p, d, q) + N_{n+1}(p, d, q)\theta_n(p, d, q) - 2H_{n+1}(p, d, q))$$

where

$$N_{n+1}(p, d, q) = N_n(p, d, q) + \phi_n(p, d, q)\phi_n^\tau(p, d, q), \quad N_0(p, d, q) = 0,$$

$$H_{n+1}(p, d, q) = H_n(p, d, q) + \phi_n(p, d, q) y_{n+1}, \quad H_0(p, d, q) = 0.$$

Therefore, we can also compute $CIC_1(p)_n$, $CIC_2(q)_n$, and $CIC_3(d)_n$ in a recursive way

$$CIC_1(p)_{n+1} = CIC_1(p)_n + p(s_{n+1} - s_n) + G(p, d^*, q^*)_n \quad (2.22)$$

$$CIC_2(q)_{n+1} = CIC_2(q)_n + q(s_{n+1} - s_n) + G(p^*, d^*, q)_n \quad (2.23)$$

and

$$CIC_3(d)_{n+1} = CIC_3(d)_n - d(s_{n+1} - s_n) + G(p^*, d, p^*)_n \quad (2.24)$$

where

$$G(p, d, q)_n = (y_{n+1} - \theta_n^\tau(p, d, q)\phi_n(p, d, q))^2 + (\theta_{n+1}(p, d, q) - \theta_n(p, d, q))^\tau \times (N_{n+1}(p, d, q)\theta_{n+1}(p, d, q) + N_{n+1}(p, d, q)\theta_n(p, d, q) - 2H_{n+1}(p, d, q)). \quad (2.25)$$

B. Deterministic Systems

In this section, we discuss the following deterministic system:

$$A(z) y_n = B(z) u_n, \; n > 0; \qquad y_n = u_n = 0, \; n \le 0 \quad (2.26)$$

where y_n, u_n are the scalar output and input, respectively; $A(z)$ and $B(z)$ are given by (2.2), (2.3); the unknown time-delay d_0 and the orders p_0 and q_0 are subject to (2.4), (2.5). The purpose of this section is to present a method similar to that used in the above section for estimating the unknown time-delay d_0, orders p_0 and q_0, and coefficients a_i $(i = 1, \cdots, p_0)$ and b_j $(j = d_0, \cdots, q_0)$.

For any $(p, q) \in M_0$ and $d \in M_d$ let $\theta_n(p, d, q)$,

$\phi_n(p, d, q)$, and $\sigma_n(p, d, q)$ be given by (2.13)–(2.15), where $\theta_n(p, d, q)$ is defined by

$$\theta_n(p, d, q) = \left(I + \sum_{i=0}^{n-1}\phi_i(p, d, q)\phi_i^\tau(p, d, q)\right)^{-1} \cdot \sum_{i=0}^{n-1}\phi_i(p, d, q)y_{i+1} \quad (2.27a)$$

or recursively given as follows:

$$\theta_{n+1}(p, d, q) = \theta_n(p, d, q) + b_n(p, d, q)P_n(p, d, q)\phi_n(p, d, q) \times (y_{n+1} - \phi_n^\tau(p, d, q)\theta_n(p, d, q)).$$

$$P_{n+1}(p, d, q) = P_n(p, d, q) - b_n(p, d, q)P_n(p, d, q) \times \phi_n(p, d, q)\phi_n^\tau(p, d, q)P_n(p, d, q),$$

$$b_n(p, d, q) = (1 + \phi_n^\tau(p, d, q)P_n(p, d, q)\phi_n(p, d, q))^{-1}$$

$$P_0(p, d, q) = I.$$

The estimates p_n, d_n, and q_n for p_0, d_0, and q_0 are given as follows:

$$(p_n, q_n) = \underset{\substack{0 \le p \le p^* \\ d^* \le q \le q^*}}{\arg\min}\, CIC_4(p, q)_n, \quad (2.27b)$$

$$d_n = \underset{d^* \le d \le q_n}{\arg\min}\, CIC_5(d)_n \quad (2.27c)$$

where

$$CIC_4(p, q)_n = \sigma_n(p, d^*, q) + (p + q)s_n, \quad (2.28a)$$

$$CIC_5(d)_n = \sigma_n(p_n, d, q_n) - ds_n \quad (2.28b)$$

and $\sigma_n(p, d, q)$ is given by (2.15).

Obviously, we can compute $CIC_4(p, q)_n$ recursively in a way similar to (2.22)–(2.25).

Remark 2.1: It is worth noticing that in the above order or delay estimation procedure, $CIC_1(p)_n$, $CIC_2(q)_n$, and $CIC_3(d)_n$ correspond to estimating p_0, q_0, and d_0, respectively, and can be carried out separately. Estimating p_n, d_n, and q_n here is searched only among $p^* + 2(q^* - d^*) + 2$ points at each time instant n, rather than $(p^* + 1)q^*$ points as in [7]–[9]. We also note that the time-delay d_0 is important for some adaptive control systems [10] and is not estimated in [7]–[9].

Remark 2.2: The algorithm for computing CIC in [7], [8] is nonrecursive, while here computing $CIC_1(p)_n$, $CIC_2(q)_n$, $CIC_3(d)_n$, and $CIC_4(p, q)_n$ is carried out recursively as time goes on. For stochastic systems, the criterion $CIC_4(p, q)_n$ can be applied to replace $CIC_1(p)_n$ and $CIC_2(q)_n$ as is shown in [7], [8]. Similarly, $CIC_5(d)_n$ can replace $CIC_3(d)_n$. However, the converse is not true, i.e., for deterministic systems, the criteria $CIC_4(p, q)_n$ and $CIC_5(d)_n$ cannot be replaced by $CIC_1(p)_n$, $CIC_2(q)_n$, and $CIC_3(d)_n$. This is because for deterministic systems, $\lambda^*_{\min}$ introduced in Theorem 3.1 does not go to infinity as $n \to \infty$ and hence the estimates p_n, q_n, and d_n given by (2.19)–(2.21) are not consistent.

III. Consistency Theorems of the Estimates

In this section, we give conditions guaranteeing consistency of p_n, d_n, q_n and $\theta_n(p_n, d_n, q_n)$, and state convergence results. The proof is given in Appendix A.

One may ask what the advantage is of estimating system orders and time-delay if the consistence of the parameter estimates is established since it includes convergence to zero for zero parameters. This can be explained by the fact that convergence of the order and time-delay estimates implies that the estimates (integers) exactly match the true orders and time-delay after a finite time, while one can hardly expect the coefficient estimates to be identical to the true ones even if they are consistent.

We first consider the stochastic case.

We assume that

H_1: $\{w_n, F_n\}$ is a martingale difference sequence with the following properties:

$$\sup_n E\{w_{n+1}^2 | F_n\} < \infty, \quad \text{a.s.,} \quad (3.1)$$

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} w_i^2 < \infty, \quad \text{a.s.,} \quad (3.2)$$

$$\liminf_{n\to\infty} \frac{1}{n^{1-\epsilon^*}}\sum_{i=0}^{n-1} w_i^2 > 0, \quad \text{a.s.} \quad (3.3)$$

where $\{F_n\}$ is a family of nondecreasing σ-algebras, and

$$\epsilon^* = 1/[2(t+1)], \qquad t = 2p^* + q^*. \quad (3.4)$$

Example 3.1: Let $\{w_n, F_n\}$ be a martingale difference sequence with the following properties:

$$\sup_n E\{w_{n+1}^2 | F_n\} < \infty, \quad \text{a.s.}$$

and

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} w_i^2 = R > 0,$$

then Assumption H_1 holds. Clearly, in this case

$$\liminf_{n\to\infty} \frac{1}{n^{1-\epsilon^*}}\sum_{i=0}^{n-1} w_i^2 = \infty > 0, \quad \text{a.s.}$$

Example 3.2: Let $\{w_n, F_n\}$ satisfy the conditions of Example 3.1 and let h_n be F_n-measurable and

$$C_1 n^{-\epsilon^*/2} \le h_n \le C_2, \qquad \text{for any } n \ge 1$$

where C_1 and C_2 are two constants.

Then (3.1)–(3.3) are true with w_n replaced by $e_n = h_{n-1}w_n$, for any $n \ge 2$ and $e_1 = w_1$, $e_0 = w_0$.

Theorem 3.1: If H_1 holds, u_n is F_n-measurable, and $r_n^* \triangleq 1 + \sum_{i=0}^{n}\|\phi_i^*\|^2$ satisfies

$$\frac{(\log r_n^*)(\log\log r_n^*)^c}{(\log n)^2} \underset{n\to\infty}{\longrightarrow} 0, \qquad \text{for some constant } c > 1 \quad (3.5)$$

and

$$\frac{(\log n)^2}{\lambda^*_{\min}(n)} \underset{n\to\infty}{\longrightarrow} 0 \quad (3.6)$$

where $\lambda^*_{\min}(n)$ denotes the minimum eigenvalue of $\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}$, then θ_n^*, p_n, d_n, and q_n given by (2.9)–(2.21) are strongly consistent

$$\|\theta_n^* - \theta^*\|^2 = O\left(\frac{(\log r_n^*)(\log\log r_n^*)^c}{\lambda^*_{\min}(n)}\right) \underset{n\to\infty}{\longrightarrow} 0, \quad \text{a.s.} \quad (3.7)$$

$$(p_n, d_n, q_n) \underset{n\to\infty}{\longrightarrow} (p_0, d_0, q_0), \quad \text{a.s.} \tag{3.8}$$

For deterministic systems we similarly have the following.

Theorem 3.2: If $\lambda_{\min}(n) \triangleq \min(\lambda_{\min}^{(p_0,q^*)}(n), \lambda_{\min}^{(p^*,q_0)}(n))$ satisfies

$$\frac{(\log n)^2}{\lambda_{\min}(n)} \underset{n\to\infty}{\longrightarrow} 0 \tag{3.9}$$

where $\lambda_{\min}^{(p,q)}(n)$ denotes the minimum eigenvalue of $\sum_{i=0}^{n-1} \phi_i(p, d^*, q)\phi_i^\tau(p, d^*, q)$, then (p_n, d_n, q_n) given by (2.27b), (2.27c), and (2.28) are consistent

$$(p_n, d_n, q_n) \underset{n\to\infty}{\longrightarrow} (p_0, d_0, q_0), \quad \text{a.s.} \tag{3.10}$$

and the estimate $\theta_n(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0)$ given by (2.27) is also consistent in the sense that

$$\|\theta_n(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0) - \theta(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0)\| = O((\lambda_{\min}(n))^{-1}) \tag{3.11}$$

where $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$.

Remark 3.1: Comparing the above two theorems we can see that in the deterministic case, in order to guarantee (3.9)–(3.11), coprimeness of $A(z)$ and $B(z)$ is implicitly required since otherwise condition (3.9) fails whatever u_n is; but in the stochastic case coprimeness of $A(z)$ and $B(z)$ is not necessary for consistency of estimates. This is because the coefficient polynomials $A(z)$, $B(z)$, and 1 have no common factor whatever $A(z)$ and $B(z)$ are. Condition H_1 means that the noise w_n should be neither too strong [see (3.2)] nor too weak [see (3.3)]. Too strong noise may heavily corrupt the system data, while too weak noise cannot sufficiently excite the system in order to get consistent parameter estimates. In the latter case, we then have to require some other *a priori* information. For example, in the case where $w_n \equiv 0$, we require that $A(z)$ and $B(z)$ are coprime.

We also note that the convergence $\|\theta_n^* - \theta^*\| \to 0$ [see (3.7)] is stronger than $\|\theta_n(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0) - \theta(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0)\| \to 0$ [see (3.11)] because for the latter case we know nothing about convergence of a_i, $i > p_n \vee p_0$ and b_j, $j < d_n \wedge d_0$, $j > q_n \vee q_0$.

Remark 3.2: From the proofs of Theorems 3.1 and 3.2 (see Appendix A), we know that for Theorem 3.1 s_n in criteria $CIC_1(p)_n$, $CIC_2(q)_n$, and $CIC_3(d)_n$ can be replaced by any real number sequence $\{s_n^*\}$ satisfying

$$\frac{(\log r_n^*)(\log\log r_n^*)^c}{s_n^*} \underset{n\to\infty}{\longrightarrow} 0 \text{ and } \frac{s_n^*}{\lambda_{\min}^*(n)} \underset{n\to\infty}{\longrightarrow} 0 \tag{3.12}$$

and that for Theorem 3.2 s_n in $CIC_4(p, q)_n$ and $CIC_5(d)_n$ can be replaced by any real number sequence $\{s_n^*\}$ satisfying

$$s_n^* \underset{n\to\infty}{\longrightarrow} \infty \text{ and } \frac{s_n^*}{\lambda_{\min}(n)} \underset{n\to\infty}{\longrightarrow} 0. \tag{3.13}$$

In practice, conditions (3.5), (3.6) in Theorem 3.1 and condition (3.9) in Theorem 3.2 are difficult to verify. In the following, we remove them and use alternative conditions that are easier to verify.

We know that performances of long-run average type will not be worsened if the attenuating excitation control [12], [13] is applied. Inspired by this method, as an excitation source, we take a sequence of mutually independent variables $\{v_n\}$ that is independent of $\{w_n\}$ and satisfies

$$Ev_n = 0,\ Ev_n^2 \le n^{-\epsilon},\ v_n^2 \le \sigma^2/n^\epsilon,\ \epsilon \in (0, 1/\{2(t+1)\}) \tag{3.14}$$

where t is given by (3.4) and $\sigma^2 > 0$ is a constant which can be determined by the designer.

Without loss of generality we assume that

$$F_n = \sigma\{w_i, v_i, i \le n\}, \text{ and } F_n' = \sigma\{w_i, v_{i-1}, i \le n\}.$$

Let u_n^0 be F_n'-adapted desired control. The attenuating excitation method suggests to apply

$$u_n = u_n^0 + v_n \tag{3.15}$$

to the system.

We now assume that

H_2:

$$\sum_{i=0}^{n} (u_i^0)^2 = O(n^{1+\delta}), \quad \text{for } \delta = \{1 - 2\epsilon(t+1)\}/(2t+3) \tag{3.16}$$

and

$$\sum_{i=0}^{n} y_i^2 = O(n^b), \quad \text{for some } b > 0. \tag{3.17}$$

It is clear that Assumption H_2 is not a restrictive one. For example, (3.16) is satisfied for bounded u_i^0.

For stochastic systems we then have the following theorem.

Theorem 3.3: If H_1 and H_2 hold with u_n given by (3.15), then (2.9)–(2.21) lead to

$$\|\theta_n^* - \theta^*\|^2 = O\left(\frac{(\log n)(\log\log n)^c}{n^{1-(t+1)(\epsilon+\delta)}}\right), \quad \text{for any } c > 1 \tag{3.18}$$

and

$$(p_n, d_n, q_n) \underset{n\to\infty}{\longrightarrow} (p_0, d_0, q_0), \quad \text{a.s.} \tag{3.19}$$

For deterministic systems we have the following results.

Theorem 3.4: If $A(z)$ and $B(z)$ are coprime, and system input u_n is given by (3.15) with (3.14) and (3.16) satisfied, then (2.27) and (2.28) lead to

$$(p_n, d_n, q_n) \underset{n\to\infty}{\longrightarrow} (p_0, d_0, q_0), \quad \text{a.s.} \tag{3.20}$$

and

$$\|\theta_n(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0) - \theta(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0)\| = O(n^{-(1-(t+1)(\epsilon+\delta))}), \quad \text{a.s.} \tag{3.21}$$

In these theorems (3.19) and (3.20) mean that p_n, q_n, and d_n are consistent, while (3.18) and (3.21) indicate the convergence rates of coefficient estimates to the true values.

Remark 3.3: From Remark 3.2 we know that s_n used in (2.16)–(2.18) and (2.28) can be replaced by any real number sequence $\{s_n^*\}$ satisfying

$$\frac{(\log n)(\log\log n)^c}{s_n^*} \underset{n\to\infty}{\longrightarrow} 0 \text{ and } \frac{s_n^*}{n^{1-(t+1)(\epsilon+\delta)}} \underset{n\to\infty}{\longrightarrow} 0 \tag{3.22}$$

for Theorem 3.3 and can be replaced by any real number sequence $\{s_n^*\}$ satisfying

$$s_n^* \underset{n\to\infty}{\longrightarrow} \infty \text{ and } \frac{s_n^*}{n^{1-(t+1)(\epsilon+\theta)}} \underset{n\to\infty}{\longrightarrow} 0 \tag{3.23}$$

for Theorem 3.4.

In order to avoid underestimation of orders for small n it is desirable to take smaller s_n, for example, $s_n = \log n$ in [3], [4], $s_n = \log\log n$ in [5] for stationary ARMA processes. Unfortunately for stochastic adaptive control systems, we have to take larger s_n as is shown in (3.22).

Remark 3.4: Since $t = 2p^* + q^* > \max(p_0, q_0) + p_0 - 1$, from [11, eq. (40)] it follows that the real number δ in (3.16) can be any one satisfying

$$\delta \in \left[0, \frac{1 - 2\epsilon(t+1)}{2t+3}\right]. \tag{3.24}$$

Remark 3.5: From the proof (see Appendix A) we see that one can easily generalize the results of this paper to multidimensional systems.

Remark 3.6: We note that the estimates for $\theta(p_0, d_0, q_0)$ in Theorems 3.1–3.4 consist of components of estimates for θ^*. Obviously, if the estimation is carried out off-line, we may re-estimate $\theta(p_n, d_n, q_n)$, say by ELS, after having obtained estimates p_n, d_n, q_n in order to improve the efficiency of the coefficient estimates as is done in [15]. We may also do this for adaptive control systems, but it would greatly increase the computational load.

Remark 3.7: We now compare conditions used in Theorem 3.1 of this paper to those used in [9, Theorem 2.1].

In [9], in addition to conditions used in Theorem 3.1 of this paper, it is assumed that $E\{w_n^2 | F_{n-1}\} = \sigma^2$, a.s. and

$$\phi_n^\tau(p, 1, q)\left(\sum_{i=0}^{n}\phi_i(p, 1, d)\phi_i^\tau(p, 1, q)\right)^{-1} \cdot \phi_n(p, 1, q) \underset{n\to\infty}{\longrightarrow} 0 \qquad \text{a.s.}$$

for any $(p, q) \in M_0$. Those conditions are no longer required in this paper. The condition $\sup_n E\{w_n^\alpha | F_{n-1}\} < \infty$, a.s. for some $\alpha > 2$ required in [9] is weakened to $\alpha = 2$ and the existence of the limit for $1/n\sum_{i=0}^{n} w_n^2$ is not required here. Finally, [9] requires the following conditions:

$$\lambda_{\min}^n(p, q) \underset{n\to\infty}{\longrightarrow} \infty, \qquad \text{a.s.}$$

and

$$\lambda_{\max}^n(p, q) = O(\lambda_{\min}^n(p, q)(\log \lambda_{\min}^n(p, q))^\gamma), \qquad \text{a.s.}, \gamma > 0$$

for any $(p, q) \in M_0$, where $\lambda_{\min}^n(p, q)$ and $\lambda_{\min}^n(p, q)$ denote the minimum and maximum eigenvalue of $\sum_{i=0}^{n-1} \phi_i(p, 1, q)\phi_i^\tau(p, 1, q)$, respectively. Clearly, these conditions imply

$$\frac{(r_n^*)^{1/2}}{\lambda_{\min}(n)} \underset{n\to\infty}{\longrightarrow} 0 \text{ and } \frac{(\log r_n^*)(\log\log r_n^*)^c}{(r_n^*)^{1/2}} \underset{n\to\infty}{\longrightarrow} 0$$

since r_n^* and $\lambda_{\max}^n(p^*, q^*)$ are of the same order. Therefore, (3.12) is fulfilled with $s_n = (r_n^*)^{1/2}$ and the conclusions of Theorem 3.1 follow from Remark 3.2.

Remark 3.8: Comparing Theorem 3.3 of this paper to [9, Theorem 3.1] one finds a situation similar to that described in Remark 3.7: weaker conditions are required here and s_n can be taken as n^a, for any $a \in (0, 1)$.

Remark 3.9: In the case where we pay no attention to control performance and where only parameter estimation is required, we may take $u_n^0 \equiv 0$. Then H_2 is satisfied for stochastic systems if $A(z) \neq 0$ for $|z| < 1$. It is worth noting that we allow $A(z) = 0$ at $|z| = 1$. For deterministic systems, even this weaker than stability condition, is not required. Finally, the minimum-phase condition is unnecessary for either stochastic or deterministic systems.

IV. Adaptive Tracking

We now design the input for a deterministic or stochastic system with unknown orders, time-delay, and coefficients so that the system output follows a given bounded deterministic reference signal y_n^*. Specifically, we shall design u_n^0 in (3.15) so that Condition H_2 in Theorem 3.3 holds and the output $\{y_n\}$ of the stochastic system minimizes

$$\lim_{n\to\infty} \sup \frac{1}{n}\sum_{i=0}^{n}(y_i - y_i^*)^2$$

or so that (3.16) in Theorem 3.4 holds and $y_n - y_n^* \to_{n\to\infty} 0$ for the deterministic system (2.26). It is clear that consistence of parameter estimates does not necessarily imply asymptotic optimality of the adaptive closed-loop.

In this section we assume that v_n in (3.15) has independent components with continuous distributions, and that b_{d^*0} in the initial value θ_0^* [see (2.8)] is a nonzero constant. When the attenuating excitation control (3.15) is applied, we have [16]

$$b_{d^*n} \neq 0, \qquad \text{a.s., for any } n \geq 0. \tag{4.1}$$

Let $F(z) = 1 + f_1 z + \cdots + f_{d_0-1}z^{d_0-1}$ and $G(z) = g_0 + g_1 z + \cdots + g_{p_0-1}z^{p_0-1}$ be the solutions of the Diophantine equation

$$1 = F(z)A(z) + G(z)z^{d_0}. \tag{4.2}$$

Then the stochastic system (2.1) can be written as

$$y_{n+d_0} = F(z)B(z)z^{-d_0}u_n + G(z)y_n + F(z)w_{n+d_0} \tag{4.3}$$

and the deterministic system (2.26) as

$$y_{n+d_0} = F(z)B(z)z^{-d_0}u_n + G(z)y_n. \tag{4.4}$$

For the stochastic system, since $F(z)w_{n+d_0}$ and $F(z)B(z)z^{-d_0}u_n + G(z)y_n$ are uncorrelated, and the leading coefficient of $F(z)B(z)z^{-d_0}$ is b_{d_0} which is a nonzero constant, the optimal tracking control u_n should be defined from

$$y_{n+d_0}^* = F(z)B(z)z^{-d_0}u_n + G(z)y_n \tag{4.5}$$

when the system parameters are all known. Similarly, for deterministic systems, the optimal tracking control should also be given by (4.5). This motivates us to construct adaptive tracking control u_n in the following way.

Let u_n' be the solution of the following equation:

$$b_{d_n n}u_n' = y_{n+d_n}^* - (G_n(z)y_n + (F_nB_n)(z)z^{-d_n}u_n - b_{d_n n}u_n) \tag{4.6}$$

where

$$F_n(z) = 1 + f_{1n}z + \cdots + f_{d_n-1n}z^{d_n-1}$$

and

$$G_n(z) = g_{0n} + g_{1n}z + \cdots + g_{p_n-1n}z^{p_n-1}$$

are the solutions of the Diophantine equation

$$1 = F_n(z)A_n(z) + G_n(z)z^{d_n} \tag{4.7}$$

where

$$A_n(z) = 1 + a_{1n}z + \cdots + a_{p_n n}z^{p_n}, \tag{4.8}$$

$$B_n(z) = b_{d_n n}z^{d_n} + \cdots + b_{q_n n}z^{q_n} \tag{4.9}$$

and $(F_nB_n)(z)$ denotes the product of polynomials $F_n(z)$ and $B_n(z)$.

Clearly, u_n' is a good candidate for adaptive control. However, u_n' may grow too fast so that H_2 may not be satisfied and thus, the parameter estimates may be inconsistent. To overcome this difficulty we proceed, roughly speaking, as follows: we define the desired control u_n^0 equal to u_n' until a stopping time defined such that the growth rate conditions required in H_2 are satisfied. After this, we simply set $u_n^0 = 0$ until a stopping time so that the system output will be reduced to a certain extent with the help of the minimum-phase condition. After this, we again apply $u_n^0 = u_n'$. To be specific, we define the adaptive tracking control u_n of system (2.1) by (3.15) with u_n^0 defined as follows:

$$u_n^0 = \begin{cases} u_n', & \text{if } n \text{ belongs to some } [\tau_k, \sigma_k) \cap \Lambda, \\ 0, & \text{if } n \text{ belongs to some } [t\tau_k, \sigma_k) \cap \Lambda^c \text{ or } [\sigma_k, \tau_{k+1}) \end{cases} \tag{4.10}$$

where Λ is an integer set

$$\Lambda = [j: (u_j')^2 \le j^{1+\delta}] \tag{4.11}$$

and $\{\tau_k\}$ and $\{\sigma_k\}$ are two stopping time sequences defined by

$$1 = \tau_1 < \sigma_1 < \tau_2 < \sigma_2 < \cdots \tag{4.12a}$$

$$\sigma_k = \sup \left\{ \tau > \tau_k : \sum_{i=\tau_k}^{j-1} y_i^2 \le (j-1)^{1+\delta/2} + y_{\tau_k}^2, \ \forall\, j \in (\tau_k, \tau] \right\}. \tag{4.12b}$$

$$\tau_{k+1} = \inf \left\{ \tau > \sigma_k : \sum_{i=\sigma_k}^{\tau} y_i^2 \le \frac{\tau \log \tau}{2^k}, \sum_{i=\tau_k}^{\sigma_k - 1} y_i^2 \le \frac{\tau \log \tau}{2^k}, \right.$$
$$\left. \sum_{i=\sigma_k}^{\tau} u_i^2 \le \frac{\tau \log \tau}{2^k}, \sum_{i=\tau_k}^{\sigma_k-1} u_i^2 \le \frac{\tau \log \tau}{2^k} \right\} \tag{4.12c}$$

where δ is given by (3.16).

For the deterministic system (2.26) the adaptive tracking control u_n is given by (3.15) with u_n^0 defined by (4.10) but with Λ, $\{\tau_k\}$, and $\{\sigma_k\}$ replaced as follows:

$$\Lambda = \{j: (u_j')^2 \le j^{\delta/2}\}, \tag{4.13}$$

$$1 = \tau_1 < \sigma_1 < \tau_2 < \sigma_2 < \cdots \tag{4.14a}$$

$$\sigma_k = \sup \{\tau > \tau_k : y_{j-1}^2 \le (j-1)^{\delta/2} + y_{\tau_k}^2, \ \forall\, j \in (\tau_k, \tau]\} \tag{4.14b}$$

$$\tau_{k+1} = \inf \left\{ \tau > \sigma_k : \sum_{i=\tau-d^*-p^*+1}^{\tau} y_i^2 \le \log \tau, \right.$$
$$\left. \sum_{i=\tau-d^*-q^*+1}^{\tau} u_i^2 \le \log \tau \right\} \tag{4.14c}$$

where σ is given by (3.16).

By induction it is easy to see that u_n^0 is F_n-measurable.

Remark 4.1: In the definition of the desired control $\{u_n^0\}$ the solvability of (4.6) is essential. For this we now show

$$b_{d_n n} \ne 0 \qquad \text{a.s., for any } n \ge 1. \tag{4.15}$$

We first consider the case where d_n is generated by (2.27c). Suppose the converse were true, i.e., $b_{d_n n} = 0$. If $d_n < q_n$, then from (2.15) we see $\sigma_n(p_n, d_n, q_n) = \sigma_n(p_n, d_n + 1, q_n)$, and hence from (2.28b) $CIC_5(d_n+1)_n < CIC_5(d_n)_n$, which contradicts (2.27c). Therefore, $b_{d_n n} = 0 (n \ge 1)$ implies $d_n = q_n$, and hence

$$b_{q_n n} = 0. \tag{4.16}$$

We now show that (4.16) implies $q_n = d^*$. Otherwise, if $q_n > d^*$, then $\sigma_n(p_n, d^*, q_n - 1) = \sigma_n(p_n, d^*, q_n)$ by (2.15) and $CIC_4(p_n, q_n - 1)_n < CIC_4(p_n, q_n)_n$ by (2.28a). The last inequality contradicts (2.27b). Hence, (4.16) is impossible and (4.15) holds.

For the case where d_n is given by (2.21), (4.15) can be proved similarly.

For the stochastic system (2.1) we have the following.

Theorem 4.1: If Condition H_1 holds, $A(z)$ and $B(z)z^{-d_0}$ are stable, u_n is given by (3.15) and (4.6)–(4.12), θ_n^*, p_n, d_n, q_n are defined by (2.9)–(2.21), then

$$\|\theta_n^* - \theta\|^2 = O\left(\frac{(\log n)(\log\log n)^c}{n^{1-(t+1)\epsilon}}\right), \qquad \text{a.s.,} \tag{4.17}$$

$$(p_n, d_n, q_n) \underset{n\to\infty}{\longrightarrow} (p_0^*, d_0, q_0), \qquad \text{a.s.} \tag{4.18}$$

$$\lim_{n\to\infty} \sup \frac{1}{n}\sum_{i=0}^{n} (u_i)^2 < \infty \qquad \text{a.s.} \tag{4.19a}$$

and

$$\frac{1}{n}\sum_{i=0}^{n} (y_i - y_i^*)^2 = \frac{1}{n}\sum_{i=0}^{n} (F(z)w_i)^2 + O(n^{-\epsilon/2}). \tag{4.19b}$$

For deterministic system (2.26) we have the following.

Theorem 4.2: If $A(z)$ and $B(z)$ are coprime, $A(z)$ and $B(z)z^{-d_0}$ are stable, u_n is given by (3.15), (4.6)–(4.10), and (4.13), (4.14), then (2.27), (2.28) lead to

$$\|\theta_n(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0) - \theta(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0)\| = O(n^{-(1-(t+1)\epsilon)}), \tag{4.20}$$

$$(p_n, d_n, q_n) \underset{n\to\infty}{\longrightarrow} (p_0, d_0, q_0), \qquad \text{a.s.} \tag{4.21}$$

$$\lim_{n\to\infty} \sup |u_n| < \infty, \qquad \text{a.s.} \tag{4.22a}$$

and

$$|y_n - y_n^*| = O(n^{-\epsilon}), \qquad \text{a.s.} \tag{4.22b}$$

The proofs of Theorems 4.1 and 4.2 are given in Appendix B.

Remark 4.2: In both Theorems 4.1 and 4.2, we have assumed that $A(z)$ and $B(z)z^{-d_0}$ are stable, but the stability requirement for $A(z)$ can be removed by a treatment similar to that used in [16]. On the other hand, the stability assumption on $B(z)z^{-d_0}$ is unavoidable in a certain sense. To see this, we give a simple example for system (2.26):

$A(z) = 1$, $B(z) = z - 2z^2$, $y_n^* \equiv 1$. Clearly, $B(z)z^{-1}$ is unstable. Then the exact following $(y_n \equiv y_n^*)$ with $u_0 = 0$ leads to

$$u_n = 2u_{n-1} + 1 = \cdots = 2^n - 1$$

which grows without bound.

Remark 4.3: From [16] we know that for any u_n measurable with respect to $F_n = \sigma\{w_i, i \le n\}$ we have

$$\lim_{n\to\infty} \sup \frac{1}{n}\sum_{i=0}^{n} (y_i - y_i^*)^2 \ge \lim_{n\to\infty} \sup \frac{1}{n}\sum_{i=0}^{n} (F(z)w_i)^2.$$

So (4.17)–(4.19) mean that the adaptive control u_n defined by

(3.15) and (4.6)–(4.12) is optimal for stochastic systems, while for deterministic systems, the optimality of the adaptive control u_n given by (3.15), (4.6)–(4.10), and (4.13), (4.14) follows immediately from (4.20)–(4.22).

Remark 4.4: If the conditions of Theorem 4.1 are satisfied with

$$\lim_{n\to\infty} \sup \frac{1}{n}\sum_{i=0}^{n} w_i^2 = 0, \qquad \text{a.s.}$$

then the conclusions of Theorem 4.1 become (4.17), (4.18) and

$$\lim_{n\to\infty} \sup \frac{1}{n}\sum_{i=0}^{n} (y_i - y_i^*)^2 = 0.$$

Remark 4.5: Theorems 4.1 and 4.2 remain valid if, in lieu of stability of $A(z)$, we use a weaker condition: all zeros of $A(z)$ are outside the open unit disk and $G(z)$ is stable, where $G(z)$ is given by (4.2). To see that the latter condition is really weaker than stability of $A(z)$, it is enough to take $d_0 = 1$ and $A(z) = 1 - z$ as an example, for which $G(z) = 1$. A similar remark can be made also for deterministic systems.

Remark 4.6: For Theorem 4.1 we can show, in a way similar to that for [16, Lemma 4], that there exists a k (depending on sample) such that $\tau_k < \infty$ and $\sigma_k = \infty$. This means that after a finite number of steps the desired control u_n^0 is identical to u_n' defined from (4.6). For Theorem 4.2 the similar property will be proved in Appendix B.

V. Adaptive LQ Problem

In this section we shall consider an adaptive LQ problem for both the stochastic system (2.1) and the deterministic system (2.26). The loss function is

$$J(u) = \lim_{n\to\infty} \sup J_n(u) \tag{5.1}$$

where

$$J_n(u) = \frac{1}{n}\sum_{i=0}^{n-1} (Q_1 y_i^2 + Q_2 u_i^2), \qquad Q_1 \geq 0,\ Q_2 > 0 \tag{5.2}$$

for systems (2.1) and (2.26) with orders, time-delay, and coefficients all unknown.

In this section we assume that

H_3:

$$\frac{1}{n}\sum_{i=0}^{n} w_i^2 = R + O(n^{-\rho}), \qquad \text{a.s., for some } \rho > 0 \text{ and } R \geq 0. \tag{5.3}$$

Example 5.1: Suppose that $\{w_n, F_n\}$ is a martingale difference sequence with the following properties:

$$\sup_n E\{|w_{n+1}|^{2+\alpha} | F_n\} < \infty, \qquad \text{a.s.}$$

and

$$E\{w_{n+1}^2 | F_n\} = R, \qquad \text{a.s.}$$

where $\alpha > 0$ and $R \geq 0$ are deterministic constants, then Assumption H_3 holds for any $\rho \in (0, \alpha/(2+\alpha))$.

To see this we note that

$$\sum_{n=1}^{\infty} E\left\{\left|\frac{w_n^2 - E\{w_n^2 | F_{n-1}\}}{n^{1-\rho}}\right|^{1+\alpha/2} \bigg| F_n\right\}$$

$$\leq 2^{1+\alpha/2} \sup_n E\{|w_{n+1}|^{2+\alpha} | F_n\} \sum_{n=1}^{\infty} n^{-(1-\rho)(1+\alpha/2)}$$

$$< \infty, \qquad \text{a.s.}$$

and by the martingale convergence theorem [18] we get

$$\sum_{n=1}^{\infty} \frac{w_n^2 - R}{n^{1-\rho}} = \sum_{n=1}^{\infty} \frac{w_n^2 - E\{w_n^2 | F_{n-1}\}}{n^{1-\rho}} < \infty, \qquad \text{a.s.}$$

Then (5.3) follows by using the Kronecker lemma.

We first write (2.1) in the state-space form

$$x_{k+1} = Ax_k + Bu_k + Cw_{k+1}, \qquad x_0 = 0, \tag{5.4}$$

$$y_k = C^\tau x_k \tag{5.5}$$

and (2.26) in the form

$$x_{k+1} = Ax_k + Bu_k, \qquad x_0 = 0, \tag{5.6}$$

$$y_k = C^\tau x_k \tag{5.7}$$

with

$$A = \begin{bmatrix} -a_1 & 1 & 0 & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ -a_h & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{5.8a}$$

$$B^\tau = [0 \ \cdots \ 0 \ b_{d_0} \ \cdots \ b_h]_{1\times h}, \tag{5.8b}$$

$$C^\tau = [1 \ 0 \ \cdots \ 0]_{1\times h} \tag{5.8c}$$

where $h = \max(p_0, q_0, 1)$.

From [13] it is known that

$$\inf_{u\in U} J(u) = RC^\tau SC, \qquad \text{for system (2.1)}, \tag{5.9}$$

$$\inf_{u\in U} J(u) = 0, \qquad \text{for system (2.26)} \tag{5.10}$$

and the optimal control is

$$u_n = Lx_n \tag{5.11}$$

where

$$U = \left\{u\colon \sum_{i=0}^{n} u_i^2 = O(n),\ u_n^2 = o(n), \qquad \text{a.s. } u_n \in F_n\right\}, \tag{5.12}$$

$$L = -(B^\tau SB + Q_2)^{-1} B^\tau SA \tag{5.13}$$

S satisfies

$$S = A^\tau SA - A^\tau SB(B^\tau SB + Q_2)^{-1} B^\tau SA + CQ_1C^\tau \tag{5.14}$$

for which there is a unique positive definite solution S if (A, B, D) is controllable and observable for some D fulfilling $D^\tau D = CQ_1C^\tau$.

Based on the estimates p_n, d_n, q_n and $\theta_n(p_n, d_n, q_n)$ given by (2.27), (2.28) or (2.9)–(2.21), we estimate A, B, C, S, and x_n for a deterministic or a stochastic system by $A(n)$, $B(n)$,

$C(n)$, $S(n)$, and $\hat{x}_n$, respectively, as follows:

$$A(n) = \begin{bmatrix} -a_{1n} & 1 & 0 & \cdots & 0 \\ -a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ -a_{h_n n} & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad (5.15a)$$

$$B^\tau(n) = [0 \quad \cdots \quad 0 \quad b_{d_n n} \quad \cdots \quad b_{h_n n}]_{1\times h_n}, \quad (5.15b)$$

$$C^\tau(n) = [1 \quad 0 \quad \cdots \quad 0]_{1\times h_n}, \quad (5.15c)$$

$$h_n = \max(p_n, q_n, 1). \quad (5.16)$$

$$\begin{aligned} S(n) &= A^\tau(n)S'(n-1)A(n) \\ &\quad - A^\tau(n)S'(n-1)B(n)(B^\tau(n)S'(n-1)B(n)+Q_2)^{-1} \\ &\quad \times B^\tau(n)S'(n-1)A(n) + C(n)Q_1C^\tau(n). \quad (5.17) \end{aligned}$$

Here $S(0) = 0$, $S'(n-1)$ is a square matrix of dimension $h_n \times h_n$

$$S'(n-1) = \begin{cases} \begin{bmatrix} S(n-1) & 0 \\ 0 & 0 \end{bmatrix}, & \text{if } h_{n-1} < h_n, \\ M^\tau(n)S(n-1)M(n), & \text{if } h_{n-1} \ge h_n, \end{cases}$$

$$M^\tau(n) = [I \quad 0]_{h_n \times h_{n-1}}$$

and finally,

$$\begin{aligned} \hat{x}_{n+1} &= A(n)\hat{x}'_n + B(n)u_n + C(n)(y_{n+1} \\ &\quad - C^\tau(n)A(n)\hat{x}'_n - C^\tau(n)B(n)u_n), \quad (5.18) \end{aligned}$$

$$\hat{x}_0 = y_0 = 0$$

where $\hat{x}'_n$ is of dimension h_n and is defined by

$$\hat{x}'_n = \begin{cases} [\hat{x}_n^\tau \quad 0]^\tau, & \text{if } h_{n-1} < h_n, \\ M^\tau(n)\hat{x}_n, & \text{if } h_{n-1} \ge h_n. \end{cases} \quad (5.19)$$

We now have the estimate L_n for the optimal gain L given by (5.13)

$$L_n = -(B^\tau(n)S(n)B(n)+Q_2)^{-1}B^\tau(n)S(n)A(n). \quad (5.20)$$

However, we cannot directly take $L_n x_n$ as the desired control u_n^0 because $L_n x_n$ may grow too fast so that H_2 is not satisfied. Therefore, define

$$L_n^0 = \begin{cases} L_n, & \text{if } n \in [\tau_k, \sigma_k) \text{ for some } k, \\ 0, & \text{if } n \in [\sigma_k, \tau_{k+1}) \text{ for some } k, \end{cases} \quad (5.21)$$

$$u_n^0 = L_n^0 \hat{x}'_n \quad (5.22)$$

where stopping times $\{\tau_k\}$ and $\{\sigma_k\}$ are defined by

$$1 = \tau_1 < \sigma_1 < \tau_2 < \sigma_2 < \cdots,$$

$$\sigma_k = \sup\left\{\tau > \tau_k \colon \sum_{i=\tau_k}^{j-1}(L_i\hat{x}'_i)^2 \le (j-1)^{1+\delta} + (L_{\tau_k}\hat{x}'_{\tau_k})^2,\ \forall j \in (\tau_k, \tau]\right\}, \quad (5.23)$$

$$\tau_{k+1} = \inf\left\{\tau > \sigma_k \colon \sum_{i=\tau_k}^{\sigma_k - 1}(L_i\hat{x}'_i)^2 \le \frac{\tau^{1+\delta}}{2^k},\ \sum_{i=1}^{\tau}(\hat{x}'_i)^2 \le \tau^{1+\delta/2},\ (L_\tau\hat{x}'_\tau)^2 \le \tau^{1+\delta}\right\}. \quad (5.24)$$

For the stochastic system (2.1) and the deterministic system (2.26) we have the following two theorems, respectively.

Theorem 5.1: If H_1 and H_3 hold, $A(z)$ is stable, (A, B, D) is controllable and observable for some D satisfying $D^\tau D = CQ_1C^\tau$, θ_n^* and p_n, d_n, q_n are defined by (2.9)–(2.21), then u_n defined by (3.15) and (5.15)–(5.24) is optimal in the sense that

$$(p_n, d_n, q_n) \underset{n\to\infty}{\longrightarrow} (p_0, d_0, q_0), \quad \text{a.s.} \quad (5.25)$$

$$\|\theta_n^* - \theta^*\|^2 = O\left(\frac{(\log n)(\log\log n)^c}{n^{1-(t+1)\epsilon}}\right), \quad \text{a.s.} \quad (5.26)$$

and

$$J_n(u) = RC^\tau SC + O(n^{-\rho\wedge\epsilon}), \quad \text{a.s.} \quad (5.27)$$

Theorem 5.2: If $A(z)$ is stable, (A, B, D) is controllable and observable for some D satisfying $D^\tau D = CQ_1C^\tau$, p_n, d_n, q_n and $\theta_n(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0)$ are defined by (2.27), then u_n defined by (3.15) and (5.15)–(5.24) is optimal in the sense that

$$(p_n, d_n, q_n) \underset{n\to\infty}{\longrightarrow} (p_0, d_0, q_0), \quad \text{a.s.} \quad (5.28)$$

$$\|\theta_n(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0) - \theta(p_n \vee p_0, d_n \wedge d_0, q_n \vee q_0)\| = O(n^{-(1-(t+1)\epsilon)}) \quad (5.29)$$

and

$$J_n(u) = O(n^{-\epsilon}), \quad \text{a.s.} \quad (5.30)$$

Proof of Theorem 5.1: By an argument similar to the proof of Theorem 4.1 (see Appendix A), we have

$$\sum_{i=0}^{n}(L_i^0\hat{x}'_i)^2 = O(n^{1+\delta}). \quad (5.31)$$

From this and stability of $A(z)$ we have

$$\sum_{i=0}^{n} y_i^2 = O(n^{1+\delta}).$$

Then Theorem 3.3 asserts (3.18) and (3.19) by which we know that $(p_n, d_n, q_n) \equiv (p_0, d_0, q_0)$ for n starting from some $n_0 > 0$.

Hence, (5.27) can be shown in a way similar to that used in [13, Theorem 1].

Noticing that controllability of (A, B) implies coprimeness of $A(z)$ and $B(z)$, we can prove Theorem 5.2 similarly.

It is worth noting that under the conditions of Theorems 5.1 and 5.2 there is a k (depending on sample) such that $\tau_k < \infty$ and $\sigma_k = \infty$. This can be shown by a method similar to that of [13, Lemma 6].

Remark 5.1: From (3.15) and (5.15)–(5.24), it is easy to see that here only one computing procedure is needed for constructing the optimal linear quadratic adaptive control, whereas [9] required $(p^* + 1)q^*$ computing procedures.

VI. Conclusion

This paper gives recursive parameter estimates for systems (2.1) and (2.26) under the assumption that a lower bound of the time-delay and upper bounds of system orders are known. Optimal adaptive controls are designed for both tracking and LQ problems when the system coefficients, orders, and time-delay are all unknown, and the rate of convergence both of the estimates to their true values and of the loss functions to their minima are derived. We have simplified the estimation algorithms and essentially weakened the conditions used in [7]–[9]. The criteria used in the paper can be used for estimating time-delay, system orders, and coefficients for stochastic systems with correlated noise. This will be published elsewhere.

Appendix A

This section proves Theorems 3.1–3.4. We first present some properties of $CIC_1(p)_n$, $CIC_2(q)_n$, and $CIC_3(d)_n$.

Lemma A.1: Under the conditions of Theorem 3.1 we have

$$CIC_1(p)_n - CIC_1(p_0)_n \geq \begin{cases} s_n(p - p_0 + o(1)), \text{ a.s.,} & \text{if } p \geq p_0, \\ \lambda^*_{\min}(\hat{\alpha}_0 + o(1)), \text{ a.s.,} & \text{if } p < p_0; \end{cases} \quad (A.1)$$

$$CIC_2(q)_n - CIC_2(q_0)_n \geq \begin{cases} s_n(q - q_0 + o(1)), \text{ a.s.,} & \text{if } q \geq q_0, \\ \lambda^*_{\min}(\hat{\alpha}_0 + o(1)), \text{ a.s.,} & \text{if } q < q_0; \end{cases} \quad (A.2)$$

$$CIC_3(d)_n - CIC_3(d_0)_n \geq \begin{cases} s_n(d_0 - d + o(1)), \text{ a.s.,} & \text{if } d \leq d_0, \\ \lambda^*_{\min}(\hat{\alpha}_0 + o(1)), \text{ a.s.,} & \text{if } d > d_0 \end{cases} \quad (A.3)$$

where $\hat{\alpha}_0 > 0$ is a constant.

Proof: We first prove (A.1). For any $0 \leq p \leq p*$, set

$$H(p) = \begin{bmatrix} I_p & 0_1 & 0 \\ 0 & 0 & I_{q^*} \end{bmatrix} \quad (A.4)$$

where I_p and I_{q^*} are identity matrices of dimension p and $q^* - d^* + 1$, respectively, while 0_1 is a zero matrix of dimension $p \times (p^* - 1)$.

If $p \geq p_0$, then

$$y_{n+1} = \theta^\tau(p, d^*, q^*)\phi_n(p, d^*, q^*) + w_{n+1} \quad (A.5)$$

and

$$\begin{aligned} \sigma_n(p, d^*, q^*) &= \sum_{i=0}^{n-1}(\tilde{\theta}_n^\tau(p, d^*, q^*)\phi_i(p, d^*, q^*) + w_{i+1})^2 \\ &= \tilde{\theta}_n^\tau(p, d^*, q^*)\sum_{i=0}^{n-1}\phi_i(p, d^*, q^*) \\ &\quad \cdot \phi_i^\tau(p, d^*, q^*)\tilde{\theta}_n(p, d^*, q^*) \\ &\quad + 2\tilde{\theta}_n^\tau(p, d^*, q^*)\sum_{i=0}^{n-1}\phi_i(p, d^*, q^*)w_{i+1} \\ &\quad + \sum_{i=0}^{n-1} w_{i+1}^2 \end{aligned} \quad (A.6)$$

where $\tilde{\theta}_n(p, d, q) \triangleq \theta(p, d, q) - \theta_n(p, d, q)$.

Noticing (A.4) and (2.9), we have

$$\tilde{\theta}_n^* \triangleq \theta^* - \theta_n^* = -\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1}\left(\sum_{i=0}^{n-1}\phi_i^* w_{i+1} - \theta^*\right) \quad (A.7)$$

$$\tilde{\theta}_n(p, d^*, q^*) = H(p)\tilde{\theta}_n^*, \ \phi_n(p, d^*, q^*) = H(p)\phi_n^*$$

and

$$\begin{aligned} &\sigma_n(p, d^*, q^*) \\ &= \left(\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2}\left(\sum_{i=0}^{n-1}\phi_i^* w_{i+1} - \theta^*\right)\right)^\tau \\ &\quad \cdot \left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2} H^\tau(p)H(p) \\ &\quad \cdot \left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right) H^\tau(p)H(p)\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2} \\ &\quad \times \left(\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2}\left(\sum_{i=0}^{n-1}\phi_i^* w_{i+1} - \theta^*\right)\right) \\ &\quad - 2\left(\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2}\left(\sum_{i=0}^{n-1}\phi_i^* w_{i+1} - \theta^*\right)\right)^\tau \\ &\quad \cdot \left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2} H^\tau(p)H(p)\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{1/2} \\ &\quad \cdot \left(\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2}\sum_{i=0}^{n-1}\phi_i^* w_{i+1}\right) + \sum_{i=0}^{n-1} w_{i+1}^2. \end{aligned} \quad (A.8)$$

Let $T(p)$ be an orthogonal matrix such that

$$T(p)H^\tau(p)H(p)T^\tau(p) = \begin{bmatrix} I_p & 0 & 0 \\ 0 & I_{q^*} & 0 \\ 0 & 0 & 0 \end{bmatrix} \triangleq F(p).$$

We have

$$\begin{aligned} &\left\| \left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2} H^\tau(p)H(p)\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right) H^\tau(p)H(p) \right\| \\ &\quad \cdot \left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2} \\ &\leq \operatorname{tr}\left(\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau} + I\right)^{-1/2} H^\tau(p)H(p)\right. \end{aligned}$$

$$\cdot \left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right) H^\tau(p)H(p)\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}+I\right)^{-1/2}\Bigg)$$

$$= \operatorname{tr}(T(p)H^\tau(p)H(p)T^\tau(p)T(p) \cdot \left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}+I\right)^{-1} T^\tau(p)T(p)H^\tau(p)H(p) \cdot T^\tau(p)T(p)\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right) \cdot T^\tau(p)T(p)H^\tau(p)H(p)T^\tau(p))$$

$$= \operatorname{tr}\left(F(p)\left(T(p)\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right) \cdot T^\tau(p)+I\right)^{-1} F(p)T(p)\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right) \cdot T^\tau(p)F(p)\right) = O(1) \tag{A.9}$$

and

$$\left\|\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}+I\right)^{-1/2} H^\tau(p)H(p) \cdot \left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}+I\right)^{1/2}\right\|^2 = O(1). \tag{A.10}$$

By [8, Lemma 2] we also have

$$\left\|\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}+I\right)^{-1/2}\sum_{i=0}^{n-1}\phi_i^* w_{i+1}\right\|^2 = O((\log r_n^*)(\log\log r_n^*)^c). \tag{A.11}$$

Combining (A.8)–(A.11) yields

$$\sigma_n(p, d^*, q^*) = O((\log r_n^*)(\log\log r_n^*)^c) + \sum_{i=0}^{n-1} w_{i+1}^2. \tag{A.12}$$

From this and (3.5) we obtain the first part of (A.1).

$$\tilde{\theta}_n^*(p) = [a'_{1n} - a_1 \cdots a'_{p^*n} - a_{p^*} b_{d^*} - b_{d^*n} \cdots b_{q^*} - b_{q^*n}]^\tau \tag{A.13}$$

where $a'_{in} = a_{in}$ if $i \le p$, $a'_{in} = 0$ if $p < i \le p^*$.

When $p < p_0$, we have

$$\|\tilde{\theta}_n^*(p)\|^2 \ge \min\{a_{p_0}^2, b_{d_0}^2, b_{q_0}^2\} \triangleq \hat{\alpha}_0 > 0 \tag{A.14}$$

and hence by [8, Lemma 2]

$$\sigma_n(p, d^*, q^*) = \tilde{\theta}_n^{*\tau}(p)\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right)\tilde{\theta}_n^*(p) + 2\tilde{\theta}_n^{*\tau}(p)\sum_{i=0}^{n-1}\phi_i^* w_{i+1} + \sum_{i=0}^{n-1} w_{i+1}^2$$

$$\ge \left\|\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right)^{1/2}\tilde{\theta}_n^*(p)\right\|^2 + \sum_{i=0}^{n-1} w_{i+1}^2 - O\left(\left\|\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right)^{1/2}\tilde{\theta}_n^*(p)\right\| \cdot ((\log r_n^*)(\log\log r_n^*)^c)^{1/2}\right)$$

$$\ge \|\tilde{\theta}_n^*(p)\|^2 \lambda_{\min}^*(n) \cdot \left(1 - O\left(\frac{((\log r_n^*)(\log\log r_n^*)^c)^{1/2}}{\left\|\left(\sum_{i=0}^{n-1}\phi_i^*\phi_i^{*\tau}\right)^{1/2}\tilde{\theta}_n^*(p)\right\|}\right)\right) + \sum_{i=0}^{n-1} w_{i+1}^2$$

$$\ge \hat{\alpha}_0 \lambda_{\min}^*(n) \cdot \left(1 - O\left(\left(\frac{(\log r_n^*)(\log\log r_n^*)^c}{\lambda_{\min}^*(n)}\right)^{1/2}\right)\right) + \sum_{i=0}^{n-1} w_{i+1}^2$$

$$= \lambda_{\min}^*(n)(\hat{\alpha}_0 + o(1)) + \sum_{i=0}^{n-1} w_{i+1}^2$$

which together with (A.12) implies the second part of (A.1), while (A.2) and (A.3) can be obtained similarly.

Proof of Theorem 3.1: Noticing (A.7) we immediately conclude (3.7) by using [8, Lemma 2].

Since (p_n, d_n, q_n) belongs to a finite set, for (3.8) we need only show that any limit point of the sequence $\{(p_n, d_n, q_n)\}$ is precisely (p_0, d_0, q_0). To this end, let p' be the limit of a subsequence $\{p_{n_k}\}$ of $\{p_n\}$.

If $p' < p_0$, then (A.1) tells us that for all sufficiently large k.

$$0 \ge CIC_1(p_{n_k})_{n_k} - CIC_1(p_0)_{n_k} = CIC_1(p')_{n_k} - CIC_1(p_0)_{n_k} \ge \lambda_{\min}^*(n_k)(\hat{\alpha}_0 + o(1)) \xrightarrow[k\to\infty]{} \infty.$$

The obtained contradiction means that $p' \geq p_0$. Again, by (A.1) we know

$$0 \geq CIC_1(p')_{n_k} - CIC_1(p_0)_{n_k} \geq s_{n_k}(p' - p_0 + o(1)), \quad \text{for } p' \geq p_0.$$

Thus, we must have $p' = p_0$, otherwise the last inequality leads to a contradiction as $k \to \infty$. Since p' is any limit point of p_n, we conclude that $p_n \to p_0$, a.s. $n \to \infty$.

Similarly, by (A.2) and (A.3) it is not difficult to assert $q_n \to q_0$, a.s. and $d_n \to d_0$, a.s. as $n \to \infty$.

In order to prove Theorem 3.2, we need the following lemma.

Lemma A.2: Under the conditions of Theorem 3.2 we have

$$CIC_1(p)_n - CIC_1(p_0)_n \geq \begin{cases} s_n(p - p_0 + o(1)), \text{ a.s.} & \text{if } p \geq p_0, \\ \lambda_{\min}(\hat{\alpha}_0 + o(1)), \text{ a.s.} & \text{if } p < p_0; \end{cases} \quad \text{(A.15)}$$

$$CIC_2(q)_n - CIC_2(q_0)_n \geq \begin{cases} s_n(q - q_0 + o(1)), \text{ a.s.} & \text{if } q \geq q_0, \\ \lambda_{\min}(\hat{\alpha}_0 + o(1)), \text{ a.s.} & \text{if } q < q_0 \end{cases} \quad \text{(A.16)}$$

where $\lambda_{\min}(n)$ is given by Theorem 3.2 and $\hat{\alpha}_0$ is defined by (A.14).

The proof of this lemma is similar to that of Lemma A.1.

Proof of Theorem 3.2: Similar to the proof procedure of Theorem 3.1, we have $p_n \to p_0$ and $q_n \to q_0$, a.s. as $n \to \infty$. Noticing that M_0 which contains $\{(p_n, q_n)\}$ is a finite set, we know that

$$(p_n, q_n) \equiv (p_0, q_0) \quad \text{(A.17)}$$

for all n starting from a large enough number $n_0 > 0$.

By using (A.17) and the method used in the proof of Lemma A.1, we have

$$CIC_5(d)_n - CIC_5(d_0)_n \geq \begin{cases} s_n(d_0 - d + o(1)), \text{ a.s.,} & \text{if } d \leq d_0, \\ \lambda_{\min}(\hat{\alpha}_0 + o(1)), \text{ a.s.,} & \text{if } d > d_0 \end{cases} \quad \text{(A.18)}$$

and hence we can prove $d_n \to d_0$, a.s. as $n \to \infty$. Therefore, (3.10) holds.

Proof of Theorem 3.3: Obviously, from Theorem 3.1 we need only show that

$$\lim_{n\to\infty} \inf n^{-1+(t+1)(\epsilon+\delta)} \lambda^*_{\min}(n) \neq 0, \text{ a.s.} \quad \text{(A.19)}$$

By the argument used in [11], for this it suffices to show that there does not exist a nonzero vector

$$\eta = [\alpha_0 \cdots \alpha_{p^*-1} \quad \beta_0 \cdots \beta_{q^*-1}]^\tau$$

such that

$$\sum_{i=0}^{p^*-1} \alpha_i z^i B(z) = \sum_{i=0}^{q^*-1} \beta_i z^i A(z) \text{ and } \sum_{i=0}^{p^*-1} \alpha_i z^i = 0. \quad \text{(A.20)}$$

This is true indeed, because the second equation of (A.20) leads to $\alpha_i = 0$ $(i = 0, \cdots, p^* - 1)$, then by the first equation of (A.20) $\beta_j = 0$ $(j = 0, \cdots, q^* - 1)$.

Proof of Theorem 3.4: Similar to Theorem 3.3, here we need only show that

$$\lim_{n\to\infty} \inf n^{-1+(t+1)(\epsilon+\delta)} \lambda_{\min}(n) \neq 0, \quad \text{a.s.} \quad \text{(A.21)}$$

where $\lambda_{\min}(n)$ is given by Theorem 3.2.

If (A.21) were not true, then it would be either

$$\lim_{n\to\infty} \inf n^{-1+(t+1)(\epsilon+\delta)} \lambda_{\min}^{(p_0, q^*)}(n) \neq 0, \quad \text{a.s.} \quad \text{(A.22)}$$

or

$$\lim_{n\to\infty} \inf n^{-1+(t+1)(\epsilon+\delta)} \lambda_{\min}^{(p^*, q_0)}(n) \neq 0, \quad \text{a.s.} \quad \text{(A.23)}$$

If (A.22) were true, then by the same argument as that used in [14] we can prove that there would exist a $(p_0 + q^* - d^*)$-dimensional vector

$$\eta = [\alpha'_0 \cdots \alpha'_{p_0-1} \quad \beta'_{d^*} \cdots \beta'_{q^*-1}]^\tau, \qquad \|\eta'\| = 1$$

satisfying

$$0 = \sum_{i=0}^{p_0-1} \alpha'_i z^i B(z) = \sum_{i=d^*}^{q^*-1} \beta'_i z^i A(z). \quad \text{(A.24)}$$

Since $A(z)$ and $B(z)$ are coprime, there are two polynomials $M(z)$ and $N(z)$ such that

$$A(z)M(z) + B(z)N(z) = 1. \quad \text{(A.25)}$$

Then from (A.24) we have

$$\sum_{i=0}^{p_0-1} \alpha'_i z^i = \sum_{i=0}^{p_0-1} \alpha'_i (A(z)M(z) + B(z)N(z)) z^i = A(z) \left(\sum_{i=0}^{p_0-1} \alpha'_i z^i M(z) - \sum_{i=d^*}^{q^*-1} \beta'_i z^i N(z) \right) \quad \text{(A.26)}$$

which implies that

$$\sum_{i=0}^{p_0-1} \alpha'_i z^i = 0. \quad \text{(A.27)}$$

Since $\deg(A(z)) = p_0$ and $\deg(\sum_{i=0}^{p_0-1} \alpha'_i z^i) \leq p_0 - 1$ (A.27) and (A.24) yield $\eta' = 0$, which contradicts $\|\eta'\| = 1$. Therefore, (A.22) is not true.

Similarly, we can conclude that (A.23) is also not true. Therefore, (A.21) and, hence, Theorem 3.4 holds.

Appendix B

In this section we prove Theorems 4.1 and 4.2.

Proof of Theorem 4.1: Theorem 3.3 implies (4.17) and (4.18) if we can verify H_2, for which it suffices to show (3.16) because of the stability of $A(z)$. This can be done by a method similar to that used in [13]. Therefore, (4.17) and (4.18) hold. From (4.18), we know that $(p_n, d_n, q_n) \equiv (p_0, d_0, q_0)$, starting from a sufficiently large n_0. Hence, (4.19) can be shown in a way similar to that for [16, Theorem 1].

Proof of Theorem 4.2: From (4.10) and (4.13) it is easy to see that

$$u_n^2 = O(n^{\delta/2}) \quad \text{(B.1)}$$

which together with the stability of $A(z)$ implies

$$y_n^2 = O(n^{\delta/2}). \quad \text{(B.2)}$$

Thus, by Theorem 3.4 we know that (4.21) and (3.21) hold.

Notice that (4.21) implies that $(p_n, d_n, q_n) \equiv (p_0, d_0, q_0)$ starting from a large enough n_0. Hence, without loss of generality we may assume $(p_n, d_n, q_n) \equiv (p_0, d_0, q_0)$ for all $n \geq 1$.

We now show that there exists a k (depending on sample) such that

$$\tau_k < \infty \text{ and } \sigma_k = \infty. \quad \text{(B.3)}$$

Obviously, we need only show the impossibility of the following two situations:

i) $\sigma_k < \infty$ and $\tau_{k+1} = \infty$, for some positive integer k;

ii) $\tau_k < \infty$ and $\sigma_k < \infty$, for every positive integer k.

We now first show the impossibility of i). If i) were true, then from (3.15), (4.10), and the stability of $A(z)$ we have $y_n \to_{n\to\infty} 0$ and $u_n \to_{n\to\infty} 0$, which contradicts $\tau_{k+1} = \infty$ [see (4.14c)].

We now prove the impossibility of ii).

Set

$$t_k = \sup\{n\colon\ j \in [\tau_k, \sigma_k) \cap \Lambda,\ \forall\, j \in [\tau_k, n]\} \quad \text{(B.4)}$$

where Λ is given by (4.13).

By (4.14c) and (4.6) we know

$$u'_{\tau_k} = O(\log \tau_k)$$

which implies that t_k exists for all sufficiently large k.

For any $n \in [\tau_k, t_k]$, it follows from (4.4), (4.6), and (4.10) that

$$\begin{aligned} y_{n+d_0} = y^*_{n+d_0} &+ (F(z)B(z)z^{-d_0} - (F_nB_n)z^{-d_0})u_n \\ &+ (G(z) - G_n(z))y_n + b_{d_0n}v_n. \end{aligned} \quad \text{(B.5)}$$

Notice that $1-(t+1)(\epsilon+\delta)-\delta > \frac{1}{2} > \epsilon$. Equations (B.5), (B.1), (B.2), (3.21), and (3.14) lead to

$$y_{n+d_0} = O(1), \qquad \text{for } n \in [\tau_k, t_k]. \quad \text{(B.6)}$$

From (4.14c) and (4.4) it follows that

$$y_n = O(\log \tau_k), \qquad \text{for } n \in [\tau_k,\ \tau_k + d_0 - 1]. \quad \text{(B.7)}$$

By induction (B.6), (B.7), (4.14c), and (4.6) lead to

$$u'_n = O(\log \tau_k), \qquad \text{for } n \in [\tau_k,\ t_k + 1]. \quad \text{(B.8)}$$

Combining this with (B.4), we have

$$t_k = \sigma_k - 1. \quad \text{(B.9)}$$

By this and (B.6), (B.7) we have

$$y_{\sigma_k} = O(\log \tau_k) = O(\log \sigma_k)$$

which contradicts (4.14b). Thus, (B.3) is proved.

Equation (4.22) follows from (B.3), (4.10), (4.4), and (4.6), while (4.20) comes from Theorem 3.4 with $\delta = 0$ (see Remark 3.5).

The proof is completed.

Acknowledgment

The authors are grateful to Prof. M. Gevers and the referees for their helpful comments and suggestions.

References

[1] G. E. Box and G. M. J. Jenkins, *Time-Series Analysis, Forecasting and Control.* San Francisco: Holden Day, 1970.

[2] H. Akaike, "Fitting autoregressive models for prediction," *Ann. Inst. Statist. Math.*, vol. 21, pp. 243-247, 1969.

[3] —, "On entropy maximization principle," in *Applications of Statistics*, P. R. Krishaiah, Ed. Amsterdam, The Netherlands: North-Holland, 1977, pp. 27-41.

[4] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 467-471, 1978.

[5] E. J. Hannan and B. C. Quinn, "The determination of the order of an autoregression," *J. Roy. Statist. Soc.*, vol. B.41, pp. 190-195, 1979.

[6] J. Rissanen, "Order estimation by accumulated prediction errors," *J. Appl. Probab.*, Essays in Time Series and Allied Process, J. Gani and M. B. Priestley, Eds. Special vol. 23A, pp. 55-61, 1986.

[7] H. F. Chen and L. Guo, "Consistent estimation of the order of stochastic control systems," *IEEE Trans. Automat. Contr.*, vol. AC-32, pp. 531-535, 1987.

[8] L. Guo, H. F. Chen, and J. F. Zhang, "Consistent order estimation for linear stochastic feedback control systems (CARMA model)," *Automatica*, vol. 25, pp. 147-151, 1989.

[9] E. M. Hemerly and M. H. A. Davis, "Recursive order estimation of stochastic control systems," *Math. Contr., Signals Syst.*, 1989.

[10] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[11] H. F. Chen and L. Guo, "Convergence rate of least-squares identification and adaptive control for stochastic systems," *Int. J. Contr.*, vol. 44, pp. 1459-1476, 1986.

[12] U. Shaked and P. R. Kumar, "Minimum variance control of discrete time multivariable ARMAX systems," *SIAM J. Contr. Optimiz.*, vol. 24, pp. 396-411, 1986.

[13] H. F. Chen and L. Guo, "Optimal adaptive control and consistent parameter estimates for ARMAX model with quadratic cost," *SIAM J. Contr. Optimiz.*, vol. 25, pp. 845-867, 1987.

[14] —, "Adaptive control via consistent estimation for deterministic systems," *Int. J. Contr.*, vol. 45, pp. 2183-2202, 1987.

[15] E. J. Hannan and J. Rissanen, "Recursive estimation of ARMA order," *Biometrika*, vol. 69, pp. 81-97, 1982.

[16] H. F. Chen and J. F. Zhang, "Convergence rates in stochastic adaptive tracking," *Int. J. Contr.*, vol. 49, pp. 1915-1935, 1989.

[17] R. H. Middleton, G. C. Goodwin, D. J. Hill, and D. Q. Mayne, "Design issues in adaptive control," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 50-58, 1988.

[18] Y. S. Chow, "Local convergence of martingales and the law of large numbers," *Ann. Math. Statist.*, vol. 36, pp. 552-558, 1965.

Han-Fu Chen received the diploma in mathematics from the University of Leningrad, Leningrad, U.S.S.R., in 1961.

From 1961 to 1980 he was with the Institute of Mathematics, Academia Sinica, Beijing, China. Since 1980 he has been with the Institute of Systems Science, Academia Sinica, where he is currently a Professor in the Laboratory of Control Theory and Applications. He has authored and/or coauthored over 90 papers and four books in the areas of stochastic control, adaptive control, identification, and stochastic approximation.

Prof. Chen serves as Vice-Chairman of the International Federation of Automatic Control Theory Committee. He is the Editor of *Systems Science and Mathematical Sciences* and is a member of the Editorial Boards of *Stochastics and Stochastics Reports*, *Systems and Control Letters*, *Adaptive Control and Signal Processing*, *Scientia Sinica*, *Acta Mathematicae Applicatae Sinica*, and *Acta Mathematica Sinica*.

Ji-Feng Zhang was born in 1961 in Shandong Province, China. He graduated from the Department of Mathematics, Shandong University, Shandong, China, in 1985, and received the M.S. degree in control theory from the Institute of Systems Science, Academia Sinica, Beijing, China, in 1988.

At present, he is a Ph.D. candidate at Academia Sinica. His research interests are in stochastic systems and singular systems. He has published papers in *Automatica*, the *International Journal of Control*, and in journals printed in China.

Mr. Zhang received the Award of the President of the Chinese Academy of Sciences in 1989.