Systems Science

Vol. 16 1990 No. 4

HAN-FU CHEN*, JI-FENG ZHANG*

CONVERGENCE ANALYSIS FOR ADAPTIVE CONTROL SYSTEMS WITH UNKNOWN ORDERS, DELAY AND COEFFICIENTS**

This paper considers identification and adaptive control for linear discrete-time stochastic systems with unknown orders, time-delay and coefficients under uncorrelated noise. Under the assumption that a lower bound for the time-delay and upper bounds for system orders are known, 1) the consistent estimates for the time-delay, system orders and coefficients are recursively given; 2) the optimal adaptive controls are designed for both tracking and the quadratic loss function and 3) the rates of convergence both of the coefficient estimates to their true values and of the loss functions to their minimums are derived.

1. INTRODUCTION

Let the *a priori* information about the plant be merely that it is linear stochastic and 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 an extensive literature devoted to estimating orders and coefficients of a stationary ARMA process from a non-recursive point of view, see BOX and JENKINS [1], AKAIKE [2], [3], RISSANEN [4] and HANNAN and QUINN [5]. Recently, however, RISSANEN [6] established results concerning the 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 the ARMAX process when the exogenous input is a feedback control so that the process is neither ergodic nor stationary.

---

* Institute of Systems Science, Academia Sinica, Beijing 100080, P.R. of China.

** Project supported by the National Natural Science Foundation of China.

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]. In their work, for systems with uncorrelated noise, by combining the PLS (Predictive Least Squares) criterion for order estimation with an adaptive control strategy minimizing a quadratic cost and applied to multidimensional ARX systems, it is shown that the combination enables us to estimate, recursively and in a strong 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 construction of an optimal quadratic adaptive control.

This paper is devoted to reducing the computing quantity and the assumptions required in [7]–[9]. Parameters we want to estimate are not only system orders but also the system time-delay which is not estimated in previous works. The knowledge about time-delay is unnecessary in some cases where adaptive tracking [10] or adaptive control with quadratic cost [11] are dealt with without paying attention to parameter estimation, but it is crucial for some control problems, for example, the minimum variance control is sensitive to time-delay [12]. The recursion is also given for criteria 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 with those in [7]–[9]. As main results of the paper, for stochastic systems of possible non-minimum-phase with unknown orders, time-delay and coefficients, optimal adaptive controls are derived for tracking and quadratic index, 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 and single-output systems only. The corresponding results for multidimensional systems can be obtained similarly. The arrangement of this paper is as follows. Section 2 presents methods and criteria for estimating system orders, time-delay and coefficients. Section 3 discusses sufficient conditions guaranteeing consistency of the estimates. Section 4 designs an optimal adaptive tracking control which makes the estimated parameters strongly consistent, while Section 5 gives an optimal quadratic adaptive control which guarantees the strong consistency of the estimated parameters and the asymptotical minimality of the loss function. The convergence rates both of the coefficient estimates to their true values and of the loss functions to their minimums are also derived in Sections 4 and 5. Finally, we conclude this paper in Section 6.

2. ESTIMATION METHODS FOR TIME-DELAY, ORDERS AND COEFFICIENTS

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

Let the single-input and single-output systems be described by a linear stochastic difference equation

$$A(z)\, y_n = B(z)\, u_n + w_n,\ n > 0; \qquad y_n = u_n = w_n = 0,\ n \leqslant 0, \tag{2.1}$$

where y_n, u_n and w_n are the output, input and noise, respectively; $A(z)$ and $B(z)$ are polynomials in shift-back operator z:

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

$$B(z) = b_{d_0} z^{d_0} + \ldots + b_{q_0} z^{q_0},\ q_0 \geqslant d_0 \geqslant 1. \tag{2.3}$$

The coefficients a_i $(i = 1, \ldots, p_0)$, b_j $(j = d_0, \ldots, q_0)$, 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^*, q^* and $q^* \geqslant d^* \geqslant 1$ are given such that

$$(p_0, q_0) \in M_0 \triangleq \{(p, q)\colon\ 0 \leqslant p \leqslant p^*,\ \ d^* \leqslant q \leqslant q^*\}, \tag{2.4}$$

$$d_0 \in M_d \triangleq \{d\colon\ d^* \leqslant d \leqslant q^*\}. \tag{2.5}$$

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

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

$$\varphi_n^* = [y_n \ldots y_{n-p^*+1}\ u_{n-d^*+1} \ldots u_{n-q^*+1}]^\tau \tag{2.6}$$

and denote unknown coefficients by

$$\theta^* = [-a_1 \ldots -a_{p^*}\ b_{d^*} \ldots b_{q^*}]^\tau, \tag{2.7}$$

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

$$\theta_n^* = [-a_{1n} \ldots -a_{p^*n}\, b_{d^*n} \ldots b_{q^*n}]^\tau \tag{2.8}$$

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

$$\theta_n^* = \Big(\sum_{i=0}^{n-1} \varphi_i^* \varphi_i^{*\tau} + I\Big)^{-1} \sum_{i=0}^{n-1} \varphi_i^* y_{i+1} \tag{2.9}$$

or recursively given by:

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

$$P_{n+1}^* = P_n^* - b_n^* P_n^* \varphi_n^* \varphi_n^{*\tau} P_n^*, \tag{2.11}$$

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

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

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

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

and

$$\sigma_n(p, d, q) = \sum_{i=0}^{n-1} (y_{i+1} + a_{1n} y_i + \ldots + a_{pn} y_{i-p+1} - b_{dn} u_{i-d+1} - \ldots - b_{qn} u_{i-q+1})^2$$

$$= \sum_{i=0}^{n-1} (y_{i+1} - \theta_n^\tau(p, d, q)\varphi_i(p, d, q))^2. \tag{2.15}$$

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

Introduce the criteria

$$\mathrm{CIC}_1(p)_n = \sigma_n(p, d^*, q^*) + p s_n, \tag{2.16}$$

$$\mathrm{CIC}_2(q)_n = \sigma_n(p^*, d^*, q) + q s_n \tag{2.17}$$

and

$$\mathrm{CIC}_3(d)_n = \sigma_n(p^*, d, q^*) - d s_n, \tag{2.18}$$

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

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

$$p_n = \arg\min_{0 \leqslant p \leqslant p^*} \mathrm{CIC}_1(p)_n, \tag{2.19}$$

$$q_n = \arg\min_{d^* \leqslant q \leqslant q^*} \mathrm{CIC}_2(q)_n \tag{2.20}$$

and

$$d_n = \arg\min_{d^* \leqslant d \leqslant q^*} \mathrm{CIC}_3(d)_n. \tag{2.21}$$

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

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

where

$$N_{n+1}(p, d, q) = N_n(p, d, q) + \varphi_n(p, d, q)\varphi_n^\tau(p, d, q),$$

$$N_0(p, d, q) = 0,$$

$$F_{n+1}(p, d, q) = F_n(p, d, q) + \varphi_n(p, d, q)y_{n+1}, \; F_0(p, d, q) = 0,$$

we can compute $\mathrm{CIC}_1(p)_n$, $\mathrm{CIC}_2(q)_n$ and $\mathrm{CIC}_3(d)_n$ also in a recursive way:

$$\mathrm{CIC}_1(p)_{n+1} = \mathrm{CIC}_1(p)_n + p(s_{n+1} - s_n) + G(p, d^*, q^*)_n, \tag{2.22}$$

$$\mathrm{CIC}_2(q)_{n+1} = \mathrm{CIC}_2(q)_n + q(s_{n+1} - s_n) + G(p^*, d^*, q)_n \tag{2.23}$$

and

$$\mathrm{CIC}_3(d)_{n+1} = \mathrm{CIC}_3(d)_n - d(s_{n+1} - s_n) + G(p^*, d, q^*)_n, \tag{2.24}$$

where

$$\begin{aligned} G(q, d, q)_n = {} & (y_{n+1} - \theta_n^\tau(p, d, q)\varphi_n(p, d, q))^2 \\ & + (\theta_{n+1}(p, d, q) - \theta_n(p, d, q))^\tau (N_{n+1}(p, d, q)\theta_{n+1}(p, d, q) \\ & + N_{n+1}(p, d, q)\theta_n(p, d, q) - 2F_{n+1}(p, d, q)). \end{aligned} \tag{2.25}$$

Remark 2.1

It is worth noticing that in the above order or delay estimation procedure, $\mathrm{CIC}_1(p)_n$, $\mathrm{CIC}_2(q)_n$ and $\mathrm{CIC}_3(d)_n$ 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 $p^* + q^* - d^* + 2$ 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 non-recursive, while here computing $\mathrm{CIC}_1(p)_n$, $\mathrm{CIC}_2(q)_n$ and $\mathrm{CIC}_3(d)_n$ is carried out recursively.

3. 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.

We assume that

H_1. $\{w_n, \mathscr{F}_n\}$ is a martingale difference sequence with properties

$$\sup_n E(w_{n+1}^2 | \mathscr{F}_n) < \infty, \text{ a.s.}, \tag{3.1}$$

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

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

where $\{\mathscr{F}_n\}$ is a family of non-decreasing σ-algebras, and

$$\varepsilon^* = \frac{1}{2(t+1)}, \quad t = 2p^* + q^*. \tag{3.4}$$

THEOREM 3.1

If H_1 holds, u_n is $\mathscr{F}_n$-measurable, and $r_n^* = 1 + \sum_{i=0}^{n} \|\varphi_i^*\|^2$ satisfies

$$\frac{(\log r_n^*)(\log\log r_n^*)^c}{(\log n)^2} \xrightarrow[n\to\infty]{} 0, \textit{ for some constant } c > 1 \tag{3.5}$$

and

$$\frac{(\log n)^2}{\lambda_{\min}^*(n)} \xrightarrow[n\to\infty]{} 0, \tag{3.6}$$

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

$$\|\theta_n^* - \theta^*\|^2 = o\left(\frac{(\log r_n^*)(\log\log r_n^*)^c}{\lambda_{\min}^*(n)}\right) \xrightarrow[n\to\infty]{} 0, \text{ a.s.,} \tag{3.7}$$

$$(p_n, d_n, q_n) \xrightarrow[n\to\infty]{} (p_0, d_0, q_0), \text{ a.s.} \tag{3.8}$$

REMARK 3.1

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 have to require that $A(z)$ and $B(z)$ are coprime.

REMARK 3.2

From the proofs of Theorem 3.1 (see Appendix A), we know that s_n in criteria $\text{CIC}_1(p)_n$, $\text{CIC}_2(q)_n$ and $\text{CIC}_3(d)_n$ can be replaced by any real number sequence $\{s_n^*\}$ satisfying

$$\frac{(\log r_n^*)(\log\log r_n^*)^c}{s_n^*} \xrightarrow[n\to\infty]{} 0 \quad \text{and} \quad \frac{s_n^*}{\lambda_{\min}^*(n)} \xrightarrow[n\to\infty]{} 0. \tag{3.9}$$

In practice, conditions (3.5)–(3.6) in Theorem 3.1 are difficult to verify. In the following we remove them and use easier varifiable ones.

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

$$Ev_n = 0, \quad Ev_n^2 = \frac{1}{n^{\varepsilon}}, \quad v_n^2 \leqslant \frac{\sigma^2}{n^{\varepsilon}}, \quad \varepsilon \in \left(0, \frac{1}{2(t+1)}\right), \tag{3.10}$$

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

Without loss of generality we assume

$$\mathscr{F}_n = \{w_i, v_i, i \leqslant n\}$$

and

$$\mathscr{F}'_n = \{w_i, v_{i-1}, i \leqslant n\}.$$

Let u_n^0 be and $\mathscr{F}'_n$-adapted desired control. The attenuating excitation method suggests to implement

$$u_n = u_n^0 + v_n \tag{3.11}$$

to system.

We now assume that

$$\mathrm{H}_2. \quad \sum_{i=0}^{n} (u_i^0)^2 = o(n^{1+\delta}), \quad \text{for } \delta = \frac{1-2\varepsilon(t+1)}{2t+3} \tag{3.12}$$

and

$$\sum_{i=0}^{n} y_i^2 = o(n^b), \quad \text{for some } b > 0. \tag{3.13}$$

THEOREM 3.2

If H_1 and H_2 hold with u_n given by (3.11), then (2.9)–(2.21) lead to

$$\|\theta_n^* - \theta^*\|^2 = o\left(\frac{(\log n)(\log\log n)^c}{n^{1-(t+1)(\varepsilon+\delta)}}\right), \quad \text{a.s.} \tag{3.14}$$

and

$$(p_n, d_n, q_n,) \xrightarrow[n\to\infty]{} (p_0, d_0, q_0), \text{ a.s.} \tag{3.15}$$

In Theorem 3.2. (3.15) means that p_n, d_n and q_n are consistent, while (3.14) indicates the convergence rates of coefficient estimates to their true values.

REMARK 3.3

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

$$\frac{(\log n)(\log\log n)^c}{s_n^*}\xrightarrow[n\to\infty]{}0 \quad \text{and} \quad \frac{s_n^*}{n^{1-(t+1)(\varepsilon+\delta)}}\xrightarrow[n\to\infty]{}0. \tag{3.16}$$

REMARK 3.4

From [12]–[13] we know that the real number δ in (3.12) can be any one satisfying $\delta\in[0,\ (1-2\ (t+1))/(2t+3)]$.

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 now compare conditions used in Theorem 3.1 of this paper with those used in Theorem 2.1 of [9].

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

$$\varphi_n^\tau(p,\,1,\,q)\Big(\sum_{i=0}^{n}\varphi_i(p,\,1,\,q)\,\varphi_i^\tau(p,\,1,\,q)\Big)^{-1}\varphi_n(p,\,1,\,q)\xrightarrow[n\to\infty]{}0,\ \text{a.s.}$$

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

$$\lambda_{\min}^n(p,\,q)\xrightarrow[n\to\infty]{}\infty,\ \text{a.s.}$$

and

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

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

$$\frac{(r_n^*)^{1/2}}{\lambda_{\min}^*(n)}\xrightarrow[n\to\infty]{}0 \quad \text{and} \quad \frac{(\log r_n^*)(\log\log r_n^*)^c}{(r_n^*)^{1/2}}\xrightarrow[n\to\infty]{}0$$

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 conclusion of Theorem 3.1 follows from Remark 3.2.

REMARK 3.7

Comparing Theorem 3.2 of this paper with Theorem 3.1 in [9] one may find a situation similar to that described in Remark 3.6: weaker conditions are applied here and s_n can be taken as n^a, $\forall a \in (0, 1)$.

REMARK 3.8

In the case where we pay no attention to control performance and concern the parameter estimation only, 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 $|z| = 1$, and that we do not require the minimum phase condition.

4. ADAPTIVE TRACKING

We now design the input for a stochastic system with unknown orders and coefficients so that the system output follows a given bounded deterministic reference signal y_n^*. Specifically, we shall design u_n^0 in (3.11) so that condition H_2 in Theorem 3.2 holds and the output $\{y_n\}$ minimizes

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

In this section we assume that v_n in (3.11) has independent components and continuous distributions, and that delay d_0 is known. Let

$$A_n(z) = 1 + a_{1n} z + \ \dots\ + a_{p_n n} z^{p_n}, \tag{4.1}$$

$$B_n(z) = b_{d_0 n} z^{d_0} + \ \dots\ + b_{q_n n} z^{q_n}, \tag{4.2}$$

where p_n, q_n, $a_{in}(i = 1, \dots, p_n)$ and $b_{jn}(j = d_0, \dots, q_n)$ are given by (2.9)–(2.17) and (2.19)–(2.20) with initial value θ_0^* satisfying $b_{d_0 0} \neq 0$ which guarantees that

$$b_{d_0 n} \neq 0 \text{ a.s.}, \ n \geqslant 0, \tag{4.3}$$

when the attenuating excitation control (3.11) is applied [14].

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

$$1 = F(z)\, A(z) + G(z)\, z^{d_0}. \tag{4.4}$$

Then the system (2.1) can be written as

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

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 $F(z)\,B(z)\,z^{-d}$ is b_{d_0} which is a non-zero constant, the optimal tracking control u_n should be defined from

$$y^*_{n+d_0}=F(z)\,B(z)\,z^{-d_0}u_n+G(z)\,y_n \tag{4.6}$$

when the system parameters are all known. This motivates us to construct adaptive tracking control u_n in the following way.

Let u'_n be the solution of the following equation:

$$b_{d_0n}u'_n=y^*_{n+d_0}-(G_n(z)\,y_n+(F_nB_n)(z)\,z^{-d_0}u_n-b_{d_0n}u_n), \tag{4.7}$$

where

$$F_n(z)=1+f_{1n}z+\ \dots\ +f_{d_0-1n}z^{d_0-1}$$

and

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

are the solutions of the Diophantine equation

$$1=F_n(z)\,A_n(z)+G_n(z)\,z^{d_0}, \tag{4.8}$$

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

Finally, the adaptive tracking control u_n of system (2.1) is given by (3.11) with u^0_n defined as follows:

$$u^0_n=\begin{cases}u'_n, & \text{if } n \text{ belongs to } [\tau_k,\sigma_k) \quad \text{for some } k,\\ 0, & \text{if } n \text{ belongs to } [\sigma_k,\tau_{k+1}) \quad \text{for some } k,\end{cases} \tag{4.9}$$

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

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

$$\sigma_k=\sup\{\mu>\tau_k\colon \sum_{i=\tau_k}^{j-1}(u'_i)^2\leqslant(j-1)^{1+\delta}+(u'_{\tau_k})^2,\quad \forall j\in(\tau_k,\mu]\}, \tag{4.10}$$

$$\tau_{k+1}=\inf\{\mu>\sigma_k\colon \sum_{i=\tau_k}^{\sigma_k-1}(u'_i)^2\leqslant\frac{\mu^{1+\delta}}{2^k},\ (u'_\mu)^2\leqslant\mu^{1+\delta}\}. \tag{4.11}$$

By induction it is easy to see that u^0_n is $\mathscr{F}'_n$-measurable.

For the system (2.1) we have

THEOREM 4.1

If condition H_1 holds; $A(z)$ is stable; u_n is given by (3.11) and (4.7)–(4.11); θ^*_n, p_n, q_n are defined by (2.9)–(2.17) and (2.19)–(2.20), then

$$\|\theta^*_n-\theta^*\|^2=o\left(\frac{(\log n)\,(\log\log n)^c}{n^{1-(t+1)(\varepsilon+\delta)}}\right),\ \text{a.s.}, \tag{4.12}$$

$$(p_n, q_n) \xrightarrow[n\to\infty]{} (p_0, q_0), \text{ a.s.} \tag{4.13}$$

and

$$\frac{1}{n}\sum_{i=0}^{n}(y_i - y_i^*)^2 = \frac{1}{n}\sum_{i=0}^{n}(F(z)\,w_i)^2 + o(n^{-1/2\varepsilon}), \text{ a.s.} \tag{4.14}$$

The proof is given in Appendix B.

REMARK 4.1

If the time-delay is also unknown, then it is difficult to know whether $b_{d_n n}$ is zero or not and hence difficult to guarantee solvability of the optimal adaptive tracking control (see (4.7)).

REMARK 4.2

From [14] we know that for any u_n measurable with respect to $\mathscr{F}_n \triangleq \{w_i, i \leqslant n\}$ we have

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

So (4.12)–(4.14) mean that the adaptive control u_n defined by (3.11) and (4.7)–(4.11) is optimal.

REMARK 4.3

If conditions of Theorem 4.1 are satisfied with

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

then the conclusions of Theorem 4.1 become (4.12), (4.13) and

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}(y_i - y_i^*)^2 = 0 \text{ a.s.}$$

REMARK 4.4

Theorem 4.1 remains valid if in lieu of stability of $A(z)$ we use a weaker condition: all zeros of $A(z)$ are outside open unit disk and $G(z)$ is stable, where $G(z)$ is given by (4.4). 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 $G(z) = 1$.

5. ADAPTIVE LQ PROBLEM

In this section we shall consider adaptive LQ problem for the system (2.1). The loss function is

$$J(u) = \limsup_{n\to\infty} J_n(u), \tag{5.1}$$

where

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

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

In this section we assume that

$$\text{H}_3. \qquad \frac{1}{n}\sum_{i=0}^{n} w_i^2 = R + o(n^{-\varrho}) \text{ a.s. for some } \varrho > 0 \text{ and } R \geqslant 0. \tag{5.3}$$

We first write (2.1) in the state space form

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

$$y_k = C^\tau x_k, \quad x_0 = 0, \tag{5.5}$$

where

$$A = \begin{bmatrix} -a_1 & 1 & 0 & \cdots & \cdots & 0 \\ -a_2 & 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ & & & & \ddots & 1 \\ -a_h & 0 & 0 & \cdots & \cdots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ b_{d_0} \\ \vdots \\ b_h \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{5.6}$$

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

From [13] it is known that

$$\inf_{u \in U} J(u) = RC^\tau SC, \quad \text{for system (2.1)}, \tag{5.7}$$

and the optimal control is

$$u_n = Lx_n, \tag{5.8}$$

where

$$U = \{u\colon \sum_{j=0}^{n} u_i^2 = o(n),\ u_n^2 = o(n) \text{ a.s.}, u_n \in \mathscr{F}_n\}, \tag{5.9}$$

$$L = -(B^\tau SB + Q_2)^{-1} B^\tau SA \tag{5.10}$$

and S satisfies

$$S = A^\tau SA - A^\tau SB(B^\tau SB + Q_2)^{-1} B^\tau SA + CQ_1C^\tau \tag{5.11}$$

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 (2.9)–(2.21) we estimate A, B, C, S and x_n by $A(n)$, $B(n)$, $C(n)$, $S(n)$ and $\hat{x}_n$, respectively, as follows:

$$A(n) = \begin{bmatrix} -a_{1n} & 1 & 0 & \cdots & \cdots & 0 \\ -a_{2n} & 0 & 1 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ & & & & \ddots & \\ \vdots & \vdots & & & \ddots & 1 \\ -a_{h_n n} & 0 & 0 & \cdots & \cdots & 0 \end{bmatrix}, \quad B(n) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ b_{d_n n} \\ \vdots \\ b_{h_n n} \end{bmatrix}, \tag{5.12}$$

$$C(n) = [1\ 0\ \ldots\ldots\ 0]^{\tau}_{1\times h_n}, \qquad h_n = \max(p_n, q_n, 1), \tag{5.13}$$

$$S(n) = A^{\tau}(n)S'(n-1)A(n) - A^{\tau}(n)S'(n-1)B(n)\big(B^{\tau}(n)S'(n-1)B(n)+Q_2\big)^{-1} \times B^{\tau}(n)S'(n-1)A(n) + C(n)Q_1C^{\tau}(n) \tag{5.14}$$

with $S(0) = 0$ and $S'(n-1)$ being a square matrix of dimension $h_n \times h_n$:

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

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

$$\hat{x}_{n+1} = A(n)\hat{x}'_n + B(n)u_n + C(n)\big(y_{n+1} - C^{\tau}(n)A(n)\hat{x}'_n - C^{\tau}(n)B(n)u_n\big),$$

$$\hat{x}_0 = y_0 = 0, \tag{5.15}$$

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

$$\hat{x}'_n = \begin{cases} [\hat{x}^{\tau}_n\ \ 0]^{\tau}, & \text{if } h_{n-1} < h_n, \\ M^{\tau}(n)\hat{x}_n, & \text{if } h_{n-1} \geqslant h_n. \end{cases} \tag{5.16}$$

We now have the estimate L_n for the optimal gain L given by (5.10)

$$L_n = -\big(B^{\tau}(n)S(n)B(n) + Q_2\big)^{-1}B^{\tau}(n)S(n)A(n). \tag{5.17}$$

However, we cannot directly take $L_n\hat{x}'_n$ as the desired control u^0_n, because $L_n\hat{x}'_n$ may grow too fast so that H_2 is not satisfied.

Define

$$L^0_n = \begin{cases} L_n, & \text{if } n \in [\tau_k, \sigma_k) \quad \text{for some } k, \\ 0, & \text{if } n \in [\sigma_k, \tau_{k+1}) \quad \text{for some } k, \end{cases} \tag{5.18}$$

$$u^0_n = L^0_n \hat{x}'_n, \tag{5.19}$$

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

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

$$\sigma_k=\sup\{\mu>\tau_k\colon \sum_{i=\tau_k}^{j-1}(L_i\hat{x}_i')^2\leqslant(j-1)^{1+\delta}+(L_{\tau_k}\hat{x}_{\tau_k}')^2,\ \forall j\in(\tau_k,\mu]\},\tag{5.20}$$

$$\tau_{k+1}=\inf\left\{\mu>\sigma_k\colon \sum_{i=\tau_k}^{\sigma_k-1}(L_i\hat{x}_i')^2\leqslant\frac{\mu^{1+\delta}}{2^k},\right.$$

$$\left.\sum_{j=1}^{\mu}\|\hat{x}_j'\|^2\leqslant\mu^{1+1/2\delta},\ \frac{(L_\mu\hat{x}_\mu')^2}{\mu^{1+\delta}}\leqslant 1\right\}.\tag{5.21}$$

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 and q_n are defined by (2.9)–(2.21), then u_n defined by (3.11) and (5.18)–(5.21) is optimal in the sense that

$$(p_n,d_n,q_n)\xrightarrow[n\to\infty]{}(p_0,d_0,q_0),\ \text{a.s.},\tag{5.22}$$

$$\|\theta_n^*-\theta^*\|^2=o\left(\frac{(\log n)(\log\log n)^c}{n^{1-(t+1)\varepsilon}}\right),\ \text{a.s.}\tag{5.23}$$

and

$$J_n(u)=RC^\tau SC+o(n^{-(\varrho\wedge\varepsilon)}),\ \text{a.s.}\tag{5.24}$$

PROOF

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

$$\sum_{i=0}^{n}(L_i^0\hat{x}_i')^2=o(n^{1+\delta}).\tag{5.25}$$

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

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

Then theorem 3.2 asserts (3.14) and (3.15) by which we know that $(p_n,d_n,q_n)=(p_0,d_0,q_0)$ for n starting from some $\bar{N}_0>0$.

Hence Theorem 4 of [12] applies to the present case and (5.22), (5.23) follow immediately.

REMARK 5.1

From (3.11) and (5.12)–(5.21) 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.

REMARK 5.2

If the lower bound for time-delay and upper bounds for system orders are treated as the true delay and true orders, respectively, in the adaptive control design the structure of the controller is at least as complex as that of the system to be controlled [11]. Consistently estimating time-delay and system orders not only gives better knowledge about the system which possibly is important for the user, but also provides simpler structure of the controller. Furthermore, if the true time-delay is strictly greater than its lower bound, then the estimate for the leading coefficient in $B(z)$ will tend to zero. As consequence, the control for adaptive tracking will be unboundedly increased and this is unacceptable in the practice.

6. CONCLUSION

The paper gives recursive parameter estimates for system (2.1) 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 rates of convergence both of the estimates to their true values and of the loss functions to their minimums 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 $\mathrm{CIC}_1(p)_n$, $\mathrm{CIC}_2(q)_n$ and $\mathrm{CIC}_3(d)_n$.

LEMMA A.1

Under the conditions of Theorem 3.1 we have

$$\mathrm{CIC}_1(p)_n-\mathrm{CIC}_1(p_0)_n \geqslant \begin{cases} s_n(p-p_0+o(1)), \text{ a.s.,} & \text{if } p \geqslant p_0, \\ \lambda^*_{\min}(n)(\bar{\alpha}_0+o(1)), \text{ a.s.,} & \text{if } p < p_0; \end{cases} \tag{A.1}$$

$$\mathrm{CIC}_2(q)_n-\mathrm{CIC}_2(q_0)_n \geqslant \begin{cases} s_n(q-q_0+o(1)), \text{ a.s.,} & \text{if } q \geqslant q_0, \\ \lambda^*_{\min}(n)(\bar{\alpha}_0+o(1)), \text{ a.s.,} & \text{if } q < q_0; \end{cases} \tag{A.2}$$

$$\mathrm{CIC}_3(d)_n-\mathrm{CIC}_3(d_0)_n \geqslant \begin{cases} s_n(d_0-d+o(1)), \text{ a.s.,} & \text{if } d \leqslant d_0, \\ \lambda^*_{\min}(n)(\bar{\alpha}_0+o(1)), \text{ a.s.,} & \text{if } d > d_0, \end{cases} \tag{A.3}$$

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

PROOF

We first prove (A.1). For any $0 \leqslant p \leqslant p^*$, set

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

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 $p \times (p^* - p)$.

If $p \geqslant p_0$, then

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

and

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

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

$$\tilde{\theta}_n^* \triangleq \theta^* - \theta_n^* = -\Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau} + I\Big)^{-1}\Big(\sum_{i=0}^{n-1} \varphi_i^* w_{i+1} - \theta^*\Big), \tag{A.7}$$

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

and

$$\begin{aligned}\sigma_n(p, d^*, q^*) = &\Big(\Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau} + I\Big)^{-1/2}\Big(\sum_{i=0}^{n-1} \varphi_i^* w_{i+1} - \theta^*\Big)\Big)^\tau \Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau} + I\Big)^{-1/2} H^\tau(p) \\ &H(p)\Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau}\Big)H^\tau(p)H(p)\Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau} + I\Big)^{-1/2}\Big(\Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau} + I\Big)^{-1/2} \\ &\Big(\sum_{i=0}^{n-1} \varphi_i^* w_{i+1} - \theta^*\Big)\Big) - 2\Big(\Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau} + I\Big)^{-1/2}\Big(\sum_{i=0}^{n-1} \varphi_i^* w_{i+1} - \theta^*\Big)\Big)^\tau \\ &\Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau} + I\Big)^{-1/2} H^\tau(p)H(p)\Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau} + I\Big)^{1/2}\Big(\Big(\sum_{i=0}^{n-1} \varphi_i^*\varphi_i^{*\tau} + I\Big)^{-1/2} \\ &\sum_{i=0}^{n-1} \varphi_i^* w_{i+1}\Big) + \sum_{i=0}^{n-1} w_{i+1}^2.\end{aligned} \tag{A.8}$$

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

$$\|(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau}+I)^{-1/2}H^\tau(p)H(p)(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau})H^\tau(p)H(p)(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau}+I)^{-1/2}\|$$

$$\leqslant \operatorname{tr}\Big((\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau}+I)^{-1/2}H^\tau(p)H(p)(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau})H^\tau(p)H(p)(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau}+I)^{-1/2}\Big)$$

$$= \operatorname{tr}\Big((T(p)H^\tau(p)H(p)T^\tau(p)T(p)(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau}+I)^{-1}T^\tau(p)T(p)H^\tau(p)H(p)T^\tau(p)T(p)$$
$$(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau})T^\tau(p)T(p)H^\tau(p)H(p)T^\tau(p)\Big)$$

$$= \operatorname{tr}\Big(F(p)(T(p)(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau})T^\tau(p)+I)^{-1}F(p)T(p)(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau})T^\tau(p)F(p)\Big)=o(1) \tag{A.9}$$

and

$$\|(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau}+I)^{-1/2}H^\tau(p)H(p)(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau}+I)^{1/2}\|^2=o(1). \tag{A.10}$$

By Lemma 2 of [8] we also have

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

Combining (A.8)–(A.11) yields

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

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

Set

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

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

When $p<p_0$, we have

$$\|\tilde\theta_n^*(p)\|^2\geqslant \min(a_{p_0}^2,\,b_{d_0}^2,\,b_{q_0}^2)\triangleq\bar\alpha_0>0 \tag{A.14}$$

and hence by Lemma 2 of [8]

$$\sigma_n(p,d^*,q^*)=\tilde\theta_n^{*\tau}(p)(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau})\tilde\theta_n^*(p)+2\tilde\theta_n^{*\tau}(p)\sum_{i=0}^{n-1}\varphi_i^*w_{i+1}+\sum_{i=0}^{n-1}w_{i+1}^2$$

$$\geqslant\|(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau})^{1/2}\tilde\theta_n^*(p)\|^2+\sum_{i=0}^{n-1}w_{i+1}^2-o$$
$$\Big(\|(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau})^{1/2}\tilde\theta_n^*(p)\|((\log r_n^*)(\log\log r_n^*)^c)^{1/2}\Big)$$

$$\geqslant\|\tilde\theta_n^*(p)\|^2\lambda_{\min}^*(n)\left(1-o\left(\frac{((\log r_n^*)(\log\log r_n^*)^c)^{1/2}}{\|(\sum_{i=0}^{n-1}\varphi_i^*\varphi_i^{*\tau})^{1/2}\tilde\theta_n^*(p)\|}\right)\right)+\sum_{i=0}^{n-1}w_{i+1}^2$$

$$\geqslant \bar{\alpha}_0 \lambda^*_{\min}(n)\left(1-o\left(\left(\frac{(\log r^*_n)(\log\log r^*_n)^c}{\lambda^*_{\min}(n)}\right)^{1/2}\right)\right)+\sum_{i=0}^{n-1} w^2_{i+1}$$

$$= \lambda^*_{\min}(n)\left(\bar{\alpha}_0+o(1)\right)+\sum_{i=0}^{n-1} w^2_{i+1},$$

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 Lemma 2 of [8].

Since (p_n, d_n, q_n) belongs to a finite set, for (3.8) we need only to show that any limit point of sequence $\{(p_n, d_n, q_n)\}$ is nothing but (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

$$0 \geqslant \mathrm{CIC}_1(p_{n_k})_{n_k}-\mathrm{CIC}_1(p_0)_{n_k} = \mathrm{CIC}_1(p')_{n_k}-\mathrm{CIC}_1(p_0)_{n_k}$$

$$\geqslant \lambda^*_{\min}(n_k)\left(\bar{\alpha}_0+o(1)\right)\xrightarrow[k\to\infty]{}\infty.$$

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

$$0 \geqslant \mathrm{CIC}_1(p')_{n_k}-\mathrm{CIC}_1(p_0)_{n_k} \geqslant s_{n_k}(p'-p_0+o(1)), \qquad \text{for } p' \geqslant p_0.$$

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. as $n\to\infty$.

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 $n \to \infty$.

PROOF OF THEOREM 3.2

Obviously, from Theorem 3.1 we need only to show that

$$\liminf_{n\to\infty} n^{-1+(t+1)(\varepsilon+\delta)} \lambda^*_{\min}(n) \neq 0, \text{ a.s.} \tag{A.15}$$

By the argument used in [12] for this it suffices to show that there does not exist a non-zero vector

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

such that

$$\sum_{i=0}^{p^*-1} \alpha_i z^i B(z) = \sum_{i=0}^{q^*-1} \beta_i z^i A(z) \quad \text{and} \quad \sum_{i=0}^{p^*-1} \alpha_i z^i = 0. \tag{A.16}$$

This is true indeed, because the second equation of (A.16) leads to $\alpha_i = 0$ $(i = 0, \ldots, p^*-1)$, then by the first equation of (A.16) $\beta_j = 0$ $(j = 0, \ldots, q^*-1)$.

APPENDIX B

PROOF OF THEOREM 4.1

Theorem 3.2 implies (4.12)–(4.13) if we can verify H_2, for which it suffices to show (3.12) because of stability of $A(z)$. This can be done by a method similar to that used in [13]:

If $\tau_k < \infty$ and $\sigma_k = \infty$ for some k then by (4.9) $u^0_i \equiv u'_i$ for $i \geqslant \tau_k$ and (3.13) follows from (4.10).

If $\sigma_k < \infty$ and $\tau_{k+1} = \infty$ for some k, then by (4.9) $u^0_i \equiv 0$ for $i \geqslant \sigma_k$ and (3.12) is trivial.

If $\tau_k < \infty$ and $\sigma_k < \infty$ for all k, then by (4.9)–(4.11) for any $k \geqslant 1$

$$\sup_{\tau_k \le n < \tau_{k+1}} \frac{1}{n^{1+\delta}} \sum_{i=1}^{n} (u_i^0)^2 = \sup_{\tau_k \le n \le \sigma_k - 1} \frac{1}{n^{1+\delta}} \sum_{i=\tau_1}^{n} (u_i^0)^2$$

$$= \sup_{\tau_k \le n \le \sigma_k - 1} \frac{1}{n^{1+\delta}} \Big(\sum_{i=\tau_1}^{\sigma_1 - 1} (u_i')^2 + \ldots + \sum_{i=\tau_{k-1}}^{\sigma_{k-1}-1} (u_i')^2 + \sum_{i=\tau_k}^{n} (u_i')^2 \Big)$$

$$\le \frac{1}{\tau_2^{1+\delta}} \sum_{i=\tau_1}^{\sigma_1 - 1} (u_i')^2 + \ldots + \frac{1}{\tau_k^{1+\delta}} \sum_{i=\tau_{k-1}}^{\sigma_{k-1}-1} (u_i')^2 + \sup_{\tau_k \le n \le \sigma_k - 1} \frac{1}{n^{1+\delta}} \sum_{i=\tau_k}^{n} (u_i')^2$$

$$\le \sum_{i=1}^{k-1} \frac{1}{2^i} + \sup_{\tau_k \le n \le \sigma_k - 1} \frac{1}{n^{1+\delta}} (n^{1+\delta} + (u'_{\tau_k})^2) \le 3, \text{ a.s.},$$

which verifies (3.12), and hence (4.12) and (4.13) hold.

We now prove (4.14). By (4.5) and (4.7) we have

$$y_{n+d_0} - y^*_{n+d_0} = F(z) B(z) z^{-d_0} u_n - (F_n B_n)(z) z^{-d_0} u_n + G(z) y_n - G_n(z) y_n + b_{d_0} v_n + F(z) w_{n+d_0}. \tag{B.1}$$

By (4.16) we have $(p_n, q_n) \equiv (p_0, q_0)$ for sufficiently large n, and by stability of $A(z)$ we find that

$$\sum_{i=0}^{n} y_i^2 = o(n^{1+\delta}), \text{ a.s.}$$

Then by the argument used for proving (F8) and (52) in [14] we see

$$\sum_{i=0}^{n} \big(F(z) B(z) z^{-d_0} u_i - (F_i B_i)(z) z^{-d_0} u_i + G(z) y_i - G_i(z) y_i + b_{d_0} v_i \big)^2 = o(n^{1-\varepsilon})$$

and obtain the estimate (4.17).

REFERENCES

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

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

[3] Akaike H., *On entropy maximization principle*, [in:] *Applications of Statistics* (ed. P. R. Krishnaiah), North Holland, Amsterdam 1977, pp. 27–41.

[4] Rissanen J., *Modelling by shortest data description*, Automatica, 14, 467–471, 1978.

[5] Hannan E. J., Quinn B. G., *The determination of the order of an autoregression*, J. Roy. Statist. Soc., B.41, 190–195, 1979.

[6] Rissanen J., *Order estimation by accumulated prediction errors*, Essays in Time Series and Allied Process (eds J. Gani, M. B. Priestley), J. Appl. Probab., Special Vol. 23A, 55–61, 1986.

[7] Chen H. F., Guo L., *Consistent estimation of the order of stochastic control systems*, IEEE Trans. Automatic Control AC-32, No. 6, 531–535, 1987.

[8] Guo L., Chen H. F., Zhang J. F., *Consistent order estimation for linear stochastic feedback control systems (CARMA model)*, Automatica, 25, No. 1, 147–151, 1989.

[9] Hemerly E. M., Davis M. H. A., *Recursive order estimation of stochastic control systems*, Mathematics of Control, Signals and Systems, 1989 (in press).

[10] Goodwin G. C., Sin K. S., *Adaptive filtering, prediction and control*, Prentice Hall, Englewood Cliffs 1984.

[11] Chen H. F., Guo L., *Convergence rate of least-squares identification and adaptive control for stochastic systems*, Int. J. Contr., 44, No. 5, 1459–1476, 1986.

[12] SHAKED U., KUMAR P. R., *Minimum variance control of discrete time multivariable ARMAX systems*, Siam J. Contr. a. Optim., 24, 396–411, 1986.

[13] CHEN H. F., GUO L., *Optimal adaptive control and consistent parameter estimates for ARMAX model with quadratic cost*, Siam J. Contr. a. Optim., 25, No. 4, 845–867, 1987.

[14] CHEN H. F., ZHANG J. F., *Convergence rates in stochastic adaptive tracking*, Int. J. Control, 1989 (in press).

[15] MIDDLETON R. H., GOODWIN G. C., HILL D. J., MAYNE D. Q., *Design issues in adaptive control*, IEEE Trans. Automatic Control AC-33, 50–58, 1988.

[16] NIU S. H., XIAO D. Y., *A recursive algorithm for parameter estimation and order identification of SISO model*, J. Tsinghua University, 28, 24–31, 1988.

[17] CHEN H. F., GUO L., *Adaptive control via consistent estimation for deterministic systems*, Int. J. Control, 45, 2183–2202, 1987.

Received May 7, 1989

ANALIZA ZBIEŻNOŚCI DLA SYSTEMÓW ADAPTACYJNEGO STEROWANIA O NIE ZNANYCH RZĘDACH, OPÓŹNIENIACH I WSPÓŁCZYNNIKACH

Analizowane są zadania identyfikacji i adaptacyjnego sterowania dla liniowych, dyskretnych systemów stochastycznych o nie znanych rzędach, opóźnieniach czasowych i współczynnikach w obecności nieskolerowanego szumu. Przy założeniu, że znane są: dolne ograniczenie na wartość opóźnienia i górne ograniczenie na rząd systemu, podano metodę obliczania zbieżnych estymat czasu opóźnienia, rzędu systemu i współczynników; zaprojektowano algorytm rozwiązywania zadania optymalnego śledzenia i zadania optymalnego sterowania adaptacyjnego dla kwadratowej funkcji strat; oszacowano szybkość zbieżności zarówno estymat współczynników do ich rzeczywistej wartości, jak i wartości funkcji strat do jej minimalnej wartości.