International Journal of Control

ISSN: 0020-7179 (Print) 1366-5820 (Online) Journal homepage: https://www.tandfonline.com/loi/tcon20

Adaptive control via consistent estimation for deterministic systems

HAN-FU CHEN & LEI GUO

To cite this article: HAN-FU CHEN & LEI GUO (1987) Adaptive control via consistent estimation for deterministic systems, International Journal of Control, 45:6, 2183-2202, DOI: 10.1080/00207178708933875

To link to this article: https://doi.org/10.1080/00207178708933875

Published online: 16 May 2007.

INT. J. CONTROL, 1987, VOL. 45, NO. 6, 2183–2202

Adaptive control via consistent estimation for deterministic systems

HAN-FU CHEN† and LEI GUO†

For multidimensional discrete-time deterministic systems the optimal adaptive control has been derived by use of a probabilistic method so that when the reference signal is an arbitrary bounded random sequence, the tracking error and the estimation error based on a projection algorithm go to zero with a near-exponential convergence rate. For this, the basic step is to prove the consistency of estimates when the condition number of the matrix consisting of regressors diverges to infinity; in other words, when the persistent excitation condition is not satisfied.

1. Introduction

For a linear deterministic system with known parameter one can easily define the optimal control depending on the system coefficients in order that the output of the system tracks a given reference signal. Controlling systems with unknown parameter is the purpose of adaptive control. Considerable progress has been made in recent years (see e.g. Åström 1984, Anderson *et al.* 1986, Anderson and Johnson 1982, Bitmead 1984, Goodwin *et al.* 1980, Goodwin and Sin 1984, Kosut *et al.* 1985), but to the authors' knowledge the problem of simultaneously determining optimal adaptive control and consistent parameter estimates is still open. This can be explained as follows. In the analysis of the existing recursive algorithms estimating unknown coefficients of a linear deterministic system, for convergence of the estimates to the true values the persistent excitation (PE) condition is normally required, meaning that the ratio of the maximum to the minimum eigenvalue of the matrix $\sum_{i=1}^{n} \varphi_i \varphi_i^\tau$ is bounded as $n \to \infty$, where φ_i is the regression vector with components being the input–output data of the system. However, the PE condition is not usually satisfied for systems with unknown coefficients and with adaptive control derived from the optimal control with system coefficients replaced by their estimates. Hence the convergence of parameter estimates is not guaranteed, and as a result, the adaptive control obtained in this way may be far from the optimal one.

To overcome this difficulty, in the consideration of stochastic linear systems (Chen and Guo 1985 a, b, 1986 b) some consistency results on parameter estimation were first established under a condition allowing the ratio of the maximum to the minimum eigenvalue of $\sum_{i=1}^{n} \varphi_i \varphi_i^\tau$ to diverge to a certain extent. It was then shown that this condition was met when a sequence of independent random vectors with covariance matrices tending to zero was introduced to disturb the adaptive control. Since an attenuating dither cannot change the long-run average-type loss function, it is thus possible to derive simultaneously the optimal adaptive control and the consistent parameter estimates.

In this paper we consider the linear discrete-time deterministic system with

Received 21 May 1986. Revised 2 September 1986.
† Institute of Systems Science, Academia Sinica, Beijing, People's Republic of China.

unknown coefficients, which are estimated by the projection algorithm. We prove the convergence of the parameter estimates with a near-exponential rate, if the input satisfies a condition which is shown to hold when the attenuating excitation technique mentioned above is applied to the control. The proof is essentially based on estimation for the random matrix sums truncated at stopping times. Then, an adaptive tracking control is defined such that the parameter estimation error converges to zero with a near-exponential rate, and the tracking error between the system output and a given bounded reference sequence also goes to zero with a rate of convergence which we can also indicate.

2. Parameter identification

Let the l-input m-output system be described by

$$A(z)y_n = z^d B(z)u_n, \quad d \geqslant 1 \tag{1}$$

with unknown matrix coefficient θ

$$\theta^\tau = [-A_1 \quad \dots \quad -A_p \quad B_1 \quad \dots \quad B_q]$$

in the matrix polynomials

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

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

written in the shift-back operator z.

The orders p and q as well as the time-delay d are assumed known.

We estimate θ by a projection algorithm

$$\theta_{n+1} = \theta_n + \frac{\varphi_n}{1 + \|\varphi_n\|^2}(y_{n+1}^\tau - \varphi_n^\tau \theta_n) \tag{4}$$

$$\varphi_n^\tau = [y_n^\tau \quad \dots \quad y_{n-p+1}^\tau \quad u_{n-d+1}^\tau \quad \dots \quad u_{n-q+2-d}^\tau] \tag{5}$$

with arbitrary initial values θ_0 and φ_0.

Set

$$\tilde{\theta}_n = \theta - \theta_n \tag{6}$$

and

$$\Psi(n+1, i) = \left(I - \frac{\varphi_n \varphi_n^\tau}{1 + \|\varphi_n\|^2}\right)\Psi(n, i), \quad \Psi(i, i) = I \tag{7}$$

Then

$$\tilde{\theta}_{n+1} = \left(I - \frac{\varphi_n \varphi_n^\tau}{1 + \|\varphi_n\|^2}\right)\tilde{\theta}_n = \Psi(n+1, 0)\tilde{\theta}_0 \tag{8}$$

We list the conditions used subsequently.

(*a*) $A(z)$ and $B(z)$ are left-coprime and A_p is of full rank.

(*b*) There exists a sequence of nonnegative numbers δ_n (possibly tending to zero) and a sequence of integers τ_n with

$$\tau_0 = 1, \quad d_n \triangleq \tau_n - \tau_{n-1} \geqslant mp \tag{9}$$

such that

$$\sum_{i=\tau_{n-1}+mp}^{\tau_n-1} U_i U_i^\tau \geqslant \delta_n I \quad \forall n \geqslant 1 \tag{10}$$

where

$$U_i \triangleq [u_i^\tau \quad u_{i-1}^\tau \quad \dots \quad u_{i-mp-q+1}^\tau]^\tau \tag{11}$$

Theorem 1

If Conditions (*a*) and (*b*) are satisfied and if there are constants $v \geqslant 0$, $\lambda \geqslant 0$, $\delta \geqslant 0$ and $c > 0$ with $4(1+\lambda)v + 2\delta + 5\lambda < 1$ so that

$$\|\varphi_n\| = O(n^v), \quad d_n = O(n^\lambda), \quad \delta_n > \frac{c}{n^\delta}, \quad \forall n \tag{12}$$

then

$$\|\tilde{\theta}_n\| = O(\exp(-\alpha n^{[1-2\delta-5\lambda-4(1+\lambda)v]/(1+\lambda)})), \quad \text{as } n \to \infty \tag{13}$$

where α is a positive constant.

This theorem says that the parameter estimate remains consistent even if the input and output of the system grow as fast as n^v, if in (10) $\delta_n \to 0$ with rate $O(1/n^\delta)$ and if the number of summands is allowed to grow as fast as n^λ.

Corollary 1

If Conditions (*a*) and (*b*) hold and if

$$\sup_n \|\varphi_n\| < \infty, \quad \sup_n d_n < \infty, \quad \inf_n \delta_n > 0$$

then there is a constant $\gamma \in (0, 1)$ such that

$$\|\tilde{\theta}_n\| = O(\gamma^n), \quad \text{as } n \to \infty \tag{14}$$

This conclusion actually follows from (13) by setting $v = 0$, $\lambda = 0$ and $\delta = 0$ in it.

Corollary 2 (Anderson and Johnson 1982)

If Condition (*a*) holds, $\{y_n\}$ is bounded and there are constants $N > 0$, $\beta_2 > \beta_1 > 0$ such that for any $n \geqslant 0$

$$\beta_1 I \leqslant \sum_{i=n+1}^{n+N} U_i U_i^\tau \leqslant \beta_2 I$$

(a sufficiently rich condition) then (14) is valid.

To be convinced of the assertion one need only take $\tau_n = n(mp + N) + 1$ and $\delta_n = \beta_1$ in Corollary 1.

To prove Theorem 1 we present some lemmas.

Lemma 1

If Condition (*a*) is satisfied, then there is a constant $c_0 > 0$ such that

$$\lambda_{\min}\left(\sum_{i=k}^{N} \varphi_i \varphi_i^\tau\right) \geqslant \frac{c_0}{N-k-mp+1} \lambda_{\min}\left(\sum_{i=k+mp-d+1}^{N-d+1} U_i U_i^\tau\right) \tag{15}$$

for any $N \geq k+mp$, $\forall k \geq 0$, where and hereafter $\lambda_{\min}(X)$ denotes the minimum eigenvalue of a matrix X and U_i is defined by (11).

Proof

Let

$$\det A(z) = a_0 + a_1 z + \dots + a_{mp} z^{mp}, \quad a_{mp} \neq 0 \tag{16}$$

and

$$\psi_n = [\det A(z)]\varphi_n \tag{17}$$

Then

$$\psi_n = [((\text{adj } A(z))B(z)z^d u_n)^\tau \dots ((\text{adj } A(z))B(z)z^{p+d-1}u_n)^\tau z^{d-1} \det A(z)u_n^\tau \dots z^{d+q-2} \det A(z)u_n^\tau]^\tau \tag{18}$$

For any $x \in R^{mp+lq}$ from (17) it is easy to see that

$$x^\tau\left(\sum_{i=k+mp}^{N} \psi_i\psi_i^\tau\right)x = \sum_{i=k+mp}^{N} (x^\tau\psi_i)^2 = \sum_{i=k+mp}^{N}\left(\sum_{j=0}^{mp} a_j x^\tau \varphi_{i-j}\right)^2$$

$$\leq \sum_{j=0}^{mp} a_j^2 \sum_{i=k+mp}^{N} \sum_{j=0}^{mp} (x^\tau\varphi_{i-j})^2$$

$$\leq (N-k-mp+1) \sum_{j=0}^{mp} a_j^2 \sum_{i=k}^{N} x^\tau\varphi_i\varphi_i^\tau x$$

so we have

$$\lambda_{\min}\left(\sum_{i=k}^{N} \varphi_i\varphi_i^\tau\right) \geq \frac{1}{(N-k)\sum_{j=0}^{mp} a_j^2} \lambda_{\min}\left(\sum_{i=k+mp}^{N} \psi_i\psi_i^\tau\right) \tag{19}$$

Hence for (15) we only need to show that

$$\lambda_{\min}\left(\sum_{i=k+mp}^{N} \psi_i\psi_i^\tau\right) \geq c_1 \lambda_{\min}\left(\sum_{i=k+mp-d+1}^{N-d+1} U_i U_i^\tau\right) \quad \text{for some } c_1 > 0 \tag{20}$$

Write $x \in R^{mp+lq}$ in the vector-component form

$$x = [x^{1\tau} \dots x^{p\tau} \quad x^{(p+1)\tau} \dots x^{(p+q)\tau}]^\tau$$

with $x^i \in R^m$, $x^j \in R^l$, $1 \leq i \leq p$, $p+1 \leq j \leq p+q$.

Set

$$H_x(z) = x^{1\tau}(\text{adj } A(z))B(z)z^d + \dots + x^{p\tau}(\text{adj } A(z))B(z)z^{p+d-1}$$
$$+ x^{(p+1)\tau}z^{d-1} \det A(z) + \dots + x^{(p+q)\tau}z^{q+d-2} \det A(z)$$
$$\triangleq \sum_{i=0}^{mp+q-1} g_i^\tau(x) z^{i+d-1} \tag{21}$$

Then from (18) and (21) we see that

$$
\begin{aligned}
x^{\tau} \sum_{i=k+mp}^{N} \psi_i \psi_i^{\tau} x &= \sum_{i=k+mp}^{N} (H_x(z)u_i)^2 \\
&= \sum_{i=k+mp}^{N} \sum_{t=0}^{mp+q-1} \sum_{s=0}^{mp+q-1} g_s^{\tau}(x) u_{i-s-d+1} u_{i-t-d+1}^{\tau} g_t(x) \\
&= g^{\tau}(x) \sum_{i=k+mp-d+1}^{N-d+1} U_i U_i^{\tau} g(x) \\
&\geq \min_{\|x\|=1} \|g(x)\|^2 \lambda_{\min}\left(\sum_{i=k+mp-d+1}^{N-d+1} U_i U_i^{\tau} \right) \qquad (22)
\end{aligned}
$$

where by definition

$$g(x) = [g_0^{\tau}(x) \dots g_{mp+q-1}^{\tau}(x)]^{\tau}$$

Thus for (20) we only need to show

$$\min_{\|x\|=1} \|g(x)\| \neq 0 \qquad (23)$$

Suppose the converse were true. Then by continuity of $g(x)$ there exists some x such that $g(x) = 0$ and $\|x\| = 1$. For this x by (21) $H_x(z) \equiv 0$, i.e.

$$\sum_{i=1}^{p} x^{i\tau}(\text{adj } A(z))B(z)z^i = -\sum_{j=1}^{q} x^{(p+j)\tau} \det A(z) I z^{j-1} \qquad (24)$$

Setting $z = 0$ from (24) we see $x^{p+1} = 0$, and (24) can be rewritten as

$$\sum_{i=1}^{p} x^{i\tau}(\text{adj } A(z))B(z)z^i = -\sum_{j=2}^{q} x^{(p+j)\tau} \det A(z) I z^{j-1} \qquad (25)$$

By coprimeness of $A(z)$ and $B(z)$ there are matrix polynomials $M(z)$ and $N(z)$ such that

$$A(z)M(z) + B(z)N(z) = I$$

Hence from (25) it follows that

$$
\begin{aligned}
\sum_{i=1}^{p} x^{i\tau}(\text{adj } A(z))z^i &= \sum_{i=1}^{p} x^{i\tau} z^i (\text{adj } A(z))(A(z)M(z) + B(z)N(z)) \\
&= \det A(z)\left(\sum_{i=1}^{p} x^{i\tau} z^i M(z) - \sum_{j=2}^{q} x^{(p+j)\tau} N(z) z^{j-1} \right) \\
&= z(\det A(z))\left(\sum_{i=1}^{p} x^{i\tau} z^{i-1} M(z) - \sum_{j=2}^{q} x^{(p+j)\tau} N(z) z^{j-2} \right)
\end{aligned}
$$

Noticing

$$\deg\left(\sum_{i=1}^{p} x^{i\tau}(\text{adj } A(z))z^i \right) = mp < mp + 1 = \deg(z \det A(z))$$

we conclude that $x^i = 0$, $1 \leq i \leq p$, then $x^j = 0$, $p + 1 \leq j \leq p + q$ by (25). This contradicts $\|x\| = 1$ and thus (23) is proved. □

Lemma 2

If

$$\sum_{i=k}^{N-1} \frac{\varphi_i \varphi_i^\tau}{1+\|\varphi_i\|^2} \geqslant \delta I, \quad \delta > 0$$

then

$$\|\Psi(N,k)\| \leqslant \left[1 - \frac{\delta^2}{4(N-k)^3}\right]^{1/2}$$

Proof

Let x_0 be the unit eigenvector corresponding to the maximum eigenvalue ρ of $\Psi^\tau(N,k)\Psi(N,k)$.

From the difference equation

$$x_{i+1} = \left(I - \frac{\varphi_i \varphi_i^\tau}{1+\|\varphi_i\|^2}\right) x_i, \quad x_k = x_0, \quad k \leqslant i \leqslant N-1 \tag{26}$$

it is easy to see

$$x_N = \Psi(N,k) x_0$$

and

$$\|x_N\|^2 = x_0^\tau \Psi^\tau(N,k)\Psi(N,k) x_0 = \rho \tag{27}$$

From (26) we have

$$x_{i+1}^\tau x_{i+1} \leqslant x_i^\tau x_i - \frac{\|\varphi_i^\tau x_i\|^2}{1+\|\varphi_i\|^2} \tag{28}$$

then by (27) and (28) we see

$$\sum_{i=k}^{N-1} \frac{\|\varphi_i^\tau x_i\|^2}{1+\|\varphi_i\|^2} \leqslant 1 - \rho \tag{29}$$

For any integer $i \in [k, N-1]$ by (26) and (29) we get

$$\begin{aligned} \|x_i - x_0\| &\leqslant \left\| \sum_{j=k}^{i-1} \frac{\varphi_j \varphi_j^\tau}{1+\|\varphi_j\|^2} x_j \right\| \\ &\leqslant \left[\sum_{j=k}^{i-1} \frac{\|\varphi_j\|^2}{1+\|\varphi_j\|^2}\right]^{1/2} \cdot \left[\sum_{j=k}^{i-1} \frac{\|\varphi_j^\tau x_j\|^2}{1+\|\varphi_j\|^2}\right]^{1/2} \\ &\leqslant (N-k)^{1/2}(1-\rho)^{1/2} \end{aligned} \tag{30}$$

Thus, by (29) and (30) we can estimate as follows:

$$\begin{aligned} \delta &\leqslant x_0^\tau \sum_{i=k}^{N-1} \frac{\varphi_i \varphi_i^\tau}{1+\|\varphi_i\|^2} x_0 \\ &\leqslant \sum_{i=k}^{N-1} \frac{\|\varphi_i\|}{1+\|\varphi_i\|^2} \|\varphi_i^\tau x_i\| + \sum_{i=k}^{N-1} \frac{\|\varphi_i\|^2}{1+\|\varphi_i\|^2} \|x_i - x_0\| \\ &\leqslant (N-k)^{1/2}(1-\rho)^{1/2} + (N-k)^{3/2}(1-\rho)^{1/2} \\ &\leqslant 2(N-k)^{3/2}(1-\rho)^{1/2} \end{aligned}$$

Hence finally we conclude that

$$\rho \leqslant 1 - \frac{\delta^2}{4(N-k)^3} \qquad \square$$

Lemma 3

If Conditions (*a*) and (*b*) are satisfied, then

$$\|\Psi(\tau_n + d - 1, 0)\| \leqslant \exp\left(-c_1 \sum_{i=1}^{n} \frac{\delta_i^2}{M_i^2 d_i^5}\right) \tag{31}$$

where

$$M_i = \sup_{\tau_{i-1} \leqslant j \leqslant \tau_i - 1} \|\varphi_j\|^2 + 1 \quad \text{and} \quad c_1 > 0$$

Proof

By Lemma 1 and Condition (*b*) it is clear that

$$\sum_{i=\tau_{n-1}+d-1}^{\tau_n + d - 2} \frac{\varphi_i \varphi_i^{\tau}}{1 + \|\varphi_i\|^2} \geqslant \frac{c_0 \delta_n}{M_n d_n}$$

Then using Lemma 2 we know that there is a constant $c_1' > 0$ such that

$$\|\Psi(\tau_n + d - 1, \tau_{n-1} + d - 1)\| \leqslant \left[1 - c_1' \frac{\delta_n^2}{M_n^2 d_n^5}\right]^{1/2} \tag{32}$$

From (7) and (32) it follows that

$$\begin{aligned} \|\Psi(\tau_n + d - 1, 0)\| &\leqslant \prod_{i=1}^{n} \|\Psi(\tau_i + d - 1, \tau_{i-1} + d - 1\| \\ &\leqslant \left(\prod_{i=1}^{n} \left(1 - c_1' \frac{\delta_i^2}{M_i^2 d_i^5}\right)\right)^{1/2} \end{aligned} \tag{33}$$

Finally, taking notice of the elementary inequality $1 - x \leqslant e^{-x}, 0 \leqslant x \leqslant 1$, from (33) we conclude (31) with $c_1 = \frac{1}{2} c_1'$. $\square$

Proof of Theorem 1

Since $d_n = O(n^{\lambda})$ we can find $\beta > 0$ so that $d_n \leqslant \beta n^{\lambda}$. Then

$$mp \cdot n \leqslant \sum_{i=1}^{n} d_i \leqslant n \cdot \beta \cdot n^{\lambda} = \beta n^{1+\lambda}$$

or equivalently,

$$mp \cdot n + 1 \leqslant \tau_n \leqslant \beta n^{1+\lambda} + 1 \tag{34}$$

By (12) we see $\|\varphi_{\tau_n}\| = O(\tau_n^{\nu})$, then by (34) and the definition of M_i we find that

$$M_i^2 = O(\tau_i^{4\nu}) = O(i^{4\nu(1+\lambda)}) \tag{35}$$

So (12) and Lemma 3 lead to

$$\|\Psi(\tau_n + d - 1, 0)\| \leqslant \exp\left(-c_2 \sum_{i=1}^{n} \frac{1}{i^{2\delta+5\lambda+4\nu(1+\lambda)}}\right)$$
$$= O\left(\exp\left(-\frac{c_2}{1-2\delta-5\lambda-4\nu(1+\lambda)}(n+1)^{1-2\delta-5\lambda-4\nu(1+\lambda)}\right)\right) \quad (36)$$

for some $c_2 > 0$.

Since $\tau_{n-1} < \tau_n \to \infty$, for any n there exists k such that

$$\tau_k + d - 1 \leqslant n \leqslant \tau_{k+1} + d - 1 \quad (37)$$

Then by (34) and (37) it follows that

$$(k+1)^{1+\lambda} \geqslant \frac{\tau_{k+1} - 1}{\beta} \geqslant \frac{n-d}{\beta}$$

or

$$k \geqslant \left(\frac{n-d}{\beta}\right)^{1/(1+\lambda)} - 1$$

and by (36)

$$\|\Psi(n, 0)\| \leqslant \|\Psi(\tau_k + d - 1, 0)\|$$
$$= O\left(\exp\left(-\frac{c_2}{1-2\delta-5\lambda-4\nu(1+\lambda)}(k+1)^{1-2\delta-5\lambda-4\nu(1+\lambda)}\right)\right)$$
$$= O\left(\exp\left(\frac{c_2}{1-2\delta-5\lambda-4\nu(1+\lambda)}\left(\frac{n-d}{\beta}\right)^{[1-2\delta-5\lambda-4\nu(1+\lambda)]/(1+\lambda)}\right)\right)$$
$$= O(\exp(-\alpha n^{[1-2\delta-5\lambda-4\nu(1+\lambda)]/(1+\lambda)})), \quad \text{with } \alpha > 0 \quad (38)$$

which together with (8) gives the desired result.

3. Parameter identification for systems with attenuating excitation control

Let $\{\mathscr{F}_n\}$ be a family of non-decreasing σ-algebras. Take a sequence of l-dimensional i.i.d. random vectors $\{\varepsilon_n\}$ such that ε_n is $\mathscr{F}_n$-measurable and is independent of $\mathscr{F}_{n-1}$ and that

$$E(\varepsilon_n \mid \mathscr{F}_{n-1}) = 0, \quad E\varepsilon_n\varepsilon_n^{\tau} = \mu I, \quad \|\varepsilon_n\| \leqslant M \quad (39)$$

where $\mu > 0$, $M > 0$ are constants.

Let u_n^0 be a $\mathscr{F}_{n-1}$-measurable desired control. We add to u_n^0 a dither v_n tending to zero:

$$v_n = \frac{\varepsilon_n}{n^{\varepsilon}}, \quad \varepsilon \in \left(0, \frac{1}{6 + 12(mp+q)}\right) \quad (40)$$

and the resulting control

$$u_n = u_n^0 + v_n \quad (41)$$

is called *attenuating excitation control.*

In this section for systems with attenuating excitation control (41) we establish consistency of parameter estimation without requiring any condition like (10).

Lemma 4

Let $\sigma > 0$ be a stopping time with respect to $\{\mathscr{F}_i\}$. Then

$$\sum_{i=\sigma}^{n+\sigma} f_i v_i^{\tau} = o(n^{1-2(1+3(mp+q))\varepsilon}) \tag{42}$$

for any $\mathscr{F}_{i-1}$-measurable f_i with $\|f_i\| = O(i^{\nu})$ and $0 \leqslant \nu < \varepsilon$, and

$$\frac{1}{(n+\sigma)^{1-2\varepsilon} - \sigma^{1-2\varepsilon}} \sum_{i=\sigma}^{n+\sigma} v_i v_i^{\tau} > c_2 I, \quad c_2 > 0 \tag{43}$$

for $\forall n \geqslant n_0$ where n_0 and c_2 may be ω-dependent.

Proof

Since $\|f_i\| = O(i^{\nu})$, there exists ω-dependent $\xi(\omega)$ so that $\|f_i\| \leqslant \xi i^{\nu}$.

Clearly, for $\varepsilon < 1/[6 + 12(mp+q)]$

$$\sum_{i=1}^{\infty} \frac{\|f_i\|^2}{i^{2-4(1+3(mp+q))\varepsilon}} \leqslant \xi^2 \sum_{i=1}^{\infty} \frac{i^{2\varepsilon}}{i^{2-4(1+3(mp+q))\varepsilon}} < \infty$$

Then by the martingale convergence theorem

$$\sum_{i=\sigma}^{\infty} \frac{f_i^{\tau} v_i}{i^{1-2(1+3(mp+q))\varepsilon}} < \infty$$

and (42) follows from the Kronecker lemma.

For (43) we first show

$$\frac{1}{(n+\sigma)^{1-2\varepsilon} - \sigma^{1-2\varepsilon}} \sum_{i=\sigma+1}^{n+\sigma} \left(v_i v_i^{\tau} - \frac{\mu}{i^{2\varepsilon}} I \right) \xrightarrow[n\to\infty]{} 0, \quad \text{a.s.} \tag{44}$$

Since $(v_{i+\sigma} v_{i+\sigma}^{\tau} - (\mu/(i+\sigma)^{2\varepsilon} I), \mathscr{F}_{i+\sigma})$ is a martingale difference sequence and

$$\sum_{i=1}^{\infty} E\left\{ \left[\left\| v_{i+\sigma} v_{i+\sigma}^{\tau} - \frac{\mu}{(i+\sigma)^{2\varepsilon}} I \right\| \Big/ (i+\sigma)^{1-3\varepsilon} \right]^2 | \mathscr{F}_{\sigma+i-1} \right\}$$
$$\leqslant 2l(M^4 + \mu^2) \sum_{i=1}^{\infty} \frac{1}{(i+\sigma)^{2(1-\varepsilon)}} < \infty$$

Again by the martingale convergence theorem and the Kronecker lemma we obtain

$$\frac{1}{(n+\sigma)^{1-3\varepsilon}} \sum_{i=1}^{n} \left(v_{i+\sigma} v_{i+\sigma}^{\tau} - \frac{\mu}{(i+\sigma)^{2\varepsilon}} I \right) \xrightarrow[n\to\infty]{} 0, \quad \text{a.s.}$$

which implies (44).

Noticing the elementary inequalities for $n > n_0$

$$\frac{1}{1-2\varepsilon}(n^{1-2\varepsilon} - n_0^{1-2\varepsilon}) = \sum_{i=n_0}^{n-1} \int_i^{i+1} \frac{dx}{x^{2\varepsilon}} \leqslant \sum_{i=n_0}^{n-1} \frac{1}{i^{2\varepsilon}} = \sum_{i=n_0}^{n-1} \int_{i-1}^{i} \frac{dx}{i^{2\varepsilon}} \leqslant \sum_{i=n_0}^{n-1} \int_{i-1}^{i} \frac{dx}{x^{2\varepsilon}}$$
$$= \frac{1}{1-2\varepsilon}[(n-1)^{1-2\varepsilon} - (n_0-1)^{1-2\varepsilon}] \tag{45}$$

we see

$$\frac{1}{(n+\sigma)^{1-2\varepsilon} - \sigma^{1-2\varepsilon}} \sum_{i=\sigma+1}^{n+\sigma} \frac{\mu}{i^{2\varepsilon}} I \xrightarrow[n\to\infty]{} \frac{\mu}{1-2\varepsilon} I$$

which together with (44) proves (43). □

Lemma 5

Let $\{\tau_k\}$ and $\{\sigma_k\}$ be two sequences of finite stopping times with respect to $\{\mathscr{F}_k\}$, $\sigma_k \geqslant \tau_k > \sigma_{k-1}$.

(i) If there is a sequence of real numbers $b_i > 0$ so that

$$\frac{1}{b_k} \geqslant \frac{1}{(\tau_k^{1-2\varepsilon} - \sigma_{k-1}^{1-2\varepsilon})^{1/2}} \quad \text{and} \quad \sum_{i=1}^{\infty} \frac{1}{b_i^2} < \infty \tag{46}$$

then for sufficiently large k

$$\sum_{i=\sigma_{k-1}}^{\tau_k} v_i v_i^\tau \geqslant \beta(\tau_k^{1-2\varepsilon} - \sigma_{k-1}^{1-2\varepsilon})I \quad \text{for some } \beta > 0 \tag{47}$$

(ii) For any $\mathscr{F}_{i-1}$-measurable f_i with $\|f_i\| = O(i^\nu)$, $0 \leqslant \nu < \varepsilon$

$$\frac{1}{(\tau_k^{1-\varepsilon+\nu} - \sigma_{k-1}^{1-\varepsilon+\nu})^{1/2} \beta_k \log \beta_k} \sum_{i=\sigma_{k-1}}^{\tau_k} f_i v_i^\tau \xrightarrow[k\to\infty]{} 0, \quad \text{a.s.} \tag{48}$$

whenever $\beta_k > 1$, $\forall k$ and $\sum_{k=1}^{\infty} \frac{1}{\beta_k^2} < \infty$.

Proof

(i) From (45) we see

$$\frac{1}{\tau_k^{1-2\varepsilon} - \sigma_{k-1}^{1-2\varepsilon}} \sum_{i=\sigma_{k-1}}^{\tau_k - 1} \frac{1}{i^{2\varepsilon}} > \frac{1}{1-2\varepsilon}$$

So for (47) it suffices to show

$$\frac{1}{\tau_k^{1-2\varepsilon} - \sigma_{k-1}^{1-2\varepsilon}} \left\| \sum_{i=\sigma_{k-1}+1}^{\tau_k} \left(v_i v_i^\tau - \frac{\mu}{i^{2\varepsilon}} I \right) \right\| \xrightarrow[k\to\infty]{} 0, \quad \text{a.s.} \tag{49}$$

for which we only need to prove

$$\frac{1}{(\tau_k^{1-2\varepsilon} - \sigma_{k-1}^{1-2\varepsilon})^{1/2} \cdot b_k} \sum_{i=\sigma_{k-1}+1}^{\tau_k} \left(v_i v_i^\tau - \frac{\mu}{i^{2\varepsilon}} I \right) \xrightarrow[k\to\infty]{} 0, \quad \text{a.s.} \tag{50}$$

Set

$$t = \tau_k - \sigma_{k-1}, \quad a_i = \frac{1}{(i + \sigma_{k-1})^{1-2\varepsilon} - \sigma_{k-1}^{1-2\varepsilon}}, \quad a_0 = 0$$

$$x_i = \frac{\varepsilon_{\sigma_{k-1}+i} \varepsilon_{\sigma_{k-1}+i}^\tau - \mu I}{(\sigma_{k-1} + i)^{2\varepsilon}}, \quad S_n = \sum_{i=1}^{n} x_i$$

For any $\eta > 0$ by the conditional Markov inequality we have

$$\begin{aligned} P\left(\frac{1}{(\tau_k^{1-2\varepsilon} - \sigma_{k-1}^{1-2\varepsilon})^{1/2} \cdot b_k} \left\| \sum_{i=\sigma_{k-1}+1}^{\tau_k} \left(v_i v_i^\tau - \frac{\mu}{i^{2\varepsilon}} I \right) \right\| > \eta \,\middle|\, \mathscr{F}_{\sigma_{k-1}} \right) \\ \leqslant \frac{1}{b_k^2 \eta^2} E\left[\frac{1}{\tau_k^{1-2\varepsilon} - \sigma_{k-1}^{1-2\varepsilon}} \left\| \sum_{i=\sigma_{k-1}+1}^{\tau_k} \left(v_i v_i^\tau - \frac{\mu}{i^{2\varepsilon}} I \right) \right\|^2 \,\middle|\, \mathscr{F}_{\sigma_{k-1}} \right] \\ = \frac{1}{b_k^2 \eta^2} E[a_t \|S_t\|^2 \mid \mathscr{F}_{\sigma_{k-1}}] \end{aligned} \tag{51}$$

If t is a bounded stopping time, $t \leqslant n_0$, then we have

$$E[a_t \|S_t\|^2 \mid \mathscr{F}_{\sigma_{k-1}}] = E\left[\sum_{i=1}^{n_0} I_{[t=i]} a_i \|S_i\|^2 \,\middle|\, \mathscr{F}_{\sigma_{k-1}}\right] \leqslant \operatorname{tr} E\left[\sum_{i=1}^{n_0} I_{[t=i]} a_i S_i S_i^{\tau} \,\middle|\, \mathscr{F}_{\sigma_{k-1}}\right]$$

$$= \operatorname{tr} \sum_{i=1}^{n_0} E\left[I_{[t=i]} \sum_{j=1}^{i} (a_j S_j S_j^{\tau} - a_{j-1} S_{j-1} S_{j-1}^{\tau}) \,\middle|\, \mathscr{F}_{\sigma_{k-1}}\right]$$

$$= \sum_{i=1}^{n_0} E\left[\sum_{j=1}^{i} I_{[t=i]} \operatorname{tr} (a_j (S_{j-1} + x_j)(S_{j-1} + x_j)^{\tau} - a_{j-1} S_{j-1} S_{j-1}^{\tau}) \,\middle|\, \mathscr{F}_{\sigma_{k-1}}\right]$$

$$\leqslant E\left[\sum_{i=1}^{n_0} \sum_{i=1}^{i} I_{[t=i]} \operatorname{tr} (2a_j S_{j-1} x_j^{\tau} + a_j x_j x_j^{\tau}) \,\middle|\, \mathscr{F}_{\sigma_{k-1}}\right]$$

$$= \sum_{j=1}^{n_0} \operatorname{tr} E[I_{[t \geqslant j]} (2a_j S_{j-1} x_j^{\tau} + a_j x_j x_j^{\tau}) \mid \mathscr{F}_{\sigma_{k-1}}] \tag{52}$$

Noticing the properties of stopping times (cf. Lipster and Shiryayev 1978, Chow and Teicher 1978), we see

$$I_{[t \geqslant j]} = 1 - I_{[\tau_k \leqslant \sigma_{k-1} + j - 1]} \in \mathscr{F}_{\sigma_{k-1} + j - 1}$$

$$a_j \in \mathscr{F}_{\sigma_{k-1}} \subset \mathscr{F}_{\sigma_{k-1} + j - 1}, \quad S_{j-1} \in \mathscr{F}_{\sigma_{k-1} + j - 1}$$

and $E(x_j \mid \mathscr{F}_{\sigma_{k-1} + j - 1}) = 0$, we then have

$$E[I_{[t \geqslant j]} a_j S_{j-1} x_j^{\tau} \mid \mathscr{F}_{\sigma_{k-1}}] = 0$$

and

$$\operatorname{tr} E[I_{[t \geqslant j]} a_j x_j x_j^{\tau} \mid \mathscr{F}_{\sigma_{k-1}}] \leqslant 2l(M^4 + \mu^2) E\left[\frac{a_j}{(\sigma_{k-1} + j)^{4\varepsilon}} I_{[t \geqslant j]} \mid \mathscr{F}_{\sigma_{k-1}}\right]$$

then we continue to estimate (52):

$$E[a_t \|S_t\|^2 \mid \mathscr{F}_{\sigma_{k-1}}] \leqslant 2l(M^4 + \mu^2) \sum_{j=1}^{n_0} E\left[I_{[t \geqslant j]} \frac{a_j}{(\sigma_{k-1} + j)^{4\varepsilon}} \,\middle|\, \mathscr{F}_{\sigma_{k-1}}\right]$$

$$= 2l(M^4 + \mu^2) E\left[\sum_{j=1}^{t} \frac{a_j}{(\sigma_{k-1} + j)^{4\varepsilon}} \,\middle|\, \mathscr{F}_{\sigma_{k-1}}\right]$$

$$\leqslant 2l(M^4 + \mu^2) E\left[\sum_{j=1}^{\infty} \frac{a_j}{(\sigma_{k+1} + j)^{4\varepsilon}} \,\middle|\, \mathscr{F}_{\sigma_{k-1}}\right] \tag{53}$$

Thus we have proved (53) for bounded t, but it is also true for unbounded t, since it holds for $t(n) = \min\,[t, n]$, then the Fatou lemma yields the desired estimate. Then by

(53) we continue to estimate (51):

$$P\left(\frac{1}{(\tau_k^{1-2\varepsilon}-\sigma_{k-1}^{1-2\varepsilon})^{1/2}\cdot b_k}\left\|\sum_{i=\sigma_{k-1}+1}^{\tau_k}\left(v_i v_i^\tau-\frac{\mu}{i^{2\varepsilon}}I\right)\right\|>\eta\,\middle|\,\mathscr{F}_{\sigma_{k-1}}\right)$$

$$\leqslant\frac{2l(M^4+\mu)}{b_k^2\eta^2}E\left[\sum_{j=1}^{\infty}\frac{a_j}{(\sigma_{k-1}+j)^{4\varepsilon}}\,\middle|\,\mathscr{F}_{\sigma_{k-1}}\right]$$

$$=\frac{2l(M^4+\mu^2)}{b_k^2\eta^2}\sum_{j=1}^{\infty}\frac{1}{[(\sigma_{k-1}+j)^{1-2\varepsilon}-\sigma_{k-1}^{1-2\varepsilon}](\sigma_{k-1}+j)^{4\varepsilon}}$$

$$\leqslant\frac{2l(M^4+\mu^2)}{b_k^2\eta^2}\left\{\frac{1}{[(1+\sigma_{k-1})^{1-2\varepsilon}-\sigma_{k-1}^{1-2\varepsilon}](\sigma_{k-1}+1)^{4\varepsilon}}+\int_1^{\infty}\frac{dx}{(\sigma_{k-1}+x)^{4\varepsilon}[(x+\sigma_{k-1})^{1-2\varepsilon}-\sigma_{k-1}^{1-2\varepsilon}]}\right\}$$

$$\leqslant O\left(\frac{1}{b_k^2}\cdot\frac{\sigma_{k-1}^{2\varepsilon}}{\sigma_{k-1}^{4\varepsilon}}\right)+\frac{2l(M^4+\mu^2)}{(1-2\varepsilon)b_k^2\eta^2}\int_{(\sigma_{k-1}+1)^{1-2\varepsilon}-\sigma_{k-1}^{1-2\varepsilon}}^{\infty}\frac{dy}{y(y+\sigma_{k-1}^{1-2\varepsilon})^{(2\varepsilon)/(1-2\varepsilon)}}$$

$$=O\left(\frac{1}{b_k^2}\right)+\frac{2l(M^4+\mu^2)}{(1-2\varepsilon)b_k^2\eta^2}\int_{(\sigma_{k-1}+1)^{1-2\varepsilon}-\sigma_{k-1}^{1-2\varepsilon}}^{\infty}\frac{dy}{y(y+\sigma_{k-1}^{1-2\varepsilon})^{2\varepsilon+(2\varepsilon)^2/(1-2\varepsilon)}}$$

$$\leqslant O\left(\frac{1}{b_k^2}\right)+\frac{2l(M^4+\mu^2)}{(1-2\varepsilon)b_k^2\eta^2}\int_{(\sigma_{k-1}+1)^{1-2\varepsilon}-\sigma_{k-1}^{1-2\varepsilon}}^{\infty}\frac{dy}{y\cdot y^{2\varepsilon}\cdot\sigma_{k-1}^{(2\varepsilon)^2}}$$

$$=O\left(\frac{1}{b_k^2}\right)+\frac{2l(M^4+\mu^2)}{2\varepsilon(1-2\varepsilon)b_k^2\eta^2\sigma_{k-1}^{(2\varepsilon)^2}}\cdot\frac{1}{[(\sigma_{k-1}+1)^{1-2\varepsilon}-\sigma_{k-1}^{1-2\varepsilon}]^{2\varepsilon}}$$

$$=O\left(\frac{1}{b_k^2}\right)+O\left(\frac{(\sigma_{k-1})^{(2\varepsilon)^2}}{b_k^2\sigma_{k-1}^{(2\varepsilon)^2}}\right)\leqslant c_3\left(\frac{1}{b_k^2}\right)$$

where $c_3>0$ is not dependent on k.

Since $\sum_{k=1}^{\infty}(1/b_k^2)<\infty$, then by the Borel–Cantelli–Lévy lemma (Chow *et al.* 1971) we conclude (50) and hence (49).

(ii) For (48) it suffices to show

$$\frac{1}{(\tau_k^{1-\varepsilon+v}-\sigma_{k-1}^{1-\varepsilon+v})^{1/2}\beta_k\log\beta_k}\sum_{i=\sigma_{k-1}+1}^{\tau_k}f_i I_{[\|f_i\|\leqslant(\log\beta_k)i^v]}\cdot v_i^\tau\xrightarrow[k\to\infty]{}0 \qquad (54)$$

since $\beta_k\xrightarrow[k\to\infty]{}\infty$.

The proof of (54) is similar to that for (50). Instead of a_i, x_i and S_n we should set

$$a_i'=\frac{1}{(i+\sigma_{k-1})^{1-\varepsilon+v}-\sigma_{k-1}^{1-\varepsilon+v}}$$

$$x_i'=\frac{f_{\sigma_{k-1}+i}I_{[\|f_{\sigma_{k-1}+i}\|\leqslant(\log\beta_k)(\sigma_{k-1}+i)^v]}\cdot\varepsilon_{\sigma_{k-1}+i}^\tau}{(\sigma_{k-1}+i)^\varepsilon}$$

$$S_n'=\sum_{i=1}^{n}x_i'$$

Then (51) and (52) is repdaced by

$$P\left(\frac{1}{(\tau_k^{1-\varepsilon+\nu}-\sigma_{k-1}^{1-\varepsilon+\nu})^{1/2}\cdot\beta_k\log\beta_k}\left\|\sum_{i=\sigma_{k-1}+1}^{\tau_k} f_i I_{[\|f_i\|\leqslant(\log\beta_k)i^\nu]}\cdot v_i^\tau\right\|>\eta\,\middle|\,\mathscr{F}_{\sigma_{k-1}}\right)$$
$$\leqslant\frac{1}{\eta^2\beta_k^2\log^2\beta_k}\sum_{j=1}^{\infty}\operatorname{tr}E[I_{[t\geqslant j]}(2a'_jS'_{j-1}x_j'^\tau+a'_jx'_jx_j'^\tau)\,|\,\mathscr{F}_{\sigma_{k-1}}] \tag{55}$$

Again we have

$$a'_j\in\mathscr{F}_{\sigma_{k-1}}\subset\mathscr{F}_{\sigma_{k-1}+j-1},\quad S'_{j-1}\in\mathscr{F}_{\sigma_{k-1}+j-1},\quad E(x'_j\,|\,\mathscr{F}_{\sigma_{k-1}+j-1})=0$$

and

$$\operatorname{tr}E[I_{[t\geqslant j]}a'_jx'_jx_j'^\tau\,|\,\mathscr{F}_{\sigma_{k-1}}]\leqslant lM^2E\left[a'_j\frac{(\log\beta_k)^2(\sigma_{k-1}+i)^{2\nu}}{(\sigma_{k-1}+i)^{2\varepsilon}}I_{[t\geqslant j]}\,\middle|\,\mathscr{F}_{\sigma_{k-1}}\right]$$

Then (55) is estimated by

$$\begin{aligned}(55)&\leqslant\frac{lM^2}{\eta^2\beta_k^2}E\left[\sum_{j=1}^{t}\frac{a'_j}{(\sigma_{k-1}+j)^{2\varepsilon-2\nu}}\,\middle|\,\mathscr{F}_{\sigma_{k-1}}\right]\\
&\leqslant\frac{lM^2}{\eta^2\beta_k^2}\sum_{j=1}^{\infty}\frac{1}{(\sigma_{k-1}+j)^{2(\varepsilon-\nu)}[(\sigma_{k-1}+j)^{1-\varepsilon+\nu}-\sigma_{k-1}^{1-\varepsilon+\nu}]}\\
&\leqslant\frac{lM^2}{\eta^2\beta_k^2}\cdot\frac{1}{(\sigma_{k-1}+1)^{2(\varepsilon-\nu)}[(\sigma_{k-1}+1)^{1-\varepsilon+\nu}-\sigma_{k-1}^{1-\varepsilon+\nu}]}\\
&\quad+\frac{lM^2}{\eta^2\beta_k^2(1-\varepsilon+\nu)}\int_{(\sigma_{k-1}+1)^{1-\varepsilon+\nu}-\sigma_{k-1}^{1-\varepsilon+\nu}}^{\infty}\frac{dy}{y(y+\sigma_{k-1}^{1-\varepsilon+\nu})^{\varepsilon-\nu+(\varepsilon-\nu)^2/(1-\varepsilon+\nu)}}\\
&\leqslant c_4\frac{1}{\beta_k^2}\end{aligned} \tag{56}$$

where $c_4>0$ and is independent of k.

Again by the Borel–Cantelli–Lévy lemma the conclusion (48) follows. □

Define

$$\tau_0=1$$
$$\tau_n=\begin{cases}\inf\left\{k>\tau_{n-1}:\displaystyle\sum_{i=\tau_{n-1}+mp}^{k-1}U_iU_i^\tau\geqslant I\right\}\\[2ex] \infty,\quad\text{if }\lambda_{\min}\left(\displaystyle\sum_{i=\tau_{n-1}+mp}^{k-1}U_iU_i^\tau\right)<1,\quad\forall k>\tau_{n-1}\end{cases} \tag{57}$$

where U_i is given by (11).

Lemma 6

If u_n is defined by (41) and $\|u_n^0\|=O(n^\nu)$, $0\leqslant\nu<\varepsilon$, then $\tau_n<\infty$ a.s. $\forall n\geqslant1$.

Proof

Let S be the set where the conclusion of the lemma does not hold.

To be precise, assume

$$\lambda_{\min}\left(\sum_{i=\sigma}^{n+\sigma} U_i U_i^\tau\right) < 1, \quad \forall n \geqslant 0 \text{ on } S \tag{58}$$

for some $\sigma \triangleq \tau_{n_0} + mp$.

Let $x^n = [x_1^{n\tau} \ldots x_{mp+q}^{n\tau}]^\tau$ be the unit minimum eigenvector of $\sum_{i=\sigma}^{n+\sigma} U_i U_i^\tau$, where $x_i^n \in R^l$ and $\|x^n\| = 1$.

Clearly, $\|u_i\| = O(i^\nu)$, then by (42) and the boundedness of x_i^n, $i = 1, \ldots, mp+q$, we find that (no measurability of x^n is required):

$$\sum_{i=\sigma}^{n+\sigma} v_i(x_1^{n\tau}u_i^0 + x_2^{n\tau}u_{i-1} + \ldots + x_{mp+q}^{n\tau}u_{i-mp-q+1}) = o(n^{1-2(1+3(mp+q))\varepsilon}), \quad \text{a.s.} \tag{59}$$

and by (43) for $\omega \in S$,

$$\begin{aligned} 1 &> \sum_{i=\sigma}^{n+\sigma} (x^{n\tau}U_i)^2 \\ &\geqslant \sum_{i=\sigma}^{n+\sigma} x_1^{n\tau}v_i v_i^\tau x_1^n + 2\sum_{i=\sigma}^{n+\sigma} x_1^{n\tau}v_i(x_1^{n\tau}u_i^0 + x_2^{n\tau}u_{i-1} + \ldots + x_{mp+q}^{n\tau}u_{i-mp-q+1}) \\ &\geqslant c_2\|x_1^n\|^2((n+\sigma)^{1-2\varepsilon} - \sigma^{1-2\varepsilon}) + o(\|x_1^n\| n^{1-2(1+3(mp+q))\varepsilon}) \\ &= c_2\|x_1^n\|(\|x_1^n\|(n+\sigma)^{1-2\varepsilon} - \sigma^{1-2\varepsilon}) + o(n^{1-2(1+3(mp+q))\varepsilon}) \end{aligned} \tag{60}$$

From here it is easy to conclude

$$\|x_1^n\| = o\left(\frac{n^{1-2(1+3(mp+q))\varepsilon}}{(n+\sigma)^{1-2\varepsilon} - \sigma^{1-2\varepsilon}}\right) = o(n^{-6(mp+q)\varepsilon}) \tag{61}$$

We now show for $\omega \in S$

$$\|x_i^n\| = o(n^{-6(mp+q-(1/3)(i-1))\varepsilon}), \quad i = 1, \ldots, mp+q \tag{62}$$

The estimate (61) shows that (62) holds for $i = 1$. Let it be held for $i = 1, \ldots, s$, $s < mp+q$. Then we have

$$\begin{aligned} 1 &> \sum_{i=\sigma}^{n+\sigma} (x^{n\tau}U_i)^2 \\ &\geqslant \sum_{i=\sigma}^{n+\sigma} x_{s+1}^{n\tau}v_{i-s}v_{i-s}^\tau x_{s+1}^n + 2\sum_{i=1}^{n+\sigma} x_{s+1}^{n\tau}v_{i-s}(x_1^{n\tau}u_i + \ldots + x_s^{n\tau}u_{i-s+1}) \\ &\quad + 2\sum_{i=\sigma}^{n+\sigma} x_{s+1}^{n\tau}v_{i-s}(u_{i-s}^{0\tau}x_{s+1}^n + u_{i-s-1}^\tau x_{s+2}^n + \ldots + u_{i-mp-q+1}^\tau x_{mp+q}^n) \end{aligned} \tag{63}$$

Noticing there is a ω-dependent η such that

$$\left\|\sum_{i=\sigma}^{n+\sigma} x_{s+1}^{n\tau}v_{i-s}(x_1^{n\tau}u_i + \ldots + x_s^{n\tau}u_{i-s+1})\right\| \leqslant \eta\|x_{s+1}^n\| \sum_{i=\sigma}^{n+\sigma} \frac{i^\nu}{i^\varepsilon}(\|x_1^n\| + \ldots + \|x_s^n\|)$$
$$= o(\|x_{s+1}^n\| \cdot n^{1-6(mp+q-(1/3)(s-1))\varepsilon}), \quad \omega \in S$$

Then by (59) and (43) from (63) we have

$$\begin{aligned} 1 &> c_2((n+\sigma)^{1-2\varepsilon} - \sigma^{1-2\varepsilon})\|x_{s+1}^n\|^2 + o(\|x_{s+1}^n\| n^{1-6(mp+q-(1/3)(s-1))\varepsilon}) \\ &\quad + o(\|x_{s+1}^n\| \cdot n^{1-2(1+3(mp+q))\varepsilon}) \\ &= ((n+\sigma)^{1-2\varepsilon} - \sigma^{1-2\varepsilon})\|x_{s+1}^n\|(c_2\|x_{s+1}^n\| + o(n^{-6(mp+q-(1/3)s)\varepsilon})) \end{aligned}$$

and from here

$$\|x^n_{s+1}\| = o(n^{-6(mp+q-(1/3)s)\varepsilon}), \quad \omega \in S$$

Thus, (62) is valid on S, but (62) means

$$x^n \xrightarrow[n\to\infty]{} 0$$

this contradicts with $\|x^n\| = 1$, $\forall n$, hence $P(S) = 0$. This proves the lemma. □

Lemma 7

Assume $\|u_n\| = O(n^v)$, $0 \leqslant v < \varepsilon$. Then for $\{\tau_n\}$ defined by (57) the following estimate holds:

$$(\tau_n^{1-2\varepsilon} - \tau_{n-1}^{1-2\varepsilon})^{mp+q} = O((\tau_n^{1-\varepsilon+v} - \tau_{n-1}^{1-\varepsilon+v})^{mp+q-(1/2)}), \quad \text{a.s.} \tag{64}$$

Proof

Suppose (64) does not hold on a set Γ. We have to show

$$P(\Gamma) = 0$$

For $\omega \in \Gamma$, there is a subsequence $\{\tau_{n_k}\}$ such that

$$\frac{(\tau_{n_k}^{1-2\varepsilon} - \tau_{n_k-1}^{1-2\varepsilon})^{mp+q}}{(\tau_{n_k}^{1-\varepsilon+v} - \tau_{n_k-1}^{1-\varepsilon+v})^{mp+q-(1/2)}} \geqslant 2^k \tag{65}$$

Let $x^k = [x_1^{k\tau} \ldots x_{mp+q}^{k\tau}]^\tau$ be the unit minimum eigenvector of

$$\sum_{i=\tau_{n_k-1}+mp}^{\tau_{n_k}-2} U_i U_i^\tau, \quad \|x^n\| = 1, \quad x_i^n \in R^l$$

By definition of $\{\tau_n\}$ and (47), (48) with $b_k = 2^{k/(mp+q)}$, and $\beta_k \log \beta_k = 2^k$ we have

$$1 > \sum_{i=\tau_{n_k-1}+mp}^{\tau_{n_k}-2} (x^{k\tau} U_i)^2$$
$$\geqslant \sum_{i=\tau_{n_k-1}+mp}^{\tau_{n_k}-2} x_1^{k\tau} v_i v_i^\tau x_1^k + 2 \sum_{i=\tau_{n_k-1}+mp}^{\tau_{n_k}-2} x_1^{k\tau} v_i (u_i^{0\tau} x_1^k + u_{i-1}^\tau x_2^k + \ldots + x_{mp+q}^{k\tau} u_{i-mp-q+1})$$
$$\geqslant \beta \|x_1^k\|^2 (\tau_{n_k}^{1-2\varepsilon} - \tau_{n_k-1}^{1-2\varepsilon}) + o(\|x_1^k\| (\tau_{n_k}^{1-\varepsilon+v} - \tau_{n_k-1}^{1-\varepsilon+v})^{1/2} 2^k)$$

Similar to (60) and (61), we hence conclude

$$\|x_1^k\| = o\left(\frac{(\tau_{n_k}^{1-\varepsilon+v} - \tau_{n_k-1}^{1-\varepsilon+v})^{1/2} 2^k}{\tau_{n_k}^{1-2\varepsilon} - \tau_{n_k-1}^{1-2\varepsilon}}\right), \quad k \to \infty, \quad \omega \in \Gamma \tag{66}$$

Assume

$$\|x_i^k\| = o\left(\frac{(\tau_{n_k}^{1-\varepsilon+v} - \tau_{n_k-1}^{1-\varepsilon+v})^{i-1/2} \cdot 2^k}{(\tau_{n_k}^{1-2\varepsilon} - \tau_{n_k-1}^{1-2\varepsilon})^i}\right) \tag{67}$$

have been proved for i, $1 \leqslant i \leqslant s < mp + q$, we now prove it for $i = s + 1$, noticing (66) is only (67) with $i = 1$.

Paying attention to (40), (45) and $\|u_i\| = O(i^v)$ we see

$$\left\| \sum_{i=\tau_{n_k-1}+mp}^{\tau_{n_k}-2} v_{i-s}^\tau u_{i-j} \right\| = O(\tau_{n_k}^{1-\varepsilon+v} - \tau_{n_k-1}^{1-\varepsilon+v})$$

and then proceeding as for (63) and making use of (47) we have

$$
\begin{aligned}
1 > \beta\|x_{s+1}^k\|^2(\tau_{n_k}^{1-2\varepsilon}-\tau_{n_k-1}^{1-2\varepsilon}) \\
+ O(\|x_{s+1}^k\|(\|x_1^k\|+\ldots+\|x_s^k\|)(\tau_{n_k}^{1-\varepsilon+v}-\tau_{n_k-1}^{1-\varepsilon+v})) \\
+ o(\|x_{s+1}^k\|(\tau_{n_k}^{1-\varepsilon+v}-\tau_{n_k-1}^{1-\varepsilon+v})^{1/2}2^k)
\end{aligned} \tag{68}
$$

Noting the elementary inequality $a^y - b^y \leqslant a^x - b^x$, $\forall\, x \geqslant y \geqslant 0$, $a > b > 1$, by the induction assumption from (68) we find

$$
\begin{aligned}
1 > \beta\|x_{s+1}^k\|^2(\tau_{n_k}^{1-2\varepsilon}-\tau_{n_k-1}^{1-2\varepsilon}) + o\left(\|x_{s+1}^k\|\frac{(\tau_{n_k}^{1-\varepsilon+v}-\tau_{n_k-1}^{1-\varepsilon+v})^{s+1/2}\cdot 2^k}{(\tau_{n_k}^{1-2\varepsilon}-\tau_{n_k-1}^{1-2\varepsilon})^s}\right) \\
+ o(\|x_{s+1}^k\|(\tau_{n_k}^{1-\varepsilon+v}-\tau_{n_k-1}^{1-\varepsilon+v})^{1/2}\cdot 2^k)
\end{aligned}
$$

and from where it follows that (67) is valid for $i = s+1$ and $\omega \in \Gamma$. Further, from (65) we have

$$
\frac{(\tau_{n_k}^{1-\varepsilon+v}-\tau_{n_k-1}^{1-\varepsilon+v})^{mp+q-1/2}\cdot 2^k}{(\tau_{n_k}^{1-2\varepsilon}-\tau_{n_k-1}^{1-2\varepsilon})^{mp+q}} \leqslant 1
$$

then (67) says that

$$
\|x^k\| = o(1), \quad \text{as } k \to \infty, \quad \text{for } \omega \in \Gamma
$$

However, $\|x^k\| = 1$, $\forall\, k$, this means $P(\Gamma) = 0$. □

Theorem 2

For the system and algorithm described by (1)–(5) with attenuating excitation control defined by (41), if Condition (*a*) is satisfied and

$$
\|\varphi_n\| = O(n^v), \quad \text{with } v \in \left[0, \frac{1-(6+12(mp+q))\varepsilon}{12(mp+q)-2} \wedge \varepsilon\right) \tag{69}
$$

then

$$
\|\theta_n - \theta\| = O(\exp(-\alpha n^{1-12(mp+q)(\varepsilon+v)-6\varepsilon+2v})), \quad \text{a.s.} \tag{70}
$$

as $n \to \infty$ with $\alpha > 0$.

Proof

We note at once that for ε defined by (40) the interval for v is not empty and $1 - 12(mp+q)(\varepsilon+v) - 6\varepsilon + 2v > 0$.

The estimates

$$
\tau_n^{1-\varepsilon+v} - \tau_{n-1}^{1-\varepsilon+v} \leqslant 2(\tau_n^{1-2\varepsilon} - \tau_{n-1}^{1-2\varepsilon})\tau_n^{\varepsilon+v} \tag{71}
$$

$$
\tau_n - \tau_{n-1} \leqslant 2(\tau_n^{1-\varepsilon+v} - \tau_{n-1}^{1-\varepsilon+v})\tau_n^{\varepsilon-v} \tag{72}
$$

are derived from the inequality

$$
a^x - b^x \leqslant 2(a^y - b^y)a^{x-y}, \quad \forall\, a > b > 0, \quad 0 \leqslant y \leqslant x \leqslant 2y
$$

which comes from the identity

$$
a^x - b^x = (a^y - b^y)(a^{x-y} + b^{x-y}) + a^{x-y}b^y\left(1 - \left(\frac{a}{b}\right)^{2y-x}\right)
$$

By (64) and (71) we find

$$\tau_n^{1-\varepsilon+v} - \tau_{n-1}^{1-\varepsilon+v} = O((\tau_n^{1-\varepsilon+v} - \tau_{n-1}^{1-\varepsilon+v})^{(mp+q-(1/2))/(mp+q)}\tau_n^{\varepsilon+v}$$

hence

$$\tau_n^{1-\varepsilon+v} - \tau_{n-1}^{1-\varepsilon+v} = O(\tau_n^{2(mp+q)(\varepsilon+v)})$$

which together with (72) imply

$$\tau_n - \tau_{n-1} = O(\tau_n^{2(mp+q)(\varepsilon+v)+\varepsilon-v}) \tag{73}$$

Then we conclude that

$$\tau_n = \tau_0 + \sum_{i=1}^{n} (\tau_i - \tau_{i-1}) = O(n\tau_n^{2(mp+q)(\varepsilon+v)+\varepsilon-v})$$

and hence

$$\tau_n = O[n^{(1-2(mp+q)(\varepsilon+v)-\varepsilon+v)^{-1}}] \tag{74}$$

Putting (74) into (73) we finally obtain

$$\tau_n - \tau_{n-1} = O(n^{(2(mp+q)(\varepsilon+v)+\varepsilon-v)/(1-2(mp+q)(\varepsilon+v)-\varepsilon+v)})$$

Then the conclusion follows from (13) if we set $\delta = 0$ and

$$\lambda = \frac{2(mp+q)(\varepsilon+v)+\varepsilon-v}{1-2(mp+q)(\varepsilon+v)-\varepsilon+v}$$

4. Adaptive control

Let $\{y_n^*\}$ be an arbitrary bounded random reference signal. We want to design adaptive control so that the output y_n of the system (1) follows y_n^* and θ_n given by (4) converges to the true value.

We note that in the model reference adaptive control case y_n^* is generated by a reference model

$$A^*(z)y_n^* = B^*(z)u_n^*$$

with a monic matrix polynomial $A^*(z)$.

So

$$y_n^* = (I - A^*(z))y_n^* + B^*(z)u_n^*$$

and the problem is reduced to the previous one.

Write θ_n in component form

$$\theta_n^\tau = [-A_{1n} \dots -A_{pn} \quad B_{1n} \dots B_{qn}]$$

and form $A_n(z)$ and $B_n(z)$ as follows:

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

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

For any stable monic matrix polynomial $E(z)$ there are $G_n(z)$ and $F_n(z)$ such that

$$F_n(z)A_n(z) + z^d G_n(z) = E(z) \tag{75}$$

since $A_n(z)$ and $z^d I$ are coprime.

Define adaptive control

$$u_n = u_n^0 + v_n \tag{76}$$

with u_n^0 generated from

$$F_n(z)B_n(z)u_n^0 + G_n(z)y_n = E(z)y_{n+d}^* \tag{77}$$

and with $\{v_n\}$ given by (40), but we take $\{\varepsilon_n\}$ independent of $\{y_n^*\}$ and with continuous distribution.

It can be shown that B_{1n} is non-degenerate if $m = l$ (cf. Chen and Guo 1986 a). Hence u_n^0 can be defined from (77).

Theorem 3

For the system and algorithm (1)–(5) and control defined by (76) and (77), if Condition (*a*) holds and $B(z)$ is stable with $m = l$, then

(i) $\{y_n\}$ and $\{u_n\}$ are bounded a.s.

(ii) $\|y_n - y_n^*\| = O(M/n^{\varepsilon}) + O(\exp(-\alpha n^{1-12(mp+q)\varepsilon-6\varepsilon}))$, a.s.;

(iii) $\|\theta_n - \theta\| = O(\exp(-\alpha n^{1-12(mp+q)\varepsilon-6\varepsilon}))$, a.s.

where $\alpha > 0$ and ε and M are given in (39) and (40).

Proof

From (8) it is easy to see that

$$\|\tilde{\theta}_{n+1}\| \leqslant \|\tilde{\theta}_n\| \leqslant \|\tilde{\theta}_0\| < \infty \tag{78}$$

and

$$\operatorname{tr} \tilde{\theta}_{n+1}^{\tau}\tilde{\theta}_{n+1} \leqslant \operatorname{tr} \tilde{\theta}_n^{\tau}\tilde{\theta}_n - \frac{\|\tilde{\theta}_n^{\tau}\varphi_n\|^2}{1 + \|\varphi_n\|^2}$$

Thus we have

$$\sum_{i=0}^{\infty} \frac{\|\tilde{\theta}_i^{\tau}\varphi_i\|^2}{1 + \|\varphi_i\|^2} \leqslant \operatorname{tr} \tilde{\theta}_0^{\tau}\tilde{\theta}_0, \quad \text{a.s.}$$

and

$$\|\tilde{\theta}_i^{\tau}\varphi_i\|^2 = o(1 + \|\varphi_i\|^2) \quad \text{a.s.} \quad \text{as } i \to \infty \tag{79}$$

By (4) and (79) it is easy to see

$$\theta_{n+1} - \theta_n = o(1), \quad \text{a.s.}$$

and then

$$\theta_{n+k} - \theta_n = o(1), \quad \text{a.s.} \quad \text{as } n \to \infty \tag{80}$$

for any fixed integer $k \geqslant 1$.

We define polynomials $(AB)_n(z)$ and $(A_nB_n)(z)$ as follows:

$$(AB)_n(z) = \sum_{i,j} A_{in}B_{j(n-i)}z^{i+j}$$

$$(A_nB_n)(z) = \sum_{i,j} A_{in}B_{jn}z^{i+j}$$

and write $\tilde{\theta}_n^\tau \varphi_n$ as

$$\tilde{\theta}_n^\tau \varphi_n = y_{n+1} - \theta_n^\tau \varphi_n = A_n(z) y_{n+1} - B_n(z) u_{n-d+1}$$

Thus by using (75)–(77) we have

$$\begin{aligned} F_n(z)\tilde{\theta}_n^\tau \varphi_n &= (FA)_n(z) y_{n+1} - (FB)_n(z) u_{n-d+1} \\ &= (F_nA_n)(z) y_{n+1} + [(FA)_n(z) - (F_nA_n)(z)]\, y_{n+1} - (FB)_n(z) u_{n-d+1} \\ &= (E(z) - z^d G_n(z)) y_{n+1} + [(FA)_n(z) - (F_nA_n)(z)] y_{n+1} - (FB)_n(z) u_{n-d+1} \\ &= E(z) y_{n+1} - G_n(z) y_{n-d+1} - (F_nB_n)(z) u_{n-d+1} + [(FA)_n(z) - (F_nA_n)(z)]\, y_{n+1} \\ &\quad + [(F_nB_n)(z) - (FB)_n(z)] u_{n-d+1} \\ &= E(z) y_{n+1} - E(z) y^*_{n-1} - (F_nB_n)(z) v_{n-d+1} + [(FA)_n(z) - (F_nA_n)(z)]\, y_{n+1} \\ &\quad + [(F_nB_n)(z) - (FB)_n(z)] u_{n-d+1} \end{aligned} \tag{81}$$

Combining (1) and (81) we get

$$\begin{bmatrix} E(z) + [(FA)_n(z) - (F_nA_n)(z)] & [(F_nB_n)(z) - (FB)_n(z)] \\ \hline -A(z) & B(z) \end{bmatrix} \begin{bmatrix} y_{n+1} \\ \hline u_{n-d+1} \end{bmatrix} = \begin{bmatrix} F_n(z)\tilde{\theta}_n^\tau \varphi_n + E(z) y^*_{n+1} + (F_nB_n)(z) v_{n-d+1} \\ \hline 0 \end{bmatrix} \tag{82}$$

By (80) it is not difficult to see that

$$(FA)_n(z) - (F_nA_n)(z) \to 0, \quad (F_nB_n)(z) - (FB)_n(z) \to 0, \quad \text{as } n \to \infty$$

so (82) is asymptotically time-invariant and stable since $E(z)$ and $B(z)$ are both stable.

It is easy to convince oneself that the matrix coefficients in $F_n(z)$ and $G_n(z)$ are bounded since those of $A_n(z)$ are bounded by (78). Then by stability of $E(z)$ and $B(z)$ from (82) we know that

$$\|\varphi_{n+1}\|^2 = O(1) + O\left(\sup_{0 \le j \le n+1} \|\tilde{\theta}_j^\tau \varphi_j\|^2 \right)$$

and by (79)

$$\|\varphi_{n+1}\|^2 = O(1) + o\left(\sup_{0 \le j \le n+1} \|\varphi_j\|^2 \right)$$

Hence

$$\sup_{0 \le j \le n} \|\varphi_{j+1}\|^2 = O(1) + o\left(\sup_{0 \le j \le n} \|\varphi_{j+1}\|^2 \right)$$

which implies

$$\sup_{0 \le j \le n+1} \|\varphi_j\|^2 = O(1), \quad \text{a.s.} \quad \text{as } n \to \infty$$

This means that $\{y_n\}$ and $\{u_n\}$ are bounded. Then conclusion (iii) follows from Theorem 2 by setting $v = 0$, while (ii) follows from (81):

$$\begin{aligned} y_{n+1} - y^*_{n+1} = E^{-1}(z) \{ & F_n(z)\tilde{\theta}_n^\tau \varphi_n + (F_nB_n)(z) v_{n-d+1} \\ & - [(FA)_n(z) - (F_nA_n)(z)]\, y_{n+1} - [(F_nB_n)(z) - (FB)_n(z)] u_{n-d+1} \} \end{aligned}$$

if we use (iii) and (39), (40) and that $E(z)$ is stable and φ_n and matrix coefficients in $F_n(z)$ and $B_n(z)$ are bounded, and

$$(FA)_n(z) - (F_n A_n)(z) = O(\exp(-\alpha n^{1-12(mp+q)\varepsilon-6\varepsilon}))$$

$$(FB)_n(z) - (F_n B_n)(z) = O(\exp(-\alpha n^{1-12(mp+q)\varepsilon-6\varepsilon})) \qquad \square$$

Acknowledgment

This project was supported by the Science Fund of the Chinese Academy of Sciences.

References

Anderson, B. D. O., and Johnson, C. R., Jr., 1982, *Automatica*, **18**, 1.

Anderson, B. D. O., Bitmead, R. R., Johnson, C. R., Jr., Kokotovic, P. V., Kosut, R. L., Mareels, I. M. Y., Praly, L., and Riedle, B. D., 1986, *Stability of Adaptive Systems: Passivity and Averaging Analysis* (Boston, Mass.: MIT Press).

Åström, K. J., 1984, *Proc. 23rd I.E.E.E. Conf. Decision and Control*, 1276.

Bitmead, R. R., 1984, *I.E.E.E. Trans. Inf. Theory*, **30**, 183.

Chen, H.-F., and Guo, L., 1985 a, *SIAM Jl Control Optim.*; 1985 b, *Ibid.*; 1986 a, *Scientia Sinica* (*Series A*), **29**, 1145; 1986 b, *Int. J. Control*, **44**, 1459.

Chow, Y. S., Robbins, H., and Siegmund, D., 1971, *Great Expectations: the Theory of Optimal Stopping* (Boston: Houghton Mifflin).

Chow, Y. S., and Teicher, H., 1978, *Probability Theory* (New York: Springer-Verlag).

Goodwin, G. C., Raeadge, P. J., and Caines, P. E., 1980, *I.E.E.E. Trans. autom. Control*, **24**, 584.

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

Kosut, R. L., Anderson, B. D. O., and Mareels, I. M. Y., 1985, *Proc. 24th I.E.E.E. Conf. on Decision and Control*, p. 478 (see also) *I.E.E.E. Trans. autom. Control*, to be published.

Lipster, R. S., and Shiryayev, A. N., 1978, *Statistics of Random Processes*, Vol. I (New York: Springer-Verlag).