INT. J. CONTROL, 1989, VOL. 49, NO. 6, 1915–1935

Convergence rates in stochastic adaptive tracking†

HAN-FU CHEN‡ and JI-FENG ZHANG‡

For stochastic control systems described by the ARMAX model with unknown matrix coefficients, the stochastic adaptive control is designed so that the parameter estimates converge to the true values with a rate of convergence $O((\log n)(\log\log n)^c/n^\alpha)$ with $\alpha > 0$, $c > 1$ and the tracking error tends to its minimum value at a speed of $O(n^{-1/2\varepsilon})$ with $\varepsilon > 0$.

1. Introduction

Since Åström and Wittenmark introduced the self-tuning regulator in 1973, the stochastic adapative tracking problem has drawn much attention from control scientists (Goodwin *et al.* 1981, Goodwin and Sin 1979, Chen and Caines 1985, Caines and Lafortune 1984, Chen and Guo 1987 a, b, Chen 1984). The difficult and important problem of simultaneously identifying unknown parameters and tracking a reference signal was first considered by Caines and Lafortune (1984), where consistent estimation and suboptimal tracking were simultaneously obtained. Subsequently, Chen and Guo (1987 a, b, 1986 a), introducing an attenuating excitation to the control, achieved the consistent parameter estimate and the minimal tracking error simultaneously. However, these results are all based on the stochastic gradient algorithm for parameter estimation which is not as good with convergence properties as the estimate of least squares (ELS) algorithm. The essential difficulty of applying the ELS algorithm to the stochastic adaptive tracking problem consists in that *a posteriori* rather than the *a priori* information is used in the ELS algorithm, as pointed out by Kumar (1985).

The ELS-based adaptive tracker for the unit delay case is designed by Lai and Wei (1986) for single-input single-output systems with bounded noise and by Guo and Chen (1987) for stable multi-input multi-output systems but without boundedness restriction for noise. In this paper the multi-delay case is treated and the stability assumption on the system required by Guo and Chen (1987) has been removed so that the same convergence rates as those obtained by Guo and Chen (1987) are established.

2. Statement of the problem

Consider the stochastic control system

$$A(z)y_n = B(z)u_{n-d} + C(z)w_n, \quad d \geq 1; \quad y_i = w_i = 0, \quad u_i = 0, \quad i < 0 \tag{1}$$

Received 5 February 1988.

† Work supported by the National Natural Science Foundation of China and the TWAS Research Grant No. 87–43.

‡ Institute of Systems Science, Academia Sinica, Beijing, People's Republic of China.

with l inputs u_n, m outputs y_n and m-dimensional driven noise w_n, where

$$A(z) = I + A_1 z + \dots + A_p z^p, \quad p \geqslant 0 \tag{2}$$

$$B(z) = B_1 + B_2 z + \dots + B_q z^{q-1}, \quad q \geqslant 1 \tag{3}$$

$$C(z) = I + C_1 z + \dots + C_r z^r, \quad r \geqslant 0 \tag{4}$$

are matrix polynomials in a shift-back operator with unknown coefficients denoted by

$$\theta^\tau = [-A_1 \quad \dots \quad -A_p \quad B_1 \quad \dots \quad B_q \quad C_1 \quad \dots \quad C_r] \tag{5}$$

and $\{w_n, \mathscr{F}_n\}$ is a martingale difference sequence with

$$\sup_n E[\|w_{n+1}\|^2/\mathscr{F}_n] < \gamma < \infty, \quad \text{a.s.} \tag{6}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} w_i w_i^\tau = R > 0 \tag{7}$$

Problem A

In the adaptive tracking problem the $\mathscr{F}_n$-measurable control u_n is designed to force the output y_{n+d} to follow a given $\mathscr{F}_n$-measurable reference signal y^*_{n+d}. In addition, $\{y^*_n\}$ is mutually independent of $\{w_n\}$.

It is not difficult to show (see Appendix A) that

$$\lim_{n\to\infty} \sup \frac{1}{n} \sum_{i=0}^{n} \|y_i - y_i^*\|^2 \geqslant \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} (F(z)w_i)^\tau \cdot (F(z)w_i) = \operatorname{tr} \sum_{j=0}^{d-1} F_j R F_j^\tau \tag{8}$$

where $G(z)$ and $F(z) = F_0 + F_1 z + \dots + F_{d-1} z^{d-1}$ with $F_0 = I$ are the unique solution (see Appendix A) of the Diophantine equation

$$(\det C(z))I = F(z)(\operatorname{Adj} C(z))A(z) + z^d G(z) \tag{9}$$

So it is natural to call u_n leading to

$$\lim_{n\to\infty} \sup \frac{1}{n} \sum_{i=0}^{n} (y_i - y_i^*)(y_i - y_i^*)^\tau = \sum_{j=0}^{d-1} F_j R F_j^\tau$$

the optimal control and the convergence rate of

$$\left\| \frac{1}{n} \sum_{i=0}^{n} (y_i - y_i^*)(y_i - y_i^*)^\tau - \frac{1}{n} \sum_{i=0}^{n} (F(z)w_i)(F(z)w_i)^\tau \right\|$$

the tracking speed.

Problem B

In the stochastic model reference adaptive control (MRAC) problem the $\mathscr{F}_n$-measurable control u_n is designed in order to reduce system (1) to

$$A^0(z) y_n = B^0(z) u^0_{n-d} \tag{10}$$

where $A^0(z)$ and $B^0(z)$ are given matrix polynomials of orders $\bar{p}$ and $\bar{q}$ respectively, $A^0(z)$ is stable and u^0_n is $\mathscr{F}_n$-measurable external input. Obviously, (1) can be written in the form

$$A^0(z) y_n = B^0(z) u^0_{n-d} + \varepsilon_n \tag{11}$$

with

$$\varepsilon_n = (A^0(z) - A(z))y_n + B(z)u_{n-d} - B^0(z)u^0_{n-d} + C(z)w_n \tag{12}$$

It is not difficult to show (see Appendix B) that

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} \varepsilon_i^\tau \varepsilon_i \geqslant \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} (\bar{F}(z)w_i)^\tau \cdot (\bar{F}(z)w_i) = \operatorname{tr}\sum_{j=0}^{d-1} \bar{F}_j R \bar{F}_j^\tau \tag{13}$$

where

$$\bar{F}(z) = \bar{F}_0 + \bar{F}_1 z + \ldots + \bar{F}_{d-1} z^{d-1} \tag{14}$$

and

$$A^0(z)F(z) = \bar{F}_0 + \bar{F}_1 z + \ldots + \bar{F}_{d-1} z^{d-1} + z^d \bar{N}(z) \tag{15}$$

Therefore, the adaptive control u_n leading to

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} \varepsilon_i \varepsilon_i^\tau = \sum_{j=0}^{d-1} \bar{F}_j R \bar{F}_j^\tau$$

is optimal and the speed of

$$\left\| \frac{1}{n}\sum_{i=0}^{n} \varepsilon_i \varepsilon_i^\tau - \frac{1}{n}\sum_{i=0}^{n} (\bar{F}(z)w_i)(\bar{F}(z)w_i)^\tau \right\|$$

is the convergence rate of MRAC.

In the present paper we shall give optimal stochastic adaptive controls based on the ELS algorithm for both problems of A and B and characterize their convergence rates.

For the unknown matrix θ the ELS estimate θ_n is defined as follows:

$$\theta_{n+1} = \theta_n + a_n P_n \varphi_n (y_{n+1}^\tau - \varphi_n^\tau \theta_n) \tag{16}$$

$$P_{n+1} = P_n - a_n P_n \varphi_n \varphi_n^\tau P_n,\ a_n = (1 + \varphi_n^\tau P_n \varphi_n)^{-1} \tag{17}$$

$$P_o = sI, \quad s = m(p+r) + lq$$

$$\varphi_n^\tau = [y_n^\tau \quad \cdots \quad y_{n-p+1}^\tau \quad u_{n-d+1}^\tau \quad \cdots \quad u_{n-d-q+2}^\tau \\ y_n^\tau - \varphi_{n-1}^\tau \theta_n \quad \cdots \quad y_{n-r+1}^\tau - \varphi_{n-r}^\tau \theta_{n-r+1}] \tag{18}$$

with an arbitrary initial value

$$\theta_0^\tau = [-A_{10} \ldots -A_{p0} \quad B_{10} \ldots B_{q0} \quad C_{10} \ldots C_{r0}]$$

3. Adaptive control laws for $m = l$

Problem A

Set

$$A_n(z) = I + A_{1n} z + \ldots + A_{pn} z^p \tag{19}$$

$$B_n(z) = B_{1n} + B_{2n} z + \ldots + B_{qn} z^{q-1} \tag{20}$$

$$C_n(z) = I + C_{1n} z + \ldots + C_{rn} z^r \tag{21}$$

$$\det C(z) = 1 + \bar{c}_1 z + \ldots + \bar{c}_{mr} z^{mr} \tag{22}$$

$$G(z) = G_0 + G_1 z + \dots + G_{p_1} z^{p_1} \tag{23}$$

$$F(z)(\mathrm{Adj}\ C(z)) B(z) = D_0 + D_1 z + \dots + D_{p_2} z^{p_2} \tag{24}$$

$$\theta^* = [G_0 G_1 \dots G_{p_1} \quad D_0 D_1 \dots D_{p_2} \quad \bar{c}_1 I \dots \bar{c}_{mr} I]^\tau$$

where $D_0 = B_1$ is non-degenerate, $p_2 = (m-1)r + q + d - 1$ and $p_1 = \max(mr - d, (m-1)r + p - 1)$.

In Guo and Chen (1987), for an ELS-based adaptive tracker the desired control $u_n^{(1)}$ is defined by

$$B_{1n} u_n^{(1)} = y_{n+1}^* - \theta_n^\tau \varphi_n + B_{1n} u_n \tag{25}$$

where B_{1n} denotes the estimate given by θ_n for B_1 and u_n is the attenuately excited version of $u_n^{(1)}$. Clearly, (25) is reduced to the well-known equation defining adaptive tracking control if the dither is removed, i.e. $u_n \equiv u_n^{(1)}$ (Goodwin *et al.* 1981, Kumar 1985).

In this paper the delay may be greater than one and in order to define an appropriate adaptive control we should find the predicted value $\hat{y}_{n+d}$ for y_{n+d} on the basis of $\{u_i, y_i, i \leqslant n\}$. Using (1) and (9) we have

$$\begin{aligned}(\det C(z))(y_n - F(z) w_n) &= G(z) y_{n-d} - (\det C(z)) F(z) w_n \\ &\quad + F(z)(\mathrm{Adj}\ C(z)) A(z) y_n \\ &= G(z) y_{n-d} - (\det C(z)) F(z) w_n \\ &\quad + F(z)(\mathrm{Adj}\ C(z)) B(z) u_{n-d} + (\det C(z)) F(z) w_n \\ &= G(z) y_{n-d} + F(z)(\mathrm{Adj}\ C(z)) B(z) u_{n-d}\end{aligned} \tag{26}$$

Hence

$$\hat{y}_{n+d} = (\det C(z))^{-1} (G(z) y_n + F(z)(\mathrm{Adj}\ C(z)) B(z) u_n)$$

and the optimal tracking control should be defined from

$$(\det C(z)) y_{n+d}^* = G(z) y_n + F(z)(\mathrm{Adj}\ C(z)) B(z) u_n \tag{27}$$

This leads us to define the undisturbed adaptive control $u_n^{(1)}$ from the following equation

$$B_{1n} u_n^{(1)} = (\det C_n(z)) y_{n+d}^* - G_n(z) y_n - (F_n(z)(\mathrm{Adj}\ C_n(z)) B_n(z)) u_n + B_{1n} u_n \tag{28}$$

instead of (25), where $F_n(z)$ and $G_n(z)$ satisfy the Diophantine equation

$$(\det C_n(z)) I = F_n(z)(\mathrm{Adj}\ C_n(z)) A_n(z) + z^d G_n(z) \tag{29}$$

which is tantamount to (9) and is solvable as shown in Appendix A.

If the growth rates of the input and output are not too fast, then we take $u_n^{(1)}$ as the desired control with the hope of getting better parameter estimates because the ELS is implemented. If the output grows up too fast, then we switch the parameter estimate from the ELS algorithm to the stochastic gradient (SG) algorithm and the minimum-phase property guarantees y_n tracking y_n^*. Finally, if $u_n^{(1)}$ itself grows too fast, we then simply take zero as the desired control. To be precise, we specify the random intervals on which we apply one or other desired control. Let the stopping times $\{\tau_i\}$ and $\{\sigma_i\}$

be such that

$$0=\tau_1<\sigma_1<\tau_2<\sigma_2<\dots$$

$$\sigma_k=\sup\left\{\tau>\tau_k:\sum_{i=\tau_k}^{j-1}\|y_i\|^2\leqslant(j-1)^{1+1/2\delta}+\|y_{\tau_k}\|^2,\quad\forall j\in(\tau_k,\tau]\right\}\tag{30}$$

$$\tau_{k+1}=\inf\left\{\tau>\sigma_k:\sum_{i=\sigma_k}^{\tau}\|y_i\|^2\leqslant\frac{\tau\log\tau}{2^k},\sum_{i=\tau_k}^{\sigma_k-1}\|y_i\|^2\leqslant\frac{\tau\log\tau}{2^k}\right.$$
$$\left.\sum_{i=\sigma_k}^{\tau}\|u_i\|^2\leqslant\frac{\tau\log\tau}{2^k},\sum_{i=\tau_k}^{\sigma_k-1}\|u_i\|^2\leqslant\frac{\tau\log\tau}{2^k}\right\}\tag{31}$$

where

$$\delta\in\left[0,\frac{1-2\varepsilon(\mu+1)}{2\mu+3}\right),\quad\varepsilon\in\left(0,\frac{1}{2(\mu+1)}\right)\tag{32}$$

$$\mu=mp+\max(p,q,r).\tag{33}$$

Paying attention to (26), the SG algorithm estimating θ^* is given by

$$\theta^*_{nd+n_0}=\theta^*_{(n-1)d+n_0}+\frac{\bar{a}}{r^*_{(n-1)d+n_0}}\varphi^*_{(n-1)d+n_0}[y^\tau_{nd+n_0}-\varphi^{*\tau}_{(n-1)d+n_0}\theta^*_{(n-1)d+n_0}]\tag{34}$$

for $n_0=0,1,\dots,\quad d-1$; and

$$\varphi^*_n=[y^\tau_n\dots y^\tau_{n-p_1}u^\tau_n\dots u^\tau_{n-p_2}-\varphi^{*\tau}_{n-1}\theta^*_{n-1}\dots-\varphi^{*\tau}_{n-mr}\theta^*_{n-mr}]^\tau\tag{35}$$

$$r^*_n=r^*_{n-1}+\|\varphi^*_n\|^2,\quad r^*_{-1}=r^*_{-2}=\dots=r^*_{-d}=1\tag{36}$$

where the initial value $\theta^*_{n_0-d}(n_0=0,1,\dots,d-1)$ is arbitrary but with $D^*_{0n_0-d}$ being non-degenerate where D^*_{0n} is the component of θ^*_n written in the form

$$\theta^*_n=[G^*_{0n}\dots G^*_{p_1n}\quad D^*_{0n}\dots D^*_{p_2n}\quad C^*_{1n}\dots C^*_{mrn}]^\tau\tag{37}$$

Similar to (25), the SG-based adaptive tracking control $u^{(2)}_n$ should satisfy the following equation

$$D^*_{0n}u^{(2)}_n=y^*_{n+d}-\theta^{*\tau}_n\varphi^*_n+D^*_{0n}u_n.\tag{38}$$

The desired control u^s_n is then defined by

$$u^s_n=\begin{cases}u^{(1)}_n, & \text{if } n \text{ belongs to some } [\tau_k,\sigma_k)\cap\Lambda\\ 0, & \text{if } n \text{ belongs to some } [\tau_k,\sigma_k)\cap\Lambda^c\\ u^{(2)}_n, & \text{if } n \text{ belongs to some } [\sigma_k,\tau_{k+1})\end{cases}\tag{39}$$

where $u^{(1)}_n$ is defined by (28) and Λ is a set of integers

$$\Lambda=\{j:\|u^{(1)}_j\|^2\leqslant j^{1+\delta}\}\tag{40}$$

and the adaptive control u_n for tracking is the attenuately excited version of u^s_n (Chen and Guo 1987 a, b, 1986 a; Guo and Chen 1987; Kumar 1985; Lai and Wei 1986), i.e.

$$u_n=u^s_n+v_n\tag{41}$$

where $\{v_n\}$ is a sequence of m-dimensional mutually independent and independent of $\{w_n\}$ and $\{y^*_n\}$ random vectors with independent components having continuous distributions so that

$$Ev_n = 0,\ Ev_n v_n^\tau = \frac{\bar{\sigma}_1^2}{n^\varepsilon} I,\ \|v_n\|^2 \leqslant \bar{\sigma}_2^2/n^\varepsilon, \quad v_i = 0, \quad \text{for} \quad i \leqslant 0 \tag{42}$$

where $\bar{\sigma}_1$ and $\bar{\sigma}_2$ are two constant numbers.

Problem B

By (11), (15) and (26) we first rewrite

$$\begin{aligned}\varepsilon_n &= A^0(z)y_n - B^0(z)u_{n-d}^0 \\ &= A^0(z)F(z)w_n - (\det C(z))^{-1}A^0(z)(G(z)y_{n-d} \\ &\quad + F(z)(\text{Adj } C(z))B(z)u_{n-d}) - B^0(z)u_{n-d}^0 \\ &= \bar{F}(z)w_n + \bar{N}(z)w_{n-d} - B^0(z)u_{n-d}^0 - (\det C(z))^{-1}A^0(z)(G(z)y_{n-d} \\ &\quad + F(z)(\text{Adj } C(z))B(z)u_{n-d})\end{aligned} \tag{43}$$

Noticing that in the exact matching case $\varepsilon_n \equiv 0$ and $\bar{F}(z)w_n$ is independent of the other terms on the right-hand side of (43) we find that the optimal control u_n should satisfy

$$0 = \bar{N}(z)w_n - B^0(z)u_n^0 - (\det C(z))^{-1}A^0(z)(G(z)y_n + F(z)(\text{Adj } C(z))B(z)u_n)$$

Hence, replacing (28) we define $u_n^{(1)}$ from the following equation

$$\begin{aligned}B_{1n}u_n^{(1)} &= (\det C_n(z))(A^0(z))^{-1}(B^0(z)u_n^0 - \bar{N}_n(z)(y_n - \theta_n^\tau \varphi_{n-1})) \\ &\quad + G_n(z)y_n + (F_n(z)(\text{Adj } C_n(z))B_n(z))u_n + B_{1n}u_n\end{aligned} \tag{44}$$

where $\bar{N}_n(z)$ is given by

$$A^0(z)F_n(z) = \bar{F}_{0n} + \bar{F}_{1n}z + \ldots + \bar{F}_{d-1n}z^{d-1} + \bar{N}_n(z)z^d$$

Adaptive control u_n is again defined by (30)–(42) but with $u_n^{(1)}$ given by (44) rather than (28).

Lemma 1

Let u_i^0 and $u_i^{(k)}$ be measurable with respect to $\mathscr{F}_i' = \{w_j, v_{j-1}, y_{j+d}^*, j \leqslant i\}$ for any $i \leqslant n$ and $k = 1, 2$. Then $u_n^{(k)}(k = 1, 2)$ can be solved from (44), (28) and (38) respectively.

Proof

The proof is given in Appendix C.

4. Main results ($m = l$)

We first list conditions used later on.

(i) $C^{-1}(z) - \frac{1}{2}I$ is SPR.

(ii) $\det C(z) - (\bar{a}/2)$ is SPR for some $\bar{a} > 0$.

(iii) All zeros of $\det B(z)$ lie outside the closed unit disc.

(iv) $A(z)$, $B(z)$ and $C(z)$ have no common left factor and A_p is of full rank.

Lemma 2

If condition (ii) holds then

$$\sum_{i=0}^{\infty} \frac{\|z_i\|^2}{r_i^*} < \infty, \quad \text{a.s.} \tag{45}$$

and $\{\theta_i^*\}$ is bounded a.s., where

$$z_{n-d} = y_n - F(z)w_n - \theta_{n-d}^{*\tau}\varphi_{n-d}^* \tag{46}$$

with $F(z)$ given by (A 1) and (A 2) of Appendix A.

Proof

The proof of the lemma is given in Appendix D.

Lemma 3

If conditions (i)–(iv) hold and control is defined by (30)–(42), then

$$\frac{1}{n}\sum_{i=0}^{n} (\|y_i\|^2 + \|u_i\|^2) = O(n^\delta) \tag{47}$$

$$\|\theta - \theta_n\|^2 = O\left(\frac{(\log n)(\log\log n)^c}{n^{1-(\mu+1)(\varepsilon+\delta)}}\right), \quad \forall\, c > 1 \tag{48}$$

Proof

The proof is given in Appendix E.

Estimates (47) and (48) are the intermediate results and used only for proving Lemma 4. From Theorem 1 we shall see that (47) and (48) in fact hold for $\delta = 0$.

Lemma 4

Under the conditions of Lemma 3 there exists an integer k (possibly depending on sampling) so that

$$\tau_k < \infty, \quad \sigma_k = \infty, \quad \text{a.s.} \tag{49}$$

and the set Λ^c is finite. (50)

Proof

The proof is given in Appendix F.

This lemma tells us that after a finite number of iterations $u_n^s \equiv u_n^{(1)}$ (see (39)) and only the ELS algorithm will be used in the adaptive control system.

For Problem A we have the following theorem showing that minimal tracking error and consistent parameter estimates are achieving simultaneously.

Theorem 1

If conditions of Lemma 3 are satisfied, then

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n} (\|y_i\|^2 + \|u_i\|^2) < \infty, \quad \text{a.s.} \tag{51}$$

$$\left\| \frac{1}{n} \sum_{i=0}^{n} (y_i - y_i^*)(y_i - y_i^*)^\tau - \frac{1}{n} \sum_{i=0}^{n} (F(z)w_i)(F(z)w_i)^\tau \right\| = O(n^{-1/2\varepsilon}) \tag{52}$$

$$\|\theta_n - \theta\|^2 = O\left(\frac{(\log n)(\log \log n)^c}{n^{1-(\mu+1)\varepsilon}}\right), \quad \forall c > 1 \tag{53}$$

We note that (51) follows from (52), (7) and condition (iii) while the fact that (53) follows from (51) is proved completely in the same way as (48) from (47). So we need only to show (52). This is given in Appendix G.

Remark 1

For the case where $m < l$ if det $B(z)$ is replaced by

$$\det B(z)\begin{bmatrix} I_m \\ 0 \end{bmatrix}$$

in condition (iii), then Theorem 1 and Lemmas 3 and 4 remain valid. The proof is similar to that given by Chen and Guo (1986 b).

Remark 2

Theorem 1 and Lemmas 3 and 4 still hold true if condition (7) on the driven noise is weakened so that

$$\limsup_{n\to\infty} \frac{1}{n^{1-\varepsilon}} \sum_{i=0}^{n} w_i w_i^\tau \geqslant R_1 \tag{54}$$

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} w_i w_i^\tau \leqslant R_2 \tag{55}$$

for some $R_1 > 0$ and $R_2 > 0$, because Theorems 1–3 of Chen and Guo (1986 c) can be shown to be true under (54) and (55). In particular, if (54) holds and

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} w_i w_i^\tau = 0$$

then

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

Remark 3

In the unit delay case, we can give an adaptive control u_n that leads to (51), (53) and

$$\left\| \frac{1}{n} \sum_{i=0}^{n} (y_i - y_i^*)(y_i - y_i^*)^\tau - \frac{1}{n} \sum_{i=0}^{n} (F(z)w_i)(F(z)w_i)^\tau \right\| = O(n^{-\varepsilon}) \tag{56}$$

For Problem B we have the following Theorem which indicates the rates of convergence of the parameter estimates to their true values and of the model tracking error to its minimal value.

Theorem 2

Assume that conditions (i)–(iv) are fulfilled, control u_n is given by (30)–(42) with $u_n^{(1)}$ defined by (44) and the external input u_n^0 is $\mathscr{F}'_n$-measurable and satisfies

$$\frac{1}{n}\sum_{i=0}^{n} \|u_i^0\|^2 = O(1) \tag{57}$$

Then (51) and (53) hold and

$$\left\| \frac{1}{n}\sum_{i=0}^{n} \varepsilon_i \varepsilon_i^\tau - \frac{1}{n}\sum_{i=0}^{n} (\bar{F}(z)w_i)(\bar{F}(z)w_i)^\tau \right\| = O(n^{-1/2\varepsilon}) \tag{58}$$

where ε_i, $\bar{F}(z)$ are given in (12) and (15) respectively.

Proof

The proof is given in Appendix H.

Remark 4

We can weaken condition (7) in a way similar to Remark 2.

5. Conclusion

We have designed the optimal stochastic adaptive control for tracking both a given reference signal and a given reference model and the convergence rates have also been established. After finite steps the adaptive trackers are based on the ELS algorithm, but in the first steps we have invoked an adaptive controller based on the stochastic gradient algorithm to slow down the growth rate of

$$\sum_{i=0}^{n} \|y_i\|^2$$

Avoiding the second algorithm belongs to further research.

Appendix A

Proof of (8)

We first consider the existence and uniqueness of the solution for the Diophantine equation (9).

Let $[F_0 \quad F_1 \dots F_{d-1}]$ satisfy

$$[I \quad \bar{c}_1 I \dots \bar{c}_{d-1} I] = [F_0 \quad F_1 \dots F_{d-1}] \begin{bmatrix} I & \bar{A}_1 & \dots & \dots & \bar{A}_{d-1} \\ 0 & I & \bar{A}_1 & \dots & \bar{A}_{d-2} \\ 0 & 0 & I & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \bar{A}_1 \\ 0 & 0 & \dots & O & I \end{bmatrix} \tag{A 1}$$

where $\bar{c}_i$ $(i = 1, 2, \dots, d-1)$ are given by (22) and $\bar{A}_i$ $(i = 1, 2, \dots, d-1)$ satisfy

$$(\operatorname{Adj} C(z))A(z) = I + \bar{A}_1 z + \dots + \bar{A}_{(m-1)r+p} z^{(m-1)r+p}$$

with $\bar{A}_i = 0$, for $i \in ((m-1)r + p, d-1]$; $\bar{c}_i = 0$, for $i \in (mr, d-1]$.

By straightforward computation we know that all the coefficients of z^i, $i=0, 1, \ldots, d-1$ in

$$(\det C(z))I - F(z)(\mathrm{Adj}\ C(z))A(z)$$

are zero, where

$$F(z) = F_0 + F_1 z + \ldots + F_{d-1} z^{d-1} \tag{A 2}$$

Therefore, there exists a polynomial matrix $G(z)$ such that

$$(\det C(z))I - F(z)(\mathrm{Adj}\ C(z))A(z) = z^d G(z) \tag{A 3}$$

This implies that the Diophantine equation (9) is solvable.

The uniqueness of the solution with $\deg(F(z)) \leqslant d-1$ of (9) is clear.

We now prove (8). From (26) we have

$$y_n - y_n^* = F(z)w_n + (\det C(z))^{-1}(G(z)y_{n-d} + F(z)(\mathrm{Adj}\ C(z))B(z)u_{n-d}) - y_n^* \triangleq F(z)w_n + \bar{\mu}_{n-d}$$

Using Lemma 2 in Chen and Guo (1987 c) and noticing that $\bar{\mu}_{n-d}$ is uncorrelated with $F(z)w_n$ we see that

$$\frac{1}{n}\sum_{i=0}^{n}(y_i - y_i^*)^\tau \cdot (y_i - y_i^*) = \frac{1}{n}\sum_{i=0}^{n}\bar{\mu}_{i-d}^\tau \bar{\mu}_{i-d} + O\left(\frac{1}{n}\left(\sum_{i=0}^{n}\|\bar{\mu}_{i-d}\|^2\right)^{\hat{\mu}}\right) + \frac{1}{n}\sum_{i=0}^{n}(F(z)w_i)^\tau \cdot (F(z)w_i) \tag{A 4}$$

for any $\hat{\mu} \in (\frac{1}{2}, 1)$, and

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}\|y_i - y_i^*\|^2 &\geqslant \limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}\|F(z)w_i\|^2 \\ &= \mathrm{tr}\limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}(F(z)w_i)(F(z)w_i)^\tau \\ &= \mathrm{tr}\sum_{j=0}^{d-1}F_j R F_j^\tau \end{aligned}\tag{A 5}$$

Appendix B

Proof of (13)

Similar to (A 4) and (A 5), it follows from (43) that

$$\begin{aligned}\limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}\|\varepsilon_i\|^2 &\geqslant \limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n}\|\bar{F}(z)w_i\|^2 \\ &= \mathrm{tr}\sum_{j=0}^{d-1}\bar{F}_j R \bar{F}_j^\tau \end{aligned}\tag{B 1}$$

Appendix C

Proof of Lemma 1

The solvability of (38) is shown by Chen and Guo (1986 b) while the proofs of solvability for (28) and (44) are similar. We now show it for (28). For this it suffices to prove

$$\det B_{1n} \neq 0, \quad \text{a.s.} \tag{C 1}$$

Setting

$$e^\tau = [\overbrace{0 \quad \dots \quad 0}^{mp} \quad \overbrace{I}^{m} \quad \overbrace{0 \quad \dots \quad \dots \quad 0}^{m(r+q-1)}]\}m \tag{C 2}$$

we have

$$B_{1n} = B_{1n-1} + a_{n-1}(y_n - \theta_{n-1}^\tau \varphi_{n-1})\varphi_{n-1}^\tau P_{n-1} e \tag{C 3}$$

For $\omega \in \{\omega : \det B_{1n} = 0, \det B_{1n-1} \neq 0\}$, from (C 3) it is easy to see

$$\det (I + a_{n-1}(y_n - \theta_{n-1}^\tau \varphi_{n-1})\varphi_{n-1}^\tau P_{n-1} e B_{1n-1}^{-1}) = 0 \tag{C 4}$$

hence

$$1 + a_{n-1}\varphi_{n-1}^\tau P_{n-1} e B_{1n-1}^{-1}(y_n - \theta_{n-1}^\tau \varphi_{n-1}) = 0 \tag{C 5}$$

i.e.

$$1 + \varphi_{n-1}^\tau P_{n-1}\varphi_{n-1} + \varphi_{n-1}^\tau P_{n-1} e B_{1n-1}^{-1}(y_n - \theta_{n-1}^\tau \varphi_{n-1}) = 0 \tag{C 6}$$

Denoting

$$\xi_{n-1} = y_n - \theta_{n-1}^\tau \varphi_{n-1} - (B_1 - B_{1n-1})v_{n-1}$$

$$\tilde{\varphi}_{n-1} = \varphi_{n-1} - e v_{n-1}$$

we can rewrite (C 6) in the form

$$\begin{aligned} 1 &+ (\tilde{\varphi}_{n-1} + e v_{n-1})^\tau \cdot P_{n-1}(\tilde{\varphi}_{n-1} + e v_{n-1}) \\ &+ (\tilde{\varphi}_{n-1} + e v_{n-1})^\tau \cdot P_{n-1} e B_{1n-1}^{-1}(\xi_{n-1} + (B_1 - B_{1n-1})v_{n-1}) = 0 \end{aligned} \tag{C 7}$$

Setting

$$M_{n-1} = e^\tau \cdot P_{n-1} e B_{1n-1}^{-1} B_1$$

$$f_{n-1} = e^\tau \cdot P_{n-1} e B_{1n-1}^{-1} \xi_{n-1} + (B_1 + B_{1n-1})^\tau B_{1n-1}^{-\tau} \cdot e^\tau \cdot P_{n-1}\tilde{\varphi}_{n-1}$$

$$g_{n-1} = 1 + \tilde{\varphi}_{n-1}^\tau P_{n-1}\tilde{\varphi}_{n-1} + \tilde{\varphi}_{n-1}^\tau P_{n-1} e B_{1n-1}^{-1} \xi_{n-1}$$

it follows from (C 7) that

$$v_{n-1}^\tau M_{n-1} v_{n-1} + v_{n-1}^\tau f_{n-1} + g_{n-1} = 0 \tag{C 8}$$

By induction it is easy to conclude that M_{n-1}, f_{n-1} and g_{n-1} are independent of v_{n-1}. Hence the technique used in Lemma 4 of Chen and Guo (1986 b) applies to the present case and leads to that (C 8) cannot hold for almost all ω. Since $\det B_{10} \neq 0$, then by induction we conclude that $\det B_{1n} \neq 0$, almost surely. □

Appendix D

Proof of Lemma 2

From (26) we know that

$$\begin{aligned} (\det C(z))z_{(n-1)d+n_0} &= G(z)y_{(n-1)d+n_0} + F(z)(\text{Adj } C(z))B(z)u_{(n-1)d+n_0} \\ &\quad - (\det C(z))\theta_{(n-1)d+n_0}^{*\tau}\varphi_{(n-1)d+n_0}^{*} \end{aligned}$$

$$
\begin{aligned}
&= G(z)y_{(n-1)d+n_0} + F(z)(\mathrm{Adj}\ C(z))B(z)u_{(n-1)d+n_0} \\
&\quad - ((\det C(z)) - 1)\theta^{*\tau}_{(n-1)d+n_0}\varphi^*_{(n-1)d+n_0} - \theta^{*\tau}_{(n-1)d+n_0}\varphi^*_{(n-1)d+n_0} \\
&= \theta^{*\tau}\varphi^*_{(n-1)d+n_0} - \theta^{*\tau}_{(n-1)d+n_0}\varphi^*_{(n-1)d+n_0} \\
&= \tilde{\theta}^{*\tau}_{(n-1)d+n_0}\varphi^*_{(n-1)d+n_0}
\end{aligned}
$$

and $z_{(n-1)d+n_0}$ is $\mathscr{F}_{(n-1)d+n_0}$-measurable, where

$$
\begin{aligned}
\tilde{\theta}^*_{(n-1)d+n_0} &= \theta^* - \theta^*_{(n-1)d+n_0} \\
z_{(n-1)d+n_0} &= e_{nd+n_0} - h_{nd+n_0} \\
e_{nd+n_0} &= y_{nd+n_0} - \theta^{*\tau}_{(n-1)d+n_0}\varphi^*_{(n-1)d+n_0} \\
h_{nd+n_0} &= F(z)w_{nd+n_0}
\end{aligned}
$$

Hence by strictly positive realness of $(\det C(z)) - (\bar{a}/2)$ there are $\rho > 0$, and $\rho_0 \geqslant 0$ such that

$$
S_{(n-1)d+n_0} = 2\bar{a}\sum_{i=0}^{n} z^{\tau}_{(i-1)d+n_0}\left(\tilde{\theta}^{*\tau}_{(i-1)d+n_0}\varphi^*_{(i-1)d+n_0} - \frac{\bar{a}+\rho}{2} z_{(i-1)d+n_0}\right) + \rho_0 \geqslant 0.
$$

Noticing $S_{(n-1)d+n_0}$ is $\mathscr{F}_{(n-1)d+n_0}$-measurable and

$$
\begin{aligned}
&2\bar{a}z^{\tau}_{(n-1)d+n_0}\theta^*_{(n-1)d+n_0}\varphi^*_{(n-1)d+n_0} - \bar{a}^2\|z_{(n-1)d+n_0}\|^2 \\
&\qquad = S_{(n-1)d+n_0} - S_{(n-2)d+n_0} + \rho\bar{a}\|z_{(n-1)d+n_0}\|^2
\end{aligned}
$$

by (34) we find that

$$
\begin{aligned}
\mathrm{tr}\ \tilde{\theta}^{*\tau}_{nd+n_0}\tilde{\theta}^*_{nd+n_0} &= \mathrm{tr}\ \tilde{\theta}^{*\tau}_{(n-1)d+n_0}\tilde{\theta}^*_{(n-1)d+n_0} \\
&\quad - \frac{2\bar{a}}{r^*_{(n-1)d+n_0}}\mathrm{tr}\ \tilde{\theta}^{*\tau}_{(n-1)d+n_0}\varphi^*_{(n-1)d+n_0}e^{\tau}_{nd+n_0} \\
&\quad + \frac{\bar{a}^2}{r^{*2}_{(n-1)d+n_0}}\|\varphi^{*\tau}_{(n-1)d+n_0}\|^2\cdot\|e_{nd+n_0}\|^2 \\
&= \mathrm{tr}\ \tilde{\theta}^{*\tau}_{(n-1)d+n_0}\tilde{\theta}^*_{(n-1)d+n_0} \\
&\quad - \frac{2\bar{a}}{r^*_{(n-1)d+n_0}}\mathrm{tr}\ \tilde{\theta}^{*\tau}_{(n-1)d+n_0}\varphi^*_{(n-1)d+n_0}z^{\tau}_{(n-1)d+n_0} \\
&\quad - \frac{2\bar{a}}{r^*_{(n-1)d+n_0}}\mathrm{tr}\ \tilde{\theta}^{*\tau}_{(n-1)d+n_0}\varphi^*_{(n-1)d+n_0}h^{\tau}_{nd+n_0} \\
&\quad + \frac{2\bar{a}^2}{r^{*2}_{(n-1)d+n_0}}\|\varphi^*_{(n-1)d+n_0}\|^2 z^{\tau}_{(n-1)d+n_0}h_{nd+n_0} \\
&\quad + \frac{\bar{a}^2}{r^{*2}_{(n-1)d+n_0}}\|\varphi^*_{(n-1)d+n_0}\|^2\cdot\|z_{(n-1)d+n_0}\|^2 \\
&\quad + \frac{\bar{a}^2}{r^{*2}_{(n-1)d+n_0}}\|\varphi^*_{(n-1)d+n_0}\|^2\cdot\|h_{nd+n_0}\|^2
\end{aligned}
$$

Taking conditional expectation we conclude that

$$\begin{aligned}E[\operatorname{tr} \tilde{\theta}^{*\tau}_{nd+n_0}\tilde{\theta}^{*}_{nd+n_0} | \mathscr{F}_{(n-1)d+n_0}] &= \operatorname{tr} \tilde{\theta}^{*\tau}_{(n-1)d+n_0}\tilde{\theta}^{*}_{(n-1)d+n_0} \\ &\quad - \frac{2\bar{a}}{r^{*}_{(n-1)d+n_0}} \operatorname{tr} \tilde{\theta}^{*\tau}_{(n-1)d+n_0} \varphi^{*}_{(n-1)d+n_0} z^{\tau}_{(n-1)d+n_0} \\ &\quad + \frac{\bar{a}^2}{r^{*2}_{(n-1)d+n_0}} \|\varphi^{*}_{(n-1)d+n_0}\|^2 \cdot \|z_{(n-1)d+n_0}\|^2 \\ &\quad + \frac{\bar{a}^2}{r^{*2}_{(n-1)d+n_0}} \|\varphi^{*}_{(n-1)d+n_0}\|^2 \cdot \gamma \cdot \operatorname{tr} \sum_{j=0}^{d-1} F_j \cdot F_j^{\tau} \\ &\leqslant \operatorname{tr} \tilde{\theta}^{*\tau}_{(n-1)d+n_0}\tilde{\theta}^{*}_{(n-1)d+n_0} - \frac{S_{(n-1)d+n_0}}{r^{*}_{(n-1)d+n_0}} + \frac{S_{(n-2)d+n_0}}{r^{*}_{(n-1)d+n_0}} \\ &\quad - \frac{\bar{a}\rho}{r^{*}_{(n-1)d+n_0}} \|z_{(n-1)d+n_0}\|^2 \\ &\quad + \frac{\bar{a}^2}{r^{*2}_{(n-1)d+n_0}} \|\varphi^{*}_{(n-1)d+n_0}\|^2 \cdot \gamma \cdot \operatorname{tr} \sum_{j=0}^{d-1} F_j \cdot F_j^{\tau}\end{aligned}$$

and hence

$$\begin{aligned}E\left[\operatorname{tr} \tilde{\theta}^{*\tau}_{nd+n_0}\tilde{\theta}^{*}_{nd+n_0} + \frac{S_{(n-1)d+n_0}}{r^{*}_{(n-1)d+n_0}} \middle| \mathscr{F}_{(n-1)d+n_0}\right] &\leqslant \operatorname{tr} \tilde{\theta}^{*\tau}_{(n-1)d+n_0}\tilde{\theta}^{*}_{(n-1)d+n_0} + \frac{S_{(n-2)d+n_0}}{r^{*}_{(n-2)d+n_0}} - \frac{\bar{a}\rho}{r^{*}_{(n-1)d+n_0}} \|z_{(n-1)d+n_0}\|^2 \\ &\quad + \frac{\bar{a}^2}{r^{*2}_{(n-1)d+n_0}} \|\varphi^{*}_{(n-1)d+n_0}\|^2 \cdot \gamma \cdot \operatorname{tr} \sum_{j=0}^{d-1} F_j \cdot F_j^{\tau}\end{aligned}$$

It is easy to see that

$$\sum_{i=0}^{\infty} \frac{\|\varphi^{*}_{(n-1)d+n_0}\|^2}{r^{*2}_{(n-1)d+n_0}} < \infty, \quad \text{for any} \quad n_0 = 0, 1, \ldots, d-1$$

Then by the convergence theorem of supermartingales we find that $\operatorname{tr} \tilde{\theta}^{*\tau}_{nd+n_0}\tilde{\theta}^{*}_{nd+n_0}$ tends to a finite limit and hence

$$\sum_{i=0}^{\infty} \frac{\|z_{(n-1)d+n_0}\|^2}{r^{*}_{(n-1)d+n_0}} < \infty, \quad \text{for any} \quad n_0 = 0, 1, \ldots, d-1 \qquad \square$$

Appendix E

Proof of Lemma 3

For (47) by stability of $B(z)$ and (7) it suffices to show

$$\frac{1}{n} \sum_{i=0}^{n} \|y_i\|^2 = O(n^{\delta}) \tag{E 1}$$

(i) If $\tau_k < \infty$, $\sigma_k = \infty$ for some k, then (E 1) follows from (30).

(ii) If $\sigma_k < \infty$, $\tau_{k+1} = \infty$ for some k, then by (38) for $n \geqslant \sigma_k$ we have

$$\theta^{*\tau}_n \varphi^{*}_n = y^{*}_{n+d} + D^{*}_{on} v_n, \tag{E 2}$$

which combining with (46) yields

$$\frac{1}{n}\sum_{i=0}^{n+d}\|y_i\|^2 = O\left(\frac{1}{n}\sum_{i=0}^{n}\|z_i\|^2 + 1\right) \tag{E 3}$$

Noticing $q \geqslant 1$ and

$$r_n^* \geqslant \sum_{i=0}^{n}\|u_i\|^2 = \sum_{i=0}^{n}\|u_i^s\|^2 + O\left(\left(\sum_{i=0}^{n}\|u_i^s\|^2\right)^{\hat{\mu}}\right) + \sum_{i=0}^{n}\|v_i\|^2 \xrightarrow[n\to\infty]{} \infty, \quad \forall\, \hat{\mu} \in (\tfrac{1}{2}, 1)$$

from (45) we know

$$\sum_{i=0}^{n}\|z_i\|^2 = o(r_n^*) \tag{E 4}$$

Further, by (E2) we have

$$r_n^* = O\left(\sum_{i=0}^{n+d}\|y_i\|^2 + \sum_{i=0}^{n-1}\|y_{i+d}^* + D_{0i}^* v_i\|^2 + n\right) = O\left(\sum_{i=0}^{n+d}\|y_i\|^2 + n\right) \tag{E 5}$$

Combining (E 3)–(E 5) leads to

$$\sum_{i=0}^{n}\|z_i\|^2 = o\left(\sum_{i=0}^{n+d}\|y_i\|^2 + n\right) \quad \text{or} \quad \frac{1}{n}\sum_{i=0}^{n}\|z_i\|^2 = o\left(\frac{1}{n}\sum_{i=0}^{n+d}\|y_i\|^2 + 1\right) \tag{E 6}$$

From this and (E 3) we have

$$\frac{1}{n}\sum_{i=0}^{n+d}\|y_i\|^2 = O(1) \tag{E 7}$$

which implies (E 1).

(iii) If $\sigma_k < \infty$ and $\tau_k < \infty$ for all k, then for sufficiently large k we have

$$\begin{aligned}
&\sup_{\tau_k \leqslant n < \sigma_k} \frac{1}{n^{1+\delta}} \sum_{i=0}^{n}\|y_i\|^2 \\
&\leqslant \sup_{\tau_k \leqslant n < \sigma_k} \frac{1}{n^{1+\delta}} \left(\sum_{i=\tau_1}^{\sigma_1 - 1}\|y_i\|^2 + \sum_{i=\sigma_1}^{\tau_2}\|y_i\|^2 + \dots + \sum_{i=\sigma_{k-1}}^{\tau_k}\|y_i\|^2 + \sum_{i=\tau_k}^{n}\|y_i\|^2\right) \\
&\leqslant \sup_{\tau_k \leqslant n < \sigma_k} \frac{1}{n^{1+\delta}} \left(2\sum_{i=1}^{k-1}\frac{\tau_{i+1}\log\tau_{i+1}}{2^i} + \sum_{i=\tau_k}^{n}\|y_i\|^2\right) \\
&\leqslant 2 + \sup_{\tau_k \leqslant n < \sigma_k} \frac{1}{n^{1+\delta}} \sum_{i=\tau_k}^{n}\|y_i\|^2 \\
&\leqslant 2 + \sup_{\tau_k \leqslant n < \sigma_k} \frac{1}{n^{1+\sigma}} (n^{1+\delta} + \|y_{\tau_k}\|^2) \\
&\leqslant 3 + \frac{\tau_k \log \tau_k}{\tau_k^{1+\delta}} < 4
\end{aligned} \tag{E 8}$$

Similarly, we can show

$$\sum_{i=0}^{\tau_k} \|u_i\|^2 \leqslant 3\tau_k \log \tau_k, \quad \sum_{i=0}^{\tau_k} \|y_i\|^2 \leqslant 3\tau_k \log \tau_k \tag{E 9}$$

Hence, for (E 1) it suffices to prove

$$\sup_{\sigma_k \leqslant n < \tau_{k+1}} \frac{1}{n^{1+\delta}} \sum_{i=0}^{n} \|y_i\|^2 = O(1) \tag{E 10}$$

When $n \in [\sigma_k, \sigma_k + d - 1] \cap [\sigma_k, \tau_{k+1})$, by (E 8) we have

$$\begin{aligned}\sum_{i=0}^{n} \|y_i\|^2 &= \sum_{i=0}^{\sigma_k - 1} \|y_i\|^2 + \sum_{i=\sigma_k}^{n} \|y_i\|^2 = O(\sigma_k^{1+\delta}) + \sum_{i=\sigma_k}^{n} \|y_i\|^2 \\ &= O(n^{1+\delta}) + \sum_{i=\sigma_k}^{n} \|y_i\|^2 \end{aligned} \tag{E 11}$$

Using (26) as $k \to \infty$ we see

$$\begin{aligned}\sum_{i=\sigma_k}^{n} \|y_i\|^2 &= O(n) + O\Bigg(\sum_{i=\sigma_k}^{n} \|(\det C(z))^{-1} G(z) y_{i-d} \\ &\qquad + (\det C(z))^{-1} F(z)(\operatorname{Adj} C(z)) B(z) u_{i-d}\|^2\Bigg) \\ &= O(n) + O\left(\sum_{i=0}^{n-d} \|y_i\|^2 + \sum_{i=0}^{n-d} \|u_i\|^2\right)\end{aligned} \tag{E 12}$$

(*a*) When $\tau_k \geqslant n - d$, then by (E 9)

$$\sum_{i=0}^{n-d} \|u_i\|^2 \leqslant \sum_{i=0}^{\tau_k} \|u_i\|^2 = O(\tau_k^{1+\delta}) = O(n^{1+\delta}) \tag{E 13}$$

(*b*) When $n - d > \tau_k \geqslant \sigma_k - d$, we have $\sigma_k - 1 \geqslant n - d$, hence $n - d - \tau_k + 1 \leqslant \sigma_k - 1 - \tau_k + 1 = \sigma_k - \tau_k \leqslant d$ and by (39) and (40)

$$\begin{aligned}\sum_{i=0}^{n-d} \|u_i\|^2 &= \sum_{i=0}^{\tau_k} \|u_i\|^2 + \sum_{i=\tau_k+1}^{n-d} \|u_i\|^2 \\ &\leqslant \sum_{i=0}^{\tau_k} \|u_i\|^2 + d \cdot n^{1+\delta} = O(n^{1+\delta})\end{aligned} \tag{E 14}$$

(*c*) When $\tau_k < \sigma_k - d$, then

$$\begin{aligned}\sum_{i=0}^{n-d} \|u_i\|^2 &= \sum_{i=0}^{\sigma_k - d - 1} \|u_i\|^2 + \sum_{i=\sigma_k - d}^{n-d} \|u_i\|^2 \\ &= O\left(\sum_{i=0}^{\sigma_k - 1} \|y_i\|^2 + \sigma_k\right) + d \cdot n^{1+\delta} = O(n^{1+\delta})\end{aligned} \tag{E 15}$$

(E 12)–(E 15) show that

$$\sum_{i=0}^{n} \|y_i\|^2 = O(n^{1+\delta}), \quad \text{for} \quad n \in [\sigma_k, \sigma_k + d - 1] \cap [\sigma_k, \tau_{k+1})$$

For $n \in [\sigma_k + d, \tau_{k+1})$ from (46), (E 2), (39) and (E 6) we have

$$\sum_{i=0}^{n} \|y_i\|^2 = \sum_{i=0}^{\sigma_k + d - 1} \|y_i\|^2 + \sum_{i=\sigma_k + d}^{n} \|y_i\|^2$$

$$= O(n^{1+\delta}) + \sum_{i=\sigma_k+d}^{n} \|y_i^* + F(z)w_i + D_{0i-d}^* v_{i-d} + z_{i-d}\|^2$$

$$= O(n^{1+\delta}) + O\left(\sum_{i=0}^{n-d} \|z_i\|^2\right) = O(n^{1+\delta}) + o\left(\sum_{i=0}^{n} \|y_i\|^2\right)$$

which means that

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

Hence (E 10) and consequently, (E 1) are verified. □

Convergence rate (48) is proved in the same way as that used for Theorem 3 of Chen and Guo (1986 c).

Appendix F

Proof of Lemma 4

For (49) it suffices to prove the impossibility of the following situations:

(*a*) $\sigma_k < \infty$, $\tau_{k+1} = \infty$, for some positive integer k;

(*b*) $\tau_k < \infty$, $\sigma_k < \infty$, for every positive integer k.

We first show the impossibility of (*a*). If (*a*) were true, then from the stability of $B(z)$ and (ii) of the proof of Lemma 3 it follows immediately that

$$\sum_{i=0}^{n} \|y_i\|^2 = O(n) \quad \text{and} \quad \sum_{i=0}^{n} \|u_i\|^2 = O(n)$$

which contradicts $\tau_{k+1} = \infty$ (see 31)).

We now prove the impossibility of (*b*). Setting

$$t_k = \sup \{n : j \in [\tau_k, \sigma_k) \cap \Lambda, \quad \forall\, j \in [\tau_k, n]\} \tag{F 1}$$

and noticing (28), (48), (E 9) and condition (iii) we find that

$$\|u_{\tau_k}^{(1)}\|^2 = O(\tau_k \log \tau_k)$$

and hence, for sufficiently large k, the t_k given by (F 1) exists.

Combining (E 9) with (26) we have

$$\sum_{i=0}^{\tau_k+d} \|y_i\|^2 = O(\tau_k \log \tau_k) \tag{F 2}$$

which implies that

$$\sum_{i=0}^{\tau_k+d} \|(\det C(z)) y_i\|^2 = O(\tau_k \log \tau_k) \tag{F 3}$$

Setting

$$\varphi_n^0 = [y_n^\tau \quad \dots \quad y_{n-p+1}^\tau \quad u_{n-d+1}^\tau \quad \dots \quad u_{n-q-d+2}^\tau \quad w_n^\tau \quad \dots \quad w_{n-r+1}^\tau]^\tau$$

we find

$$y_{n+d} = \theta^\tau \varphi_{n+d-1}^0 + w_{n+d} = \tilde{\theta}_n^\tau \varphi_{n+d-1}^0 + \theta_n^\tau \varphi_{n+d-1}^0 + w_{n+d}$$

or

$$\tilde{\theta}_n^{\tau} \varphi_{n+d-1}^0 + C_n(z) w_{n+d} = A_n(z) y_{n+d} - B_n(z) u_n$$

Then from this and (28) it follows that for every $n \in [\tau_k, t_k]$

$$\begin{aligned}
&(F_n(z)(\text{Adj } C_n(z)))(\tilde{\theta}_n^{\tau} \varphi_{n+d-1}^0 + C_n(z) w_{n+d}) \\
&\quad = ((F_n(\text{Adj } C_n))A)_n(z) y_{n+d} - ((F_n(\text{Adj } C_n))B)_n(z) u_n \\
&\quad = ((\det C_n(z))I - G_n(z) z^d) y_{n+d} - ((F_n(\text{Adj } C_n))B)_n(z) u_n \\
&\qquad + (((F_n(\text{Adj } C_n))A)_n(z) - (F_n(\text{Adj } C_n) A_n)(z)) y_{n+d} \\
&\quad = (\det C_n(z)) y_{n+d} - (\det C_n(z)) y_{n+d}^* - B_{1n} v_n \\
&\qquad + (((F_n(\text{Adj } C_n))A)_n(z) - (F_n(\text{Adj } C_n) A_n)(z)) y_{n+d} \\
&\qquad + ((F_n(\text{Adj } C_n) B_n)(z) - ((F_n(\text{Adj } C_n))B)_n(z)) u_n
\end{aligned} \tag{F 4}$$

where for any polynomial matrices $A_n(z)$, $B_n(z)$ and $F_n(z)$, we set

$$((F_n A_n)B)_n(z) = \sum_{i,j,k} F_{in} A_{jn} B_{kn-i-j} z^{i+j+k}$$

$$(F_n A_n B_n)(z) = \sum_{i,j,k} F_{in} A_{jn} B_{kn} z^{i+j+k}$$

From (F 4) we have

$$\begin{aligned}
(\det C(z)) y_{n+d} &= (F_n(z) \text{Adj } C_n(z)))(\tilde{\theta}_n^{\tau} \varphi_{n+d-1}^0 + C_n(z) w_{n+d}) \\
&\quad + (\det C(z) - \det C_n(z)) y_{n+d} + (\det C_n(z)) y_{n+d}^* + B_{1n} v_n \\
&\quad - (((F_n(\text{Adj } C_n))A)_n(z) - (F_n(\text{Adj } C_n) A_n)(z)) y_{n+d} \\
&\quad - ((F_n(\text{Adj } C_n) B_n)(z) - ((F_n(\text{Adj } C_n))B)_n(z)) u_n
\end{aligned} \tag{F 5}$$

Setting

$$S_i = \sum_{j=1}^{i} (\|\varphi_j^0\|^2 + \|y_j\|^2), \quad \text{for} \quad i \geqslant 1$$

$$S_i = 0, \quad \text{for} \quad i < 1$$

and noticing (47) we have

$$S_i = O(i^{1+\delta}) \tag{F 6}$$

Comparing (29) and (9), similar to (A 1) we know that F_{in} $(i = 0, 1, \ldots, d-1)$ is a finite product of coefficients of $(\det C_n(z))I$ and $(\text{Adj } C_n(z))A_n(z)$. Thus from (47), (48), (F 6) and (F 5) it follows that there exists a constant n_0 so that

$$\sum_{i=\tau_k}^{t_k} \|\det C(z)) y_{i+d}\|^2 = O(n) + O\left(\sum_{j=-d}^{n_0} \sum_{i=1}^{t_k} \frac{\log^2 i}{i^{\alpha}} (\|\varphi_{i-j}^0\|^2 + \|y_{i-j}\|^2) \right) \tag{F 7}$$

where $\alpha = 1 - (\mu + 1)(\varepsilon + \delta)$.

We now show that

$$\sum_{i=1}^{n} \frac{\log^2 i}{i^{\alpha}} (\|\varphi_{i-j}^0\|^2 + \|y_{i-j}\|^2) = O(n^{1-\varepsilon}), \quad \text{for any} \quad j \in [-d, n_0] \tag{F 8}$$

Noticing

$$(\log^2 i)(1+i)^\alpha - i^\alpha \log^2 (i+1) = i^\alpha(\log^2 i)\left(\left(1+\frac{1}{i}\right)^\alpha - \left(1+\frac{\log\left(1+\frac{1}{i}\right)}{\log i}\right)^2\right)$$
$$= i^\alpha(\log^2 i)\left(\frac{\alpha}{i}+o\left(\frac{1}{i}\right)\right)$$

we have

$$\sum_{i=1}^{n}\frac{\log^2 i}{i^\alpha}(\|\varphi^0_{i-j}\|^2+\|y_{i-j}\|^2) = \frac{\log^2 n}{n^\alpha}S_{n-j}+\sum_{i=1}^{n-1}\left(\frac{\log^2 i}{i^\alpha}-\frac{\log^2(i+1)}{(i+1)^\alpha}\right)S_{i-j}$$
$$= O\left(\frac{\log^2 n}{n^\alpha}n^{1+\delta}\right)+O\left(\sum_{i=1}^{n-1}\frac{\log^2 i}{i(i+1)^\alpha}i^{1+\delta}\right)$$
$$= O(n^{1-\varepsilon}), \quad \text{for any} \quad j\in[-d,n_0]$$

since $(\mu+1)(\varepsilon+\delta)+\frac{1}{2}\delta<(\mu+1)(\varepsilon+(1-2\varepsilon(\mu+1))/(2\mu+3))+(1-2\varepsilon(\mu+1))/(2(2\mu+3))=\frac{1}{2}$ and $\mu\geqslant 1$, we see $\alpha-\delta>\varepsilon$.

From (F 8), (F 7) and (F 3) we have

$$\sum_{i=0}^{t_k}\|(\det C(z))y_{i+d}\|^2 = O(t_k\log t_k)$$

and hence

$$\sum_{i=0}^{t_k}\|y_{i+d}\|^2 = O(t_k\log t_k) \tag{F 9}$$

since $\det C(z)$ is stable, consequently, it follows from (28) and (48) that

$$\|u^{(1)}_{t_k+1}\|^2 = O\left(t_k+\sum_{i=0}^{t_k+1}\|y_i\|^2+\sum_{i=0}^{t_k}\|u_i\|^2\right)$$
$$= O\left(t_k+\sum_{i=0}^{t_k}\|y_{i+d}\|^2\right) = O(t_k\log t_k)$$

Combining this with (F 1) we see

$$t_k=\sigma_k-1 \tag{F 10}$$

and which by (F 9) implies

$$\sum_{i=0}^{\sigma_k-1+d}\|y_i\|^2 = O(\sigma_k\log\sigma_k) \tag{F 11}$$

and hence

$$\sum_{i=\tau_k}^{\sigma_k}\|y_i\|^2 \leqslant \sigma_k^{1+\frac{1}{2}\delta}+\|y_{\tau_k}\|^2$$

which contradicts (30). Thus the proof of (49) is completed, and (50) follows from (F 1), (F 10) and (49) immediately. □

Appendix G

Proof of (52)

From (F 5) we see that for $n \geqslant \tau_k$

$$\begin{aligned} y_{n+d} - y^*_{n+d} = (\det C(z))^{-1}\{&(F_n(z)(\operatorname{Adj} C_n(z)))\tilde{\theta}^\tau_n \varphi^0_{n+d-1} + B_{1n} v_n \\ &+ (((F_n(z)(\operatorname{Adj} C_n(z)))C)_n(z) - F(z)(\operatorname{Adj} C(z))C(z))w_{n+d} \\ &+ (\det C(z) - \det C_n(z))(y_{n+d} - y^*_{n+d}) \\ &- (((F_n(\operatorname{Adj} C_n))A)_n(z) - (F_n(\operatorname{Adj} C_n)A_n)(z))y_{n+d} \\ &- ((F_n(\operatorname{Adj} C_n)B_n)(z) - ((F_n(\operatorname{Adj} C_n))B)_n(z))u_n\} \\ &+ F(z)w_{n+d} \\ &\triangleq F(z)w_{n+d} + \mu_n \end{aligned} \tag{G 1}$$

and thus we have

$$\frac{1}{n}\sum_{i=0}^{n}(y_i - y^*_i)(y_i - y^*_i)^\tau - \frac{1}{n}\sum_{i=0}^{n}(F(z)w_i)(F(z)w_i)^\tau$$

$$= \frac{1}{n}\sum_{i=0}^{n}\mu_{i-d}\mu^\tau_{i-d} + \frac{1}{n}\sum_{i=0}^{n}(F(z)w_i)\mu^\tau_{i-d} + \frac{1}{n}\sum_{i=0}^{n}\mu_{i-d}(F(z)w_i)^\tau \tag{G 2}$$

Similar to (F 7) and (F 8) we can obtain

$$\sum_{i=0}^{n}\|\mu_i\|^2 = O\left(\sum_{j=-d}^{n_0}\sum_{i=1}^{n}\frac{\log^2 i}{i^\alpha}(\|\varphi^0_{i-j}\|^2 + \|y_{i-j}\|^2 + \|w_{i-j}\|^2 + 1)\right) + O\left(\sum_{i=1}^{n}\|v_i\|^2\right)$$

$$= O(n^{1-\varepsilon})$$

and hence

$$\left\|\frac{1}{n}\sum_{i=0}^{n}\mu_{i-d}(F(z)w_i)^\tau\right\| = \left\|\frac{1}{n}\sum_{i=0}^{n}(F(z)w_i)\mu^\tau_{i-d}\right\|$$

$$= O\left(\left(\frac{1}{n}\sum_{i=0}^{n}\|\mu_i\|^2\right)^{1/2}\right) = O(n^{-1/2\varepsilon}).$$

From this and (G 2), (52) follows immediately. Thus the proof is completed. □

Appendix H

Proof of Theorem 2

Similar to (G 1), by using (44) it is not difficult to show that for $n \in [\tau_k, \sigma_k)$

$$\begin{aligned} &y_{n+d} - (A^0(z))^{-1}(B^0(z)u^0_n - \bar{N}_n(z)\xi_n - \bar{N}_n(z)w_n) \\ &= (\det C(z))^{-1}\{(F_n(z)(\operatorname{Adj} C_n(z))\tilde{\theta}^\tau_n \varphi^0_{n+d-1} + B_{1n}v_n \\ &\quad + (\det C(z) - \det C_n(z))(y_{n+d} - (A^0(z))^{-1}(B^0(z)u^0_n - \bar{N}_n(z)\xi_n - \bar{N}_n(z)w_n)) \\ &\quad + (((F_n(\operatorname{Adj} C_n))C)_n(z) - (\det C(z))F(z))w_{n+d} \\ &\quad - (((F_n(\operatorname{Adj} C_n))A)_n(z) - (F_n(\operatorname{Adj} C_n)A_n)(z))y_{n+d} \\ &\quad - ((F_n(\operatorname{Adj} C_n)B_n)(z) - ((F_n(\operatorname{Adj} C_n))B)_n(z))u_n\} \\ &\quad + F(z)w_{n+d} \end{aligned} \tag{H 1}$$

where $\xi_n = y_n - \theta_n^\tau \varphi_{n-1} - w_n$ and condition (ii) implies (Chen and Guo 1986 c) that

$$\sum_{i=0}^{n} \|\xi_i\|^2 = O((\log r_n)(\log\log r_n)^c), \quad \forall c > 1 \tag{H 2}$$

where

$$r_n = 1 + \sum_{i=0}^{n} \|\varphi_i\|^2$$

From (H 2), (47) and (48), we can see that

$$\sum_{i=0}^{n} \|\xi_i\|^2 = O(\log^2 n). \tag{H 3}$$

Combining this with (H 1), similar to the proof of Lemma 4, we can conclude that there exists a positive integer k so that $\tau_k < \infty$, $\sigma_k = \infty$, almost surely and Λ^c is finite. Therefore, from (H 1) it follows that for $n \geqslant \tau_k + \bar{p}$,

$$\begin{aligned}
&A^0(z)y_{n+d} - B^0(z)u_n^0 \\
&\quad = \bar{F}(z)w_{n+d} - \bar{N}_n(z)\xi_n + (\bar{N}(z) - \bar{N}_n(z))w_n \\
&\qquad + (\det C(z))^{-1}A^0(z)\{(F_n(z)(\mathrm{Adj}\ C_n(z)))\tilde{\theta}_n^\tau \varphi_{n+d-1}^0 + B_{1n}v_n \\
&\qquad + (\det C(z) - \det C_n(z))(y_{n+d} - (A^0(z))^{-1}(B^0(z)u_n^0 - \bar{N}_n(z)\xi_n - \bar{N}_n(z)w_n)) \\
&\qquad + (((F_n(z)(\mathrm{Adj}\ C_n(z)))C)_n(z) - (\det C(z))F(z))w_{n+d} \\
&\qquad - (((F_n(\mathrm{Adj}\ C_n))A)_n(z) - (F_n(\mathrm{Adj}\ C_n)A_n)(z))y_{n+d} \\
&\qquad - ((F_n(\mathrm{Adj}\ C_n)B_n)(z) - ((F_n(\mathrm{Adj}\ C_n))B)_n(z))u_n\}
\end{aligned}$$

From this and noticing $\tau_k < \infty$, $\sigma_k = \infty$, almost surely, similar to the proof of Theorem 1, we can show that (51), (53) and (58) are true. Thus the proof is completed. □

References

ÅSTRÖM, K. J., and WITTENMARK, B., 1973, On self-tuning regulators. *Automatica*, 9, 185–199.

CAINES, P. E., and LAFORTUNE, S., 1984, Adaptive control with recursive identification for stochastic linear systems. *I.E.E.E. Transactions on Automatic Control*, **29**, 312–321.

CHEN, H. F., 1984, Recursive system identification and adaptive control by use of the modified least squares algorithm. *S.I.A.M. Journal Control and Optimization*, **22**, 758–776.

CHEN, H. F., and CAINES, P. E., 1985, The strong consistency of the stochastic gradient algorithm of adaptive control. *I.E.E.E. Transactions on Automatic Control*, **30**, 189–192.

CHEN, H. F., and GUO, L., 1987 a, Asymptotically optimal adaptive control with consistent parameter estimates. *S.I.A.M. Journal of Control and Optimization*, **25**, 558–575; 1987 b, Adaptive control with recursive identification for stochastic linear systems. *Advances in Control and Dynamic Systems*, **26**, 277–327; 1987 c, Optimal adaptive control and consistent parameter estimates for ARMAX model with quadratic cost. *S.I.A.M. Journal of Control and Optimization*, **25**, 845–967; 1986 a, Stochastic adaptive pole-zero assignment with convergence analysis. *Systems and Control Letters*, **7**, 159–164; 1986 b, Strong consistency of parameter estimates in optimal stochastic adaptive tracking systems. *Scientia Sinica (Series A)*, **29**, 1145–1156; 1986 c, Convergence rate of least-squares identification and adaptive control for stochastic systems. *International Journal of Control*, **44**, 1459–1476.

GOODWIN, G. C., RAMADGE, P. J., and CAINES, P. E., 1981, Discrete time stochastic adaptive control. *S.I.A.M. Journal of Control and Optimization*, **19**, 829–853.

Goodwin, G. C., and Sin, K. S., 1979, Stochastic adaptive control: The general delay-colored noise case. Technical Report 7904, Department of Electrical Engineering, University of Newcastle, Australia.

Guo, L., and Chen, H. F., 1987, Convergence rate of ELS-based adaptive tracker. *Technical Report*, Institute of Systems Science, Academia Sinica, P.R.C.

Kumar, P. R., 1985, A survey of some results in stochastic adaptive control. *S.I.A.M. Journal of Control and Optimization*, **23**, 329–380.

Lai, T. L., and Wei, C. Z., 1986, Extended least-squares and their applications to adaptive control and prediction in linear systems. *I.E.E.E. Transactions on Automatic Control*, **31**, 898–906.