Reprinted from JOURNAL OF MULTIVARIATE ANALYSIS
All Rights Reserved by Academic Press. New York and London

Vol. 32 No. 1, January 1990
Printed in Belgium

# On ARX($\infty$) Approximation

L. GUO, D. W. HUANG, AND E. J. HANNAN*

*Australian National University, Canberra, ACT 2601, Australia*

*Communicated by the Editors*

Given data, $u_j, y_j, j=1, ..., n$, with $u_j$ an input sequence to a system while output is $y_j$, an approximation to the structure of the system generating $y_j$ is to be obtained by regressing $y_j$ on $u_{j-i}, y_{j-i}, i=1, ..., p_n$, where $p_n$ increases with $n$. In this paper the rate of convergence of the coefficient matrices to their asymptotic values is discussed. The context is kept general so that, in particular, $u_j$ is allowed to depend on $y_i, i \leqslant j$, and no assumption of stationarity for the $y_j$ or $u_j$ sequences is made. © 1990 Academic Press. Inc.

## 1. INTRODUCTION

A main concern of statisticians working in time series analysis and of some mathematicians in the fields of systems and control has been the study of methods for fitting linear systems to data. The simplest case is that of a regression

$$y_j = B^\tau u_j + w_j, \qquad j=1, 2, ..., \tag{1.1}$$

where we have used "$\tau$" for transposition. Here initially we might take the $u_j$ as a sure sequence. The "output" sequence, $y_j$ is $m$-dimensional, as is $w_j$, and $u_j$ is $l$-dimensional. If $w_j$ is measurable $\mathscr{F}_j$ and

$$E\{w_j \mid \mathscr{F}_{j-1}\} = 0, \qquad \sup_j E\|w_j\|^2 < \infty \tag{1.2}$$

then Lai, Robbins, and Wei [11] showed that least squares (LS) estimate,

$$\hat{B}_n = V_n^{-1} \sum_{i=1}^n u_i y_i^\tau, \qquad V_n = \sum_{i=1}^n u_i u_i^\tau$$

Received June 8, 1988

AMS 1980 subject classifications; 62M10, 60F15, 60G42, 93C40.

Key words and phrases; regression, autoregression, system, adaptive control.

* Visiting fellow from the Department of Probability and Statistics, Peking University.



0047-259X/90 $3.00

converges a.s. to $B$ if the smallest eigenvalue of $V_n$, $\lambda_{\min}(n)$, diverges to infinity.

In many applications one would wish $u_j$ to be taken as a stochastic sequence and measurable $\mathscr{F}_{j-1}$, as would be the case for design vector chosen on the basis of the data to time $j-1$. Lai and Wei [10] showed that if the second part of (1.2) is replaced by

$$\sup_j E[\|w_j\|^\alpha \mid \mathscr{F}_{j-1}] < \infty, \qquad \text{a.s.} \quad \alpha > 2$$

and $\lambda_{\min}(n) \longrightarrow \infty$, $\log \lambda_{\max}(n) = o(\lambda_{\min}(n))$, a.s., then again $\hat{B}_n \longrightarrow B$, a.s. This type of result enables, in particular, an ARX model to be considered, namely, one of the form

$$y_j = \sum_{i=1}^{p} A_{pi} y_{j-i} + \sum_{i=1}^{p} B_{pi} u_{j-i} + w_j. \tag{1.3}$$

In (1.1) $u_j$ will now be composed of the $y_{j-i}$, $u_{j-i}$, $i = 1 \cdots p$, from (1.3), but since (1.1) will not be referred to again, this should cause no confusion.

It is, however, hardly realistic to assume that (1.3) is the true process generating the data and it seems preferable, as in Hannan [6] to regard (1.3) as no more than a model on which an approximation procedure for the true structure is to be based. Then $p$, also, will be depend on the data size $n$ and we write $p_n$ for that. Thus it is necessary to investigate the nature of that approximation procedure. To begin with, we assume that

$$\begin{aligned} y_j &= \sum_{i=1}^{\infty} (A_i y_{j-i} + B_i u_{j-i}) + w_j, \qquad j \geq 0, \\ y_j &= w_j = 0, \qquad u_j = 0, \qquad j < 0. \end{aligned} \tag{1.4}$$

It is also required that

$$\sum_{i=1}^{\infty} (\|A_i\| + \|B_i\|) < \infty, \tag{1.5}$$

where the norm for a real matrix $X$ is defined as the maximum singular value of $X$, i.e., $\|X\| = \{\lambda_{\max}(XX^\tau)\}^{1/2}$, and the maximum (minimum) eigenvalue of a square matrix $X$ is denoted by $\lambda_{\max}(X)$ ($\lambda_{\min}(X)$).

In this paper the rate of convergence of the LS estimates of $A_{pi}$, $B_{pi}$ in (1.3), from $y_j$, $u_j$, $j = 1, \ldots, n$, to the $A_i$, $B_i$ in (1.4) is discussed, when $p = p_n$

increases with $n$. This LS approximation procedure we shall speak of as ARX($\infty$) approximation.

Before going on to that discussion, some further developments that relate to this will be mentioned. One problem with the procedure is the large number of parameters that may be estimated, especially, if $m$ is large, namely $p_n(m^2+ml)$, apart from any residual variance parameters. One way to overcome this problem is through further approximation procedures applied to a state space representation of a relation between $y_j$ and $u_j$. The procedure of balanced truncation or optimal Hankel norm approximation, discussed, for example, in Glover [3], could be applied to such a state space representation based on the LS estimation procedure to be studied here. This procedure would provide an ARMAX approximation to the true structure, that might be specified by many fewer parameters than that of ARX($\infty$) approximation.

One alternative procedure is to fit an ARMAX model directly to the data, allowing the order of the model (say the McMillan degree) to increase with $n$. One might, indeed, at least for the stationary case consider a criterion of the form of Rissanen [17, 18],

$$\log\det(\Sigma_\theta)+\dim(\theta)\log n/n. \tag{1.6}$$

Here $\Sigma_\theta$ is the covariance matrix of the innovation sequence, $w_j$, for a model specified by $\theta$ and $\dim(\theta)$ is the number of parameters in $\theta$. The criterion (1.6) might now be optimized, say by a Gauss–Newton procedure, over the class of models to be considered (see [7, Section 6.5], for example). Initial estimates of the $w_j$ will be needed for such a procedure and these may be obtained from the LS procedure of this paper. However, it would, in such a context, be necessary to determine $p_n$ from the data. That problem is not discussed here, but for its discussion the results of this paper are necessary preliminaries. Indeed, in a companion paper [5], we have applied the results here to the estimation problems of feedback control systems described by ARMAX models. It appears that the standard strictly positive real conditions used in engineering literature (e.g., [13, 19, 1, 16]) can be removed. This is, of course, only one possible application.

Let $z$ be the backwards-shift operator, and introduce

$$A[z]=-\sum_{i=0}^{\infty}A_iz^i \qquad (A_0=-I), \qquad B(z)=\sum_{i=1}^{\infty}B_iz^i \tag{1.7}$$

and denote the "transfer function" matrix associated with (1.4) as

$$G(z)=[A(z), B(z)]. \tag{1.8}$$

We will need the following two norms for measuring the accuracy of transfer function approximations:

$$\|F(z)\|_2 = \left\{\lambda_{\max}\left[\frac{1}{2\pi}\int_0^{2\pi} F(e^{i\theta})\, F^*(e^{i\theta})\, d\theta\right]\right\}^{1/2} \tag{1.9}$$

$$\|F(z)\|_\infty = \operatorname*{ess\,sup}_{(\theta \in [0,\, 2\pi])} \{\lambda_{\max}[F(e^{i\theta})\, F^*(e^{i\theta})]\}^{1/2}, \tag{1.10}$$

where the first is the $H_2$-norm of any measurable complex matrix $F(z)$ defined in $|z| \leqslant 1$, analytic in $|z| < 1$ and such that (1.9) is finite. The second is the $H^\infty$-norm of any complex matrix $F(z)$ which is analytic in $|z| < 1$ and bounded almost everywhere on the unit circle.

Throughout the paper, we assume that the system noise $\{w_n, \mathscr{F}_n\}$ is a martingale difference sequence with respect to a sequence $\{\mathscr{F}_n\}$ of non-decreasing $\sigma$-algebras, and that the input $u_n$ is a $\mathscr{F}_n$-measurable vector for any $n \geqslant 0$, i.e.,

$$E[w_{n+1} \mid \mathscr{F}_n] = 0, \qquad u_n \in \mathscr{F}_n,\ n \geqslant 0. \tag{1.11}$$

Clearly, system (1.4) under (1.5) and (1.11) is nonstationary in general because: (i) there are no restrictions on the location of the zeros of det $A(z)$; specifically, these zeros do not necessarily lie outside the closed unit circle and (ii) the system input sequence $\{u_n\}$ may be nonstationary.

Some results related to the estimation of scalar transfer functions were reported in Ljung [14], where it is required that the system (1.4) be stable in structure (or open-loop stable) and that all signals in (1.4) be stationary. For the more realistic nonstationary cases, however, to the best of our knowledge, there are hitherto no precise results available in the literature. This is perhaps due to the fact that the existing results on nonstationary ARX($p$), $p < \infty$, model cannot be immediately generalized to the present ARX($\infty$) case. Specifically, the standard martingale limit theorems and the stochastic Lyapunov functions which are so effective in the analysis of least squares algorithms for ARX($p$) model (see, e.g., [12, 15, 19, 10, 1]) cannot be directly used in the present ARX($\infty$) case. Instead, some limit theory on double array martingales and double array stochastic Lyapunov functions need to be established first in this case. Besides, we shall see in this paper that there are also considerable differences between finite lag regressors and increasing lag regressors in the convergence analysis.

In this paper, by considering the limit behaviors of double array martingales, we establish some general theorems on the approximation of nonstationary ARX($\infty$) models. In particular, the convergence rates of estimates for the unknown transfer matrix $G(z)$ defined by (1.8) are characterized in terms of $H^\infty$- as well as $H_2$-norms. The paper is organized as follows: In Section 2 we present the approximation/estimation algorithms, main theorems, and related observations; Section 3 focuses on establishing the asymptotic properties of some double array martingales; the main theorems are proved in Section 4; and Section 5 concludes the paper with some remarks.

## 2 Main Results

We first present the approximation algorithm

Let $\{p_n\}$ be any non-decreasing sequence of positive integers, $p_n \leqslant n$, $\forall n > 0$ Set

$$\theta(n) = [A_1 \cdots A_{p_n}, B_1 \cdots B_{p_n}]^\tau \tag{2.1}$$

and

$$\phi_i(n) = [y_i^\tau, \ldots, y_{i-p_n+1}^\tau, u_i^\tau, \ldots, u_{i-p_n+1}^\tau]^\tau, \qquad 1 \leqslant i \leqslant n, \tag{2.2}$$

The least-squares estimate $\hat{\theta}(n)$ for $\theta(n)$ at time $n$ is given by

$$\hat{\theta}(n) = \left[\sum_{i=0}^{n-1} \phi_i(n)\,\phi_i^\tau(n) + \gamma I\right]^{-1} \sum_{i=0}^{n-1} \phi_i(n)\, y_{i+1}^\tau \tag{2.3}$$

with real number $\gamma > 0$ arbitrarily chosen

Let us now write $\hat{\theta}(n)$ in its component form

$$\hat{\theta}(n) = [A_1(n), \ldots, A_{p_n}(n), B_1(n), \ldots, B_{p_n}(n)]^\tau \tag{2.4}$$

and set

$$\hat{A}_n(z) = I - \sum_{i=1}^{p_n} A_i(n)\, z^i, \qquad \hat{B}_n(z) = \sum_{i=1}^{p_n} B_i(n)\, z^i, \tag{2.5}$$

Then the estimate $\hat{G}_n(z)$ for $G(z)$ at time $n$ can now be formed as

$$\hat{G}_n(z) = [\hat{A}_n(z), \hat{B}_n(z)] \tag{2.6}$$

The convergence (and divergence) rates of $\hat{G}_n(z)$ are summarized in the following theorems

Theorem 2.1 *Consider the system* (1.4), (1.5), (1.11) *and the estimation algorithm* (2.2)–(2.6). *Suppose further that the random noise* $\{w_n\}$ *satisfies*

$$\sup_n E[\|w_{n+1}\|^4 \mid \mathscr{F}_n] < \infty, \qquad \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \|w_i\|^2 \neq 0, \qquad \text{a.s.} \tag{2.7}$$

*and*

$$\|w_n\| = O(\varphi(n)), \qquad a.s. \tag{2.8}$$

*with* $\{\varphi(n)\}$ *being a positive non-decreasing deterministic sequence. Then as* $n \longrightarrow \infty$,

$$\|\hat{G}_n(z) - G(z)\|_\infty^2 = O\left(\frac{p_n}{\lambda_{\min(n)}} \{p_n \log r_n + (\varphi(n) \log n)^{2+\varepsilon} + \delta_n r_n + (p_n \log r_n)^{(1/2)+\varepsilon} \varphi^2(n) \log n\}\right), \qquad a.s. \tag{2.9}$$

*holds for any* $\varepsilon > 0$, *where* $r_n$, $\delta_n$, *and* $\lambda_{\min}(n)$ *are defined by*

$$r_n \triangleq 1 + \sum_{i=0}^{n-1} (\|y_i\|^2 + \|u_i\|^2), \tag{2.10}$$

$$\delta_n \triangleq \left(\sum_{i=p_n+1}^{\infty} \|A_i\|\right)^2 + \left(\sum_{i=p_n+1}^{\infty} \|B_i\|\right)^2, \tag{2.11}$$

$$\lambda_{\min}(n) \triangleq \lambda_{\min}\left(\sum_{i=0}^{n-1} \phi_i(n)\, \phi_i^\tau(n) + \gamma I\right). \tag{2.12}$$

The proof of this theorem is given in Section 4.

*Remark* 2.1. If in (2.9) the $H^\infty$-norm is replaced by the $H_2$-norm, a better convergence rate can be obtained, namely, the term $p_n/\lambda_{\min}(n)$ in (2.9) can be replaced by $1/\lambda_{\min}(n)$. This can be easily seen from the proof of Theorem 2.1. Similar observations hold also for the following Theorem 2.2.

*Remark* 2.2. Note that in order to keep the generality of Theorem 2.1, we have tried to impose as few restrictions as possible. Of course, with some further conditions, "simple" formula may be deduced immediately from Theorem 2.1. For example, if in (1.1) the random disturbance $\{w_n\}$ is a Gaussian white noise (i.i.d.) sequence and

$$\|A_i\| + \|B_i\| = O(\lambda^i), \qquad 0 < \lambda < 1, \forall i \geq 0, \tag{2.13}$$

$$\sum_{i=0}^{n-1} (\|y_i\|^2 + \|u_i\|^2) = O(n^b), \qquad \text{a.s., for some } b \geq 1. \tag{2.14}$$

Then by taking $p_n = \log^a n$, $a > 1$, and noting $\|w_n\| = O(\{\log n\}^{1/2})$, we see from Theorem 2.1 that

$$\|\hat{G}_n(z) - G(z)\|_\infty^2 = O\left(\frac{\log^c n}{\lambda_{\min}(n)}\right), \qquad \text{a.s.,} \tag{2.15}$$

holds for any $c > \max\{1 + 2a, (3a+5)/2\}$.

Let us now consider the natural extension of the standard notion of "*persistence of excitation* (*PE*)" in the engineering literature (e.g., [15, 16, 19, 1]). It means that for $\lambda_{\min}(n)$ defined by (2.12),

$$\liminf_{n \longrightarrow \infty} \lambda_{\min}(n)/n \neq 0, \qquad \text{a.s.} \tag{2.16}$$

Hence under the "PE" condition, the convergence rate in (2.15) can be explicitly expressed as $O(\log^c n/n)$. The next theorem shows that this rate can be further improved if the growth rate of the observation data $\{y_i, u_i\}$ is not "too fast."

THEOREM 2.2. *Consider the system* (1.4), (1.5), (1.11) *and the estimation algorithm* (2.2)–(2.6). *Suppose further that for some* $\delta > 0$,

$$\sup_n E[\|w_n\|^2 \mid \mathscr{F}_{n-1}] < \infty; \quad \sup_n E\{\|w_n\|^4 (\log^+ \|w_n\|)^{2+\delta}\} < \infty, \tag{2.17}$$

*and that for some* $b \geq 1$,

$$E\{[\|y_n\| + \|u_n\|]^4 [\log^+(\|y_n\| + \|u_n\|)]^{2+\delta}\} = O(n^{2(b-1)}), \tag{2.18}$$

*and*

$$\sum_{i=0}^{n-1} (\|y_i\|^2 + \|u_i\|^2) = O(n^b), \qquad a.s., \tag{2.19}$$

*Then as* $n \longrightarrow \infty$,

$$\|\hat{G}_n(z) - G(z)\|_\infty^2 = \begin{cases} O\left(\dfrac{p_n}{\lambda_{\min}(n)}\{p_n \log\log[n^b/\lambda_{\min}(n)] + \delta_n n^b\}\right) & \text{if } p_n = O(\log^a n),\ a > 0 \\ O\left(\dfrac{p_n}{\lambda_{\min}(n)}\{p_n \log n[n^b/\lambda_{\min}(n)] + \delta_n n^b\}\right) & \text{if } p_n = O(n), \end{cases} \tag{2.20}$$

*where* $\delta_n$ *and* $\lambda_{\min}(n)$ *are defined by* (2.11) *and* (2.12), *respectively.*

The proof is given in Section 4

Thus, for example, if (2.18) and (2.19) hold with $b=1$, then under the "PE" condition (2.16) and the conditions in Remark 2.2, we have

$$\|\hat{G}_n(z)-G(z)\|_\infty^2 = O\left([p_n]^2\left\{\frac{\log\log n}{n}\right\}\right), \qquad \text{a.s.} \tag{2.21}$$

We remark that when (1.1) reduces to an ARX($p$), $p<\infty$, model, we may take the regression lag $p_n$ as $p$. In this case, (2.21) reads

$$\|\hat{G}_n(z)-G(z)\|_\infty = O\left(\left\{\frac{\log\log n}{n}\right\}^{1/2}\right), \qquad \text{a.s.} \tag{2.22}$$

From both Theorems 2.1 and 2.2, it is seen that the growth rate of $\lambda_{\min}(n)$ plays a crucial role in the convergence of the approximation algorithm. It is clear that $\lambda_{\min}(n)$ depends essentially on the two *input signals* $\{u_i, w_i\}$, even though it is defined via the observation data $\{y_i, u_i\}$. Let us now study how the growth rate of $\lambda_{\min}(n)$ depends explicitly on theses two *input signals*.

Set

$$\phi_i^0(n) = [u_i^\tau, \ldots, u_{i-2p_n+1}^\tau, w_i^\tau, \ldots, w_{i-2p_n+1}^\tau]^\tau, \qquad 1\leqslant i\leqslant n, \tag{2.23}$$

and denote

$$\lambda_{\min}^0(n) \triangleq \lambda_{\min}\left\{\sum_{i=0}^{n-1}\phi_i^0(n)\,\phi_i^{0\tau}(n)\right\}. \tag{2.24}$$

In Section 4 we shall prove the following result.

THEOREM 2.3. *Suppose that in system* (1.4)–(1.5), (1.7), *the number of zeros of* $\det A(z)$ *on the unit circle* $|z|=1$ *is finite (with the largest multiplicities of these zeros denoted by* $d$) *and that*

$$\sum_{k=1}^{\infty} k^{(1/2)+d}(\|A_k\|+\|B_k\|)<\infty. \tag{2.25}$$

*Then as* $n \longrightarrow \infty$,

$$\lambda_{\min}(n) \geqslant c_0(p_n)^{-2d}\,\lambda_{\min}^0(n) + O(\delta_n^0 r_n^0), \qquad \text{a.s.}, \tag{2.26}$$

*where* $c_o>0$ *is a constant,* $\lambda_{min}(n)$ *and* $\lambda_{\min}^0(n)$ *are respectively defined by* (2.12) *and* (2.24), *and*

$$r_n^0 = \sum_{i=0}^{n-1} (\|u_i\|^2 + \|w_i\|^2), \tag{2.27}$$

$$\delta_n^0 = p_n \left\{ \sum_{i=[p_n/2^m]-1}^{\infty} (\|A_i\| + \|B_i\|) \right\}^2 \qquad (m = \textit{dim of the output}) \tag{2.28}$$

Hence, for example, if either $r_n^0 = O(n)$ and $\liminf_{n \longrightarrow \infty} \lambda_{\min}^0(n)/n \neq 0$; or $r_n^0 = O(n^b)$, $b \geq 1$, $p_n = \log^a n$, $a > 1$, and (2.13) holds, then

$$\lambda_{\min}(n) \geq c_1(p_n)^{-2d} \lambda_{\min}^0(n), \qquad \text{a.s. for all } n \text{ and some } c_1 > 0. \tag{2.29}$$

(Note that if there is no zeros of det $A(z)$ on the unit circle then $d = 0$.)

As is seen from the above, the growth rate of $\lambda_{\min}(n)$ can be estimated by that of $\lambda_{\min}^0(n)$, which in turn is completely determined by the two exogenous signals $\{u_i, w_i\}$, *especially* the choice of the input sequence $\{u_i\}$.

We now give some examples to illustrate the growth rate of $\lambda_{\min}^0(n)$.

EXAMPLE 2.1. Suppose that in addition to (1.11) and (2.17), the input sequence $\{u_i\}$ and the noise sequence $\{w_i\}$ are independent and that

$$\liminf_{n \longrightarrow \infty} \lambda_{\min}\{E[w_n w_n^\tau \mid \mathscr{F}_{n-1}]\} > 0, \qquad \text{a.s.}, \; p_n = O(n^{1-(b/2)}/\log n), \tag{2.30}$$

$$E\{\|u_n\|^4 (\log^+ \|u_n\|)^{2+\delta}\} = O(n^{2(b-l)}), \qquad \sum_{i=0}^{n-1} \|u_i\|^2 = O(n^b), \text{ a.s.}, \tag{2.31}$$

where $b \in [1, 2)$ is a constant. Then there exists $\alpha > 0$ such that as $n \longrightarrow \infty$,

$$\lambda_{\min}^0(n) \geq \alpha \min\{n, \lambda_{\min}^1(n)\} + O(p_n \{n^b \log n\}^{1/2}), \tag{2.32}$$

where

$$\lambda_{\min}^1(n) \triangleq \lambda_{\min} \left\{ \sum_{i=0}^{n-1} \phi_i^1(n) \phi_i^{1\tau}(n) \right\} \tag{2.33}$$

$$\phi_i^1(n) = [u_i^\tau, u_{i-1}^\tau, \ldots, u_{i-2p_n+1}^\tau]^\tau. \tag{2.34}$$

*In particular*, if $u_n$ takes the form

$$u_n = u_n^0 + v_n, \tag{2.35}$$

where $\{u_i^0\}$ and $\{v_i\}$ are two independent sequences satisfying $E[v_n \mid \mathscr{F}_{n-1}] = 0$ and

$$0 < \liminf_{n \to \infty} \lambda_{\min}\{E[v_n v_n^\tau \mid \mathscr{F}_{n-1}]\} \leqslant \limsup_{n \to \infty} E[\|v_n\|^2 \mid \mathscr{F}_{n-1}] < \infty, \quad \text{a.s.} \tag{2.36}$$

$$E(\|u_n^0\| + \|v_n\|)^4 \{\log^+(\|u_n^0\| + \|v_n\|)\}^{2+\delta} = O(n^{2(b-1)}), \quad \text{a.s. } \delta > 0. \tag{2.37}$$

$$\sum_{i=0}^{n-1} \|u_n^0\|^2 = O(n^b), \qquad \text{a.s.} \tag{2.38}$$

Then

$$\liminf_{n \to \infty} \lambda_{\min}^0(n)/n \neq 0, \qquad \text{a.s.} \tag{2.39}$$

The proof of this example is also given in Section 4. We remark that *any* bounded deterministic sequence $\{u_i^0\}$ in (2.35) satisfies the required conditions.

Let us now consider another example where the two exogenous sequences $\{u_i\}$ and $\{w_i\}$ may be correlated.

EXAMPLE 2.2 Suppose that $\eta_i \triangleq [u_i^\tau \; w_i^\tau]^\tau$ is a stationary sequence with spectral density matrix uniformly positive definite and with autocovariances satisfying

$$\max_{(0 \leqslant t,\, k \leqslant 2p_n)} \left\| \frac{1}{n} \sum_{i=0}^{n-1} \{\eta_{i-k}\eta_{i-t}^\tau - E[\eta_{i-k}\eta_{i-t}^\tau]\} \right\| = o([p_n]^{-1}), \qquad \text{a.s.} \tag{2.40}$$

Then with $\lambda_{\min}^0(n)$ defined by (2.24),

$$\liminf_{n \to \infty} \lambda_{\min}^0(n)/n \neq 0, \qquad \text{a.s.} \tag{2.41}$$

*Proof.* Let us denote

$$\phi_i^\eta(n) = [\eta_i^\tau, \ldots, \eta_{i-2p_n+1}^\tau]^\tau; \tag{2.42}$$

then by (2.40) we know that

$$\begin{aligned} &\left\| \frac{1}{n} \sum_{i=0}^{n-1} \{\phi_i^\eta(n)\, \phi_i^{\eta\tau}(n) - E[\phi_0^\eta(n)\, \phi_0^{\eta\tau}(n)\} \right\| \\ &\qquad \leqslant 4p_n \max_{(0 \leqslant k,\, t < 2p_n)} \left\| \frac{1}{n} \sum_{i=0}^{n-1} \{\eta_{i-k}\eta_{i-t}^\tau - E[\eta_{i-k}\eta_{i-t}^\tau]\} \right\| = o(1), \\ &\qquad\qquad \text{as } n \to \infty. \end{aligned}$$

But, by the standard relation between the autocovariance matrix and the spectral density function, it is easy to see that

$$\inf_n \lambda_{\min}\{E[\phi_0^\eta(n)\,\phi_0^{\eta\tau}(n)\}>0$$

Consequently, by noting that there is an orthogonal matrix $T_n$ such that $\phi_i^\eta(n)=T_n\phi_i^0(n)$, we get

$$\begin{aligned}\liminf_{n\to\infty}\lambda_{\min}^0(n)/n&=\liminf_{n\to\infty}\lambda_{\min}\left(\frac{1}{n}\sum_{i=0}^{n-1}\{\phi_i^\eta(n)\,\phi_i^{\eta\tau}(n)\}\right)\\&=\liminf_{n\to\infty}\lambda_{\min}\{E[\phi_0^\eta(n)\,\phi_0^{\eta\tau}(n)]\}>0,\qquad\text{a.s.}\quad\blacksquare\end{aligned}$$

We conclude this section by pointing out that some results which hold in ARX($p$), $p<\infty$, case, may not hold in ARX($\infty$) case. For instance, it can be shown that (see [2, Theorem 3], for related proofs) if $u_n$ takes the form $u_n=u_n^0+v_n$, with $\{v_n\}$ and $\{w_n\}$ independent and satisfying conditions in Example 2.1, but with $u_n^0$ being any measurable vectors such that

$$u_n^0\in\sigma\{w_i,v_{i-1},i\leqslant n\}\qquad\text{and}\qquad\sum_{i=0}^{n-1}\|u_i^0\|^2=O(n),\qquad\text{a.s.;}\quad(2.43)$$

then for $\lambda_{\min}^0(n)$ defined by (2.24) with $p_n=p<\infty$, $\forall n$, the assertion (2.39) holds.

In the ARX($\infty$) case, however, this conclusion does not hold, in general. In other words, the requirement for $\{u_n^0\}$ in Example 2.1 cannot be simply replaced by (2.43). This can be illustrated as follows. Take $u_n^0=v_{n-1}$, $\forall n\geqslant 1$, $p_n\longrightarrow\infty$, $p_n=O(\log^a n)$, $a\geqslant 1$, and assume that $\{v_n,-\infty<n<\infty\}$ is a scalar i.i.d. sequence with zero mean and suitably high moment. Then, $u_n=v_{n-1}+v_n$, and by Lemma 3.5 in Section 3, we have under (2.34),

$$\begin{aligned}&\left\|\frac{1}{n}\sum_{i=0}^{n-1}\{\phi_i^1(n)\,\phi_i^{1\tau}(n)-E[\phi_0^1(n)\,\phi_0^{1\tau}(n)]\}\right\|\\&\quad\leqslant 2p_n\max_{(0\leqslant k\ t<2p_n)}\left|\frac{1}{n}\sum_{i=0}^{n-1}\{u_{i-k}u_{i-t}-E[u_{i-k}u_{i-t}]\}\right|\\&\quad=O(p_n\{\log\log n/n\}^{1/2})\\&\quad=o(1),\qquad\text{as }n\longrightarrow\infty.\end{aligned}$$

On the other hand, by a result in Grenander and Szego [4, pp. 147] we know that

$$\lambda_{\min}\{E[\phi_0^1(n)\,\phi_0^{1\tau}(n)\}\xrightarrow[n\to\infty]{}\inf_{(\lambda\in[-\pi,\pi])}|1+e^{i\lambda}|^2=0.$$

Consequently, by (2.24) and (2.33),

$$\begin{aligned}\liminf_{n\to\infty}\lambda_{\min}^0(n)/n&\leqslant\liminf_{n\to\infty}\lambda_{\min}^1(n)/n\\&\leqslant\liminf_{n\to\infty}\lambda_{\min}\{E[\phi_0^1(n)\,\phi_0^{1\tau}(n)\}\\&\quad+\limsup_{n\to\infty}\left\|\frac{1}{n}\sum_{i=0}^{n-1}\{\phi_i^1(n)\,\phi_i^{1\tau}(n)-E[\phi_0^1(n)\,\phi_0^{1\tau}(n)]\}\right\|\\&=0,\qquad \text{a.s.}\end{aligned}$$

So (2.39) fails.

## 3. Limit Behaviors of Double Array Martingales

In this section, we study the limit behaviors of some double array martingales. These results not only form an indispensable part of the proof of our main results, but also are interesting of themselves.

Lemma 3.1. *Suppose that $\{w_n, \mathscr{F}_n\}$ is an m-dimensional martingale difference sequence satisfying*

$$\sup_j E[\|w_{j+1}\|^2 \mid \mathscr{F}_j]<\infty,\qquad \|w_n\|=o(\varphi(n)),\qquad a.s., \ \varphi(n)\leqslant\varphi(n+1),\ \forall n, \tag{3.1}$$

*and that for any $1\leqslant i\leqslant n$ and $n\geqslant 1$, $f_{in}$ is $\mathscr{F}_i$-measurable $p\times m$-dimensional random matrix satisfying*

$$\sum_{i=0}^{n}\|f_{in}\|^2\leqslant A<\infty,\qquad a.s., \forall n \tag{3.2}$$

*where $\varphi(n)$ and $A$ positive and deterministic numbers. Then, as $n\to\infty$,*

$$\max_{1\leqslant i\leqslant n}\left\|\sum_{j=1}^{i}f_{jn}w_{j+1}\right\|=o(\varphi(n+1)\log n),\qquad a.s. \tag{3.3}$$

*Proof.* Clearly, we need only to consider the one-dimensional case, and without loss of generality we may assume $\|f_{in}\| \leq 1$. For any $\varepsilon > 0$, let us set

$$w_j' = w_j I_{[|w_j| \leq \varepsilon\varphi(j)]}, \qquad \bar{w}_j = w_j' - E[w_j' \mid \mathscr{F}_{j-l}], \tag{3.4}$$

then we have

$$\left| \sum_{j=1}^{i} f_{jn} w_{j+1} \right|$$
$$\leq \left| \sum_{j=1}^{i} f_{jn}(w_{j+1} - w_{j+1}') \right| + \left| \sum_{j=1}^{i} f_{jn} E[w_{j+1}' \mid \mathscr{F}_j] \right| + \left| \sum_{j=1}^{i} f_{jn} \bar{w}_{j+1} \right| \tag{3.5}$$

Note that by (3.1),

$$\max_{1 \leq i \leq n} \left| \sum_{j=1}^{i} f_{jn}(w_{j+1} - w_{j+1}') \right| \leq \max_{1 \leq i \leq n} \sum_{j=1}^{i} |w_{j+1}| I_{[|w_{j+1}| > \varepsilon\varphi(j+1)]}$$
$$\leq o(\varphi(n+1)) \sum_{j=1}^{\infty} I_{[|w_{j+1}| > \varepsilon\varphi(j+1)]} = o(\varphi(n+1)), \qquad \text{a.s.}, \tag{3.6}$$

and by (3.1), (3.2), and the Schwarz inequality,

$$\max_{1 \leq i \leq n} \left| \sum_{j=1}^{i} f_{jn} E[w_{j+1}' \mid \mathscr{F}_j] \right|^2$$
$$\leq \max_{1 \leq i \leq n} A \sum_{j=1}^{i} |E\{w_{j+1} I_{[|w_{j+1}| \leq \varepsilon\varphi(j+1)]} \mid \mathscr{F}_j\}|^2$$
$$= \max_{1 \leq i \leq n} A \sum_{j=1}^{i} |E\{w_{j+1} I_{[|w_{j+1}| > \varepsilon\varphi(j+1)]} \mid \mathscr{F}_j\}|^2$$
$$\leq A \sup_j E[\|w_{j+1}\|^2 \mid \mathscr{F}_j] \max_{1 \leq i \leq n} \sum_{j=1}^{i} P\{|w_{j+1}| > \varepsilon\varphi(j+1) \mid \mathscr{F}_j\}$$
$$= O\left( \sum_{j=1}^{\infty} P\{|w_{j+1}| > \varepsilon\varphi(j+1) \mid \mathscr{F}_j\} \right) = O(1), \qquad \text{a.s.}, \tag{3.7}$$

where the last inequality holds because by (3.1) and the conditional Borel–Cantelli lemma [20, p. 55],

$$P\left\{ \sum_{j=1}^{\infty} P[|w_{j+1}| > \varepsilon\varphi(j+1) \mid \mathscr{F}_j] = \infty \right\} = P\{|w_{j+1}| > \varepsilon\varphi(j+1), \text{ i.o.}\} = 0.$$

Hence, for (3.3) we need only to consider the last term on the RHS of (3.5). Let us set

$$S_i(n)=\sum_{j=1}^{i} f_{jn}\bar{w}_{j+1},\qquad S_0(n)=0,\qquad 1\leqslant i\leqslant n,$$

$$c_n=2\varepsilon\varphi(n+1),\qquad \lambda_n=(c_n)^{-1}$$

$$T_i(n)=\exp\{\lambda_n S_i(n)\}\exp\left\{-(\lambda_n^2|2)\left[1+\frac{\lambda_n c_n}{2}\right]\sum_{j=1}^{i}(f_{jn})^2 E[(\bar{w}_{j+1})^2\mid\mathscr{F}_j]\right\},$$

$$T_0(n)=1,\qquad 1\leqslant i\leqslant n$$

then by (3.4) and Lemma 5.4.1 in Stout [20], we know that for any fixed $n$ $\{T_i(n),\ 0\leqslant i\leqslant n\}$ is a nonnegative supermartingale. Consequently, by Corollary 5.4.1 in Stout [20] we have

$$\begin{aligned}
P\{\max_{0\leqslant i\leqslant n} S_i(n)>2c_n\log n\} &= P\{\max_{0\leqslant i\leqslant n}\exp[\lambda_n S_i(n)]>\exp[2c_n\lambda_n\log n]\}\\
&\leqslant P\left[\max_{0\leqslant i\leqslant n} T_i(n)>\exp\left\{2\log n-(\lambda_n^2/2)\left[1+\frac{\lambda_n c_n}{2}\right]A(c_n)^2\right\}\right]\\
&\leqslant \exp\left\{-2\log n+\frac{A}{2}\left(1+\frac{1}{2}\right)\right\}=\frac{1}{n^2}\exp\left\{\frac{3A}{4}\right\}.
\end{aligned}$$

So by the Borel–Cantelli lemma,

$$\limsup_{n\to\infty}\max_{1\leqslant i\leqslant n} S_i(n)/\varphi(n+1)\log n\leqslant 4\varepsilon,\qquad \text{a.s.}$$

Similar results also hold with $\{S_i(n)\}$ replaced by $\{-S_i(n)\}$. Hence from here and (3.5)–(3.7),

$$\limsup_{n\to\infty}\max_{1\leqslant i\leqslant n}\left|\sum_{j=1}^{i} f_{jn}w_{j+1}\right|\Big/\varphi(n+1)\log n\leqslant 4\varepsilon,$$

and therefore the desired result (3.3) follows by the arbitrariness of $\varepsilon$. ∎

**Lemma 3.2.** *Let $\{f_{in}\}$ and $\{w_i\}$ be two random sequences satisfying conditions in Lemma 3.1 except (3.2). If $\{a_{in}, 0\leqslant i\leqslant n, n\geqslant 1\}$ is a positive random sequence such that*

$$a_{in}\in\mathscr{F}_i,\qquad a_{(i-1)n}\leqslant a_{in},\qquad \forall 1\leqslant i\leqslant n, \forall n\geqslant 1,$$

*and*

$$\sum_{i=1}^{n}\left(\frac{\|f_{in}\|}{a_{in}}\right)^2 \leqslant A < \infty, \qquad \forall n \geqslant 1,$$

*where $A$ is a deterministic constant. Then as $n \longrightarrow \infty$,*

$$\left\|\sum_{j=1}^{n} f_{jn} w_{j+1}\right\| = o(a_{nn}\varphi(n+1)\log n), \qquad a.s.$$

*Proof.* Set

$$x_{in} \triangleq \sum_{j=1}^{i} \frac{f_{jn} w_{j+1}}{a_{jn}}, \qquad 1 \leqslant i \leqslant n,\ x_{on} = 0.$$

By Lemma 3.1, we have

$$\max_{1 \leqslant i \leqslant n} \|x_{in}\| = o(\varphi(n+1)\log n), \qquad \text{a.s.}$$

Consequently,

$$\begin{aligned}
\left\|\sum_{i=1}^{n} f_{in} w_{i+1}\right\| &= \left\|\sum_{i=1}^{n} a_{in}[x_{in} - x_{(i-1)n}]\right\| \\
&= \left\|a_{nn}x_{nn} - a_{1n}x_{on} - \sum_{i=2}^{n}[a_{in} - a_{(i-1)n}]\,x_{(i-1)n}\right\| \\
&\leqslant \|a_{nn}x_{nn}\| + \max_{1 \leqslant i \leqslant n-1} \|x_{in}\| \sum_{i=2}^{n}[a_{in} - a_{(i-1)n}] \\
&= O(a_{nn} \max_{1 \leqslant i \leqslant n} \|x_{in}\|) = o(a_{nn}\varphi(n+1)\log n), \qquad \text{a.s.}
\end{aligned}$$

This completes the proof. ▮

LEMMA 3.3. *Suppose that $\{w_n, \mathscr{F}_n\}$ is an $m$-dimensional martingale difference sequence satisfying*

$$\sup_j E[\|w_{j+1}\|^2 \,|\mathscr{F}_j] < \infty \qquad \text{and} \qquad \|w_n\| = O(\varphi(n)), \qquad a.s.,$$

*where $\{\varphi(n)\}$ is a nondecreasing positive deterministic sequence, and that $f_{in}$ is any $\mathscr{F}_i$-measurable $p \times m$ random matrix for $1 \leqslant i \leqslant n$, $n \geqslant 1$. Then, for any*

$\delta \in (\frac{1}{2}, 1)$, *there exists a function* $a(\delta) > 0$ *such that* $a(\delta) \downarrow 2$ *as* $\delta \longrightarrow 1$, *and that as* $n \longrightarrow \infty$,

$$\left\| \sum_{j=1}^{n} f_{jn} w_{j+1} \right\| = o\left( \left\{ \sum_{j=1}^{n} \|f_{jn}\|^2 \right\}^{\delta} \right) + o(\{\varphi(n+1) \log n\}^{a(\delta)}), \quad a.s. \quad (3.8)$$

*Proof.* For any $\delta \in (\frac{1}{2}, 1)$, let us denote

$$a_{in} \triangleq \left\{ 1 + \sum_{j=1}^{i} \|f_{jn}\|^2 \right\}^{(1/2) + b(\delta)}, \qquad a_{on} = 1, \qquad 1 \leqslant i \leqslant n,$$

where

$$b(\delta) \triangleq \frac{(2\delta - 1)(1 - \delta)}{1 + (2\delta - 1)^2} > 0.$$

It is easy to verify that

$$\sup_{n \geqslant 1} \sum_{i=1}^{n} \left( \frac{\|f_{in}\|}{a_{in}} \right)^2 \leqslant \frac{1}{2b(\delta)} < \infty, \qquad \text{a.s.}$$

Hence by Lemma 3.2 and the following inequality,

$$|xy| \leqslant \frac{1}{p} |x|^p + \frac{1}{q} |y|^q, \qquad \frac{1}{p} + \frac{1}{q} = 1, \qquad p > 0, q > 0,$$

we know that

$$\begin{aligned} &\left\| \sum_{j=1}^{n} f_{jn} w_{j+1} \right\| \\ &\quad = o(a_{nn}\{\varphi(n+1) \log n\}^{2-\delta}) \\ &\quad = o(\{a_{nn}\}^{1+(2\delta-1^2)}) + o(\{\varphi(n+1) \log n\}^{(2-\delta)[1+(2\delta-1)^2]}/(2\delta-1)^2). \end{aligned}$$

Consequently, by setting

$$a(\delta) = \frac{(2-\delta)[1 + (2\delta - 1)^2]}{(2\delta - 1)^2},$$

we see that $\lim_{\delta \longrightarrow 1} a(\delta) = 2$, and that (3.8) holds. ∎

We remark that in contrast to the estimation developed for standard martingale case [10], here in Lemma 3.3, the second term on the RHS of

(3.8), i.e., $o(\{\varphi(n+1)\log n\}^{a(\delta)})$, cannot be removed in general. A simple example for this is

$$f_{in}=\begin{cases}1 & \text{if } i=n\\ 0 & \text{if } i<n,\end{cases}$$

together with $\{w_n\}$ being any unbounded random sequence satisfying conditions in Lemma 3.3.

We may now prove the following result.

LEMMA 3.4. *Let $\{w_n, \mathscr{F}_n\}$ satisfy conditions in Theorem 2.1, and $f_{in}$, $1\leqslant i\leqslant n$, be any $p_n$-dimensional and $\mathscr{F}_i$-measurable random vector sequence. Then for any $\varepsilon>0$, as $n\longrightarrow\infty$,*

$$\begin{aligned}\sum_{j=1}^{n} f_{in}^{\tau}(M_{in})^{-1}f_{in}\|w_{i+1}\|^2 &= O(p_n\log^+\lambda_{\max}(M_{nn}))+o(\{p_n\log[e+\lambda_{\max}(M_{nn})]\}^{(1/2)+\varepsilon}\\ &\quad\times[\log^{1+\varepsilon}n]\,\varphi^2(n+1)),\qquad a.s.\end{aligned}\tag{3.9}$$

*where $M_{in}$ is defined as*

$$M_{in}=\sum_{j=1}^{i} f_{jn}f_{jn}^{\tau}+\gamma I,\qquad M_{0n}=\gamma I,\ (\gamma>0),\qquad 1\leqslant i\leqslant n.\tag{3.10}$$

*Proof.* Let $|X|$ denote the determinant of a square matrix $X$. Then following Lai and Wei [10, Lemma 2], we know from (3.10) that for any $1\leqslant i\leqslant n$,

$$|M_{(i-1)n}|=|M_{in}-f_{in}f_{in}^{\tau}|=|M_{in}|\,|I-(M_{in})^{-1}f_{in}f_{in}^{\tau}|=|M_{in}|[1-f_{in}^{\tau}(M_{in})^{-1}f_{in}].$$

Then, we have

$$\sum_{i=1}^{n} f_{in}^{\tau}(M_{in})^{-1}f_{in}=\sum_{i=1}^{n}\frac{|M_{in}|-|M_{(i-1)n}|}{|M_{in}|}\leqslant\int_{|M_{0n}|}^{|M_{nn}|}\frac{dx}{x}\leqslant\log^+(|M_{nn}|)-\log\gamma,\tag{3.11}$$

and

$$\begin{aligned}&\sum_{i=1}^{n}\{f_{in}^{\tau}(M_{in})^{-1}f_{in}/\log^{1/2+\varepsilon}(|M_{in}|+e)\}^2\\ &\quad\leqslant\int_{|M_{0n}|}^{|M_{nn}|}\frac{dx}{x\log^{1+2\varepsilon}(x+e)}\leqslant\frac{1}{2\varepsilon\log^{2\varepsilon}(e+\gamma)}.\end{aligned}$$

Consequently, by (3.11) and Lemma 3.2 we get

$$\sum_{i=1}^{n} f_{in}^{\tau}(M_{in})^{-1} f_{in} \|w_{i+1}\|^2$$

$$= \sum_{i=1}^{n} f_{in}^{\tau}(M_{in})^{-1} f_{in} E[\|w_{i+1}\|^2 |\mathscr{F}_i]$$

$$+ \sum_{i=1}^{n} f_{in}^{\tau}(M_{in})^{-1} f_{in}\{\|w_{i+1}\|^2 - E[\|w_{i+1}\|^2| \mathscr{F}_i]\}$$

$$= O(\log^+|M_{nn}|) + o(\{\log(|M_{nn}|+e)\}^{(1/2)+\varepsilon}[\log^{1+\varepsilon} n]\, \varphi^2(n+1)).$$

Finally, the desired result (3.9) follows by noting that the number of distinct eigenvalues of $M_{nn}$ is not great than $p_n$. ∎

We now consider another type of double array martingales.

LEMMA 3.5. *Suppose that $\{w_n, \mathscr{F}_n\}$ is a vector martingale difference sequence satisfying*

$$\sup_n E[\|w_{n+1}\|^2|\mathscr{F}_n] < \infty, \qquad \sup_n E\{\|w_n\|^4(\log^+ \|w_n\|)^{2+\delta}\} < \infty, \tag{3.12}$$

*and that $\{x_n, \mathscr{F}_n\}$ is any adapted random vector sequence satisfying*

$$\sum_{i=1}^{n} \|x_i\|^2 = O(n^b), \qquad E\|x_n\|^4\{\log^+(\|x_n\|)\}^{2+\delta} = O(n^{2(b-1)}), \tag{3.13}$$

*where $\delta > 0$, $b \geq 1$ are some constants. Then as $n \longrightarrow \infty$,*

$$\max_{1 \leq t \leq (\log n)^a} \max_{1 \leq i \leq n} \left\| \sum_{j=1}^{i} x_{j-t} w_j^{\tau} \right\| = O(n^{b/2}\{\log\log n\}^{1/2}), \qquad \text{a.s.}, \forall > 0, \tag{3.14}$$

$$\max_{1 \leq t \leq n} \max_{1 \leq i \leq n} \left\| \sum_{j=1}^{i} x_{j-t} w_j^{\tau} \right\| = O(n^{b/2}\{\log n\}^{1/2}), \qquad \text{a.s.} \tag{3.15}$$

*Proof.* The conditions and conclusions of this lemma are slightly different from those in Huang [9, Lemma 1]. However, similar proof techniques can be applied. Hence we just give a *sketch* proof here.

Note that we need only to consider the one-dimensional case. Set

$$c_n = (n/\log\log n)^{1/2}$$

$$\tilde{w}_n = w_n I_{[|w_n| \leq (c_n)^{1/2}]}, \qquad \tilde{\tilde{w}}_n = w_n - \tilde{w}_n$$

$$\tilde{x}_n = x_n I_{[|x_n| \leq n^{(b-1)/2}(c_n)^{1/2}]}, \qquad \tilde{\tilde{x}}_n = x_n - \tilde{x}_n$$

$$w_t(j) = \tilde{x}_{j-t}\tilde{w}_j - E[\tilde{x}_{j-t}\tilde{w}_j | \mathscr{F}_{j-1}], \qquad S_t(i) = \sum_{j=1}^{i} w_t(j).$$

Then, by (3.12) and (3.13) it can be shown that

$$\max_{1 \leqslant t \leqslant (\log n)^a} \sum_{j=1}^{n} |x_{j-t} w_{j-t} w_j - w_t(j)| / n^{b/2} (\log \log n)^{1/2} \xrightarrow[n \to \infty]{} 0, \qquad \text{a.s.}$$

and by Corollary 5.4.1 in Stout [20], we can prove that

$$\limsup_{k \to \infty} \max_{1 \leqslant t \leqslant (\log 2^k)^a} \max_{1 \leqslant i \leqslant 2^k} H_{2^k}(i, t)/(2^{kb} \log \log 2^k)^{1/2} < \infty, \text{ a.s.},$$

where

$$H_n(i, t) = \left| \pm S_t(i) - \frac{3[n^b \log \log n]^{1/2}}{8n^b} \sum_{j=1}^{i} E\{[w_t(j)]^2 \mid \mathscr{F}_{j-t}\} \right|$$

Then (3.14) follows from (3.12), (3.13), and the fact that $H_n(i, t)$ is a monotonic function of $n$. While (3.15) can be proved by taking $c_n = (n/\log n)^{1/2}$, and following a similar argument as that used in Huang [8, Lemma 1]. ∎

We now introduce the function

$$V_i(n) \triangleq \left\| \left( \sum_{j=1}^{i} f_{jn} f_{jn}^\tau + \gamma I \right)^{-(1/2)} \sum_{j=1}^{i} f_{jn} w_{j+1}^\tau \right\|^2, \qquad \gamma > 0, 1 \leqslant i \leqslant n, n \geqslant 1. \tag{3.16}$$

This function is obviously a natural extension of the standard stochastic Lyapunov function frequently used in the literature (e.g., [15, 19, 10, 1]), Hence we may call it a "*double array stochastic Lyapunov function.*"

LEMMA 3.6. *Let $\{w_i\}$ and $\{f_{in}, 1 \leqslant i \leqslant n\}$, $n \geqslant 1$, be any $p$- and $p_n$-dimensional random sequences, respectively. Then the "Lyapunov function" defined by (3.16) has the properties*

(*i*) $V_n(n) \leqslant \sum_{j=1}^{n} \|w_{j+1}\|^2, \forall n \geqslant 1.$

(ii) $V_n(n) = O(p_n \log^+ \lambda_{\max}(M_{nn})) + o(\{\varphi(n+1) \log n\}^{2+\varepsilon} + o(\{p_n \log[e + \lambda_{\max}(M_{nn})]\}^{(1/2)+\varepsilon} [\log^{1+\varepsilon} n]\, \varphi^2(n+1)), \forall \varepsilon > 0$, *provided that* $\{f_{in}\}$ *and* $\{w_n\}$ *satisfy conditions in Lemma* 3.4.

(iii)

$$V_n(n) = \begin{cases} O\left( \dfrac{1}{\lambda_{\min}(M_{nn})} \{p_n n^b \log \log n\} \right) & \text{if } p_n = O(\log^a n), a > 0 \\ O\left( \dfrac{1}{\lambda_{\min}(M_{nn})} \{p_n n^b \log n\} \right) & \text{if } p_n = O(n) \end{cases}$$

*provided that* $f_{in} = [x_i^\tau, x_{i-1}^\tau, \ldots, x_{i-p_n}^\tau]^\tau$, *and that* $\{x_i\}$ *and* $\{w_i\}$ *satisfy conditions in Lemma* 3.5, *where* $M_{nn}$ *is defined as in* (3.10).

*Proof.* By (3.10) and the matrix inverse formula it is clear that

$$\begin{aligned} M_{in}^{-1} &= M_{(i-1)n}^{-1} - b_{(i-1)n} M_{(i-1)n}^{-1} f_{in} f_{in}^{\tau} M_{(i-1)n}^{-1}, \\ b_{(i-1)n} &\triangleq (1 + f_{in}^{\tau} M_{(i-1)n}^{-1} f_{in})^{-1}, \end{aligned} \tag{3.17}$$

hence by denoting

$$S_i(n) = \sum_{j=1}^{i} f_{jn} w_{j+1}^{\tau}, \qquad S_0(n) = 0, \qquad 1 \leqslant i \leqslant n,$$

we see that (omitting the dependence on $n$ of the variables) for any $1 \leqslant i \leqslant n$,

$$\begin{aligned} &\operatorname{tr}\{S_i^{\tau} M_i^{-1} S_i\} \\ &\quad = \operatorname{tr}\{(S_{i-1} + f_i w_{i+1}^{\tau})^{\tau} [M_{i-1}^{-1} - b_{i-1} M_{i-1}^{-1} f_i f_i^{\tau} M_{i-1}^{-1}](S_{i-1} + f_i w_{i+1}^{\tau})\} \\ &\quad = \operatorname{tr}\{S_{i-1}^{\tau} M_{i-1}^{-1} S_{i-1}\} + 2b_{i-1} w_{i+1}^{\tau} S_{i-1}^{\tau} M_{i-1}^{-1} f_i \\ &\qquad - b_{i-1} \|S_{i-1}^{\tau} M_{i-1}^{-1} f_i\|^2 + b_{i-1} f_i^{\tau} M_{i-1}^{-1} f_i \|w_{i+1}\|^2 \end{aligned}$$

and then summing up from 1 to $n$,

$$\begin{aligned} V_n(n) &\leqslant \operatorname{tr}\{S_n^{\tau} M_n^{-1} S_n\} \\ &= \sum_{i=1}^{n} \{2b_{i-1} w_{i+1}^{\tau} S_{i-1}^{\tau} M_{i-1}^{-1} f_i - b_{i-1} \|S_{i-1}^{\tau} M_{i-1}^{-1} f_i\|^2 \\ &\quad + b_{i-1} f_i^{\tau} M_{i-1}^{-1} f_i \|w_{i+1}\|^2\}. \end{aligned} \tag{3.18}$$

Hence conclusion (i) follows by noting $2|w_{i+1}^{\tau} S_{i-1}^{\tau} M_{i-1}^{-1} f_i| \leqslant \|S_{i-1}^{\tau} M_{i-1}^{-1} f_i\|^2 + \|w_{i+1}\|^2$ and the definition of $b_{i-1}$. To prove (ii), we apply Lemma 3.3 by chosing $\delta \in (\frac{1}{2}, 1)$ such that $a(\delta) < 2 + \varepsilon$, to estimate the first term on the RHS of (3.18),

$$\begin{aligned} \sum_{i=1}^{n} b_{i-1} w_{i+1}^{\tau} S_{i-1}^{\tau} M_{i-1}^{-1} f_i &= o\left(\left\{\sum_{i=1}^{n} b_{i-1} \|S_{i-1}^{\tau} M_{i-1}^{-1} f_i\|^2\right\}^{\delta}\right) \\ &\quad + o(\{\varphi(n+1) \log n\}^{2+\varepsilon}). \end{aligned} \tag{3.19}$$

The third term on the RHS of (3.18) can be estimated by using Lemma 3.4, since (3.17) implies $b_{i-1} f_i^{\tau} M_{i-1}^{-1} f_i = f_i^{\tau} M_i^{-1} f_i$. Hence conclusion (ii) follows from (3.18), (3.19), and Lemma 3.4. While conclusion (iii) may be directly deduced from Lemma 3.5, since it means

$$\left\| \sum_{j=1}^{n} f_{jn} w_{j+1}^{\tau} \right\|^2 = \begin{cases} O(\{p_n n^b \log\log n\}) & \text{if } p_n = O(\log^a n),\ a > 0 \\ O(\{p_n n^b \log n\}) & \text{if } p_n = O(n). \end{cases}$$

Hence the proof of Lemma 3.6 is completed. ∎

## 4. Proof of the Main Results

We are now in a position to prove the results listed in Section 2.

*Proof of Theorems* 2.1 *and* 2.2. Let us denote

$$G_n(z) \triangleq [A_n(z), B_n(z)], \qquad A_n(z) = I - \sum_{i=1}^{p_n} A_i z^i, \qquad B_n(z) = \sum_{i=1}^{p_n} B_i z^i$$

Then by (1.7), (1.8), (2.1), (2.4)–(2.6), and (2.11), we know that

$$\begin{aligned}
\|\hat{G}_n(z) - G(z)\|_\infty^2 &\leqslant 2\,\|\hat{G}_n(z) - G_n(z)\|_\infty^2 \pm 2\,\|G_n(z) - G(z)\|_\infty^2 \\
&\leqslant 2\left\{\sum_{i=1}^{p_n} \|[A_i(n) - A_i, B_i(n) - B_i]\|\right\}^2 + 2\left\{\sum_{i=p_n+1}^{\infty} (\|A_i\| + \|B_i\|)\right\}^2 \\
&\leqslant 2p_n \operatorname{tr}\left\{\sum_{i=1}^{p_n} [A_i(n) - A_i, B_i(n) - B_i][A_i(n) - A_i, B_i(n) - B_i]^\tau\right\} + 4\delta_n \\
&= 2p_n \operatorname{tr}\{[\hat{\theta}(n) - \theta(n)]^\tau [\hat{\theta}(n) - \theta(n)]\} + 4\delta_n \\
&= 2m\, p_n \|\hat{\theta}(n) - \theta(n)\|^2 + 4\delta_n. 
\end{aligned} \tag{4.1}$$

Set

$$\varepsilon_i(n) = \sum_{j=p_n+1}^{\infty} [A_j y_{i-j+1} + B_j u_{i-j+1}]; \tag{4.2}$$

then by (1.4), (2.1)–(2.3), and (2.12),

$$\begin{aligned}
&\|\hat{\theta}(n) - \theta(n)\|^2 \\
&= \left\| \left[\sum_{i=0}^{n-1} \phi_i(n)\,\phi_i^\tau(n) + \gamma I\right]^{-1} \left\{\sum_{i=0}^{n-1} \phi_i(n)[y_{i+1}^\tau - \phi_i^\tau(n)\,\theta(n)] - \gamma\theta(n)\right\}\right\|^2 \\
&= \left\| \left[\sum_{i=0}^{n-1} \phi_i(n)\,\phi_i^\tau(n) + \gamma I\right]^{-1} \left\{\sum_{i=0}^{n-1} \phi_i(n)[w_{i+1}^\tau + \varepsilon_i^\tau(n)] - \gamma\theta(n)\right\}\right\|^2 \\
&\leqslant \frac{3}{\lambda_{\min}(n)} \left\{ \left\| \left[\sum_{i=0}^{n-1} \phi_i(n)\,\phi_i^\tau(n) + \gamma I\right]^{-1/2} \sum_{i=0}^{n-1} \phi_i(n)\, w_{i+1}^\tau \right\|^2 \right. \\
&\quad \left. + \left\| \left[\sum_{i=0}^{n-1} \phi_i(n)\,\phi_i^\tau(n) + \gamma I\right]^{-1/2} \sum_{i=0}^{n-1} \phi_i(n)\, \varepsilon_i^\tau(n) \right\|^2 + O(1) \right\}.
\end{aligned} \tag{4.3}$$

Now for the second term on the RHS of (4.3) we may first apply Lemma 3.6(i) and then apply the Schwarz inequality; thus under (2.10), (2.11), and (4.2),

$$\left\|\left[\sum_{i=0}^{n-1} \phi_i(n)\,\phi_i^\tau(n)+\gamma I\right]^{-1/2} \sum_{i=0}^{n-1} \phi_i(n)\,\varepsilon_i^\tau(n)\right\|^2 \leqslant \sum_{i=0}^{n-1} \|\varepsilon_i(n)\|^2$$

$$\leqslant 2\sum_{i=0}^{n-1}\left\{\sum_{j=p_n+1}^{\infty} \|A_j\| \sum_{j=p_n+1}^{\infty} \|A_j\|\,\|y_{i-j+1}\|^2 \pm \sum_{j=p_n+1}^{\infty} \|B_j\| \sum_{j=p_n+1}^{\infty} \|B_j\|\,\|u_{i-j+1}\|^2\right\}$$

$$\leqslant 2\left(\sum_{j=p_n+1}^{\infty} \|A_j\|\right)^2 \sum_{i=0}^{n-1} \|y_i\|^2 + 2\left(\sum_{j=p_n+1}^{\infty} \|B_j\|\right)^2 \sum_{i=0}^{n-1} \|u_i\|^2 \leqslant 2\delta_n r_n \tag{4.4}$$

As for the first term on the RHS of (4.3) we may use Lemma 3.6(ii) and (iii) to estimate it. Hence Theorem 2.2 follows directly from (4.1), (4.3), (4.4), and Lemma 3.6(iii). While to conclude Theorem 2.1 from Lemma 3.6 (ii), we need only to note that by (1.4), (2.7), (2.10), and the Schwarz inequality ($A_0 \triangleq I$, $B_0 \triangleq 0$),

$$0 \neq \liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \|w_i\|^2 \leqslant \liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \left\{\sum_{j=0}^{\infty} (\|A_j y_{i-j}\| + \|B_j u_{i-j}\|)\right\}^2$$

$$\leqslant 2\left\{\sum_{j=0}^{\infty} (\|A_j\| + \|B_j\|)\right\}^2 \liminf_{n\to\infty} r_n/n,$$

and then that

$$\log^+ \lambda_{\max}\left\{\sum_{i=0}^{n-1} \phi_i(n)\,\phi_i^\tau(n)+\gamma I\right\} = O(\log^+\{p_n r_n + p_n\}) = O(\log r_n), \qquad \text{a.s.}$$

Hence the proof of Theorems 2.1 and 2.2 is completed. ∎

We now proceed to prove Theorem 2.3. For this, we need to prove several lemmas first.

LEMMA 4.1. *Let $f_i(z)$, $1 \leqslant i \leqslant p$, be analytic functions expanded as*

$$f_i(z) = \sum_{j=0}^{\infty} f_j^{(i)} z^j, \qquad 1 \leqslant i \leqslant p,$$

*and let the product of $f_i(i)$ be expanded as*

$$\prod_{i=1}^{p} f_i(z) = \sum_{j=0}^{\infty} c_i z^j.$$

*Then,*

(i) $\sum_{j=0}^{\infty}|f_j^{(i)}|<\infty$, $1\leqslant i\leqslant p$, *implies that* $\sum_{j=n}^{\infty}|c_j|=O(\sum_{j=[n/2^{p-1}]-1}^{\infty}\sum_{i=1}^{p}|f_j^{(i)}|)$;

(ii) $\sum_{j=0}^{\infty}j^r|f_j^{(i)}|<\infty$, $r\geqslant 0$, $1\leqslant i\leqslant p$, *implies that* $\sum_{j=0}^{\infty}j^r|c_j|<\infty$.

*Proof.* (i) Let us first consider the case of $p=2$. In this case,

$$c_i=\sum_{j=0}^{i}f_j^{(1)}f_{i-j}^{(2)} \tag{4.5}$$

and then

$$\begin{aligned}\sum_{i=n}^{\infty}|c_i|&\leqslant\sum_{i=n}^{\infty}\sum_{j=0}^{i}|f_j^{(1)}|\,|f_{i-j}^{(2)}|=\sum_{j=0}^{n}\sum_{i=n}^{\infty}|f_j^{(1)}|\,|f_{i-j}^{(2)}|+\sum_{j=n}^{\infty}\sum_{j=j}^{\infty}|f_j^{(1)}|\,|f_{j-j}^{(2)}|\\&\leqslant\sum_{j=0}^{[n/2]-1}|f_j^{(1)}|\sum_{i=n}^{\infty}|f_{i-j}^{(2)}|+\sum_{j=[n/2]}^{n}|f_j^{(1)}|\sum_{i=n}^{\infty}|f_{i-j}^{(2)}|+O\left(\sum_{j=n}^{\infty}|f_j^{(1)}|\right)\\&\leqslant O\left(\sum_{i=[n/2]}^{\infty}|f_i^{(2)}|\right)+O\left(\sum_{j=[n/2]}^{\infty}|f_j^{(1)}|\right)+O\left(\sum_{j=n}^{\infty}|f_j^{(1)}|\right).\end{aligned}$$

Thus

$$\sum_{i=n}^{\infty}|c_i|=O\left(\sum_{i=[n/2]}^{\infty}(|f_i^{(1)}|+|f_i^{(2)}|)\right).$$

This proves the case of $p=2$. The general $p>2$ case can be proved by induction via the above relationship.

(ii) By induction we need only to prove the case of $p=2$. Note that $i^r\leqslant 2^r\{(i-j)^r+j^r\}$, $\forall i\geqslant j\geqslant 0$; we have by (4.5),

$$\begin{aligned}\sum_{i=0}^{\infty}i^r|c_i|&\leqslant\sum_{i=0}^{\infty}2^r\{(i-j)^r+j^r\}\sum_{j=0}^{i}|f_j^{(1)}|\,|f_{i-j}^{(2)}|\\&\leqslant 2r\sum_{j=0}^{\infty}\sum_{i=j}^{\infty}\{(i-j)^r+j^r\}|f_j^{(1)}|\,|f_{i-j}^{(2)}|\\&=2^r\sum_{j=0}^{\infty}|f_j|\sum_{i=j}^{\infty}(i-j)^r|f_{i-j}^{(2)}|+2^r\sum_{j=0}^{\infty}j^r|f_j^{(1)}|\sum_{i=j}^{\infty}|f_{i-j}^{(2)}|\\&=2^r\left(\sum_{j=0}^{\infty}|f_j^{(1)}|\right)\sum_{i=0}^{\infty}i^r|f_i^{(2)}|+2^r\left(\sum_{j=0}^{\infty}j^r|f_j^{(1)}|\right)\sum_{i=0}^{\infty}|f_i^{(2)}|<\infty.\quad\blacksquare\end{aligned}$$

For any $x_n\in R^{(m+l)p_n}$ with

$$\|x_n\|=1,\qquad\forall n\geqslant 1, \tag{4.6}$$

let us write it in its component form

$$x_n = [\alpha_n^{(o)\tau}, \ldots, \alpha_n^{(p_n-1\tau}, \qquad \beta_n^{(o)\tau}, \ldots, \beta_n^{(p_n-1)\tau}]^\tau, \tag{4.7}$$

with $v_n^{(i)\tau} = R^m$, $\beta_n^{(i)\tau} \in R^l$, and introduce the vector complex function

$$H_n(z) \triangleq \sum_{i=0}^{p_n-1} \{\alpha_n^{(i)\tau} [\text{Adj } A(z)][B(z), I] + \beta_n^{(i)\tau} [\det A(z)\, I_l, 0)\} z^i \tag{4.8}$$

$$\triangleq \sum_{i=0}^{\infty} [h_n^{(i)\tau}, g_n^{(i)\tau}] z^i \tag{4.9}$$

Obviously, $h_n^{(i)}$ and $g_n^{(i)}$ are functions of $x_n$.

LEMMA 4.2. *Under the conditions of Theorem 2.3 and the denotation* (4.6)–(4.9),

$$\liminf_{n\to\infty} \inf_{\|x\|=1} p_n^{2d} \sum_{i=0}^{2p_n-1} (\|h_n^{(i)}\|^2 + \|g_n^{(i)}\|^2) \neq 0, \qquad a.s.$$

*Proof.* Denote

$$\det A(z) = \sum_{i=0}^{\infty} a_i z^i, \qquad \text{Adj } A(z) = \sum_{i=0}^{\infty} \bar{A}_i z^i, \qquad [\text{Adj } A(z)]\, B(z) = \sum_{i=1}^{\infty} \bar{B}_i z^i; \tag{4.10}$$

then by Lemma 4.1 (ii) and (2.25) it is easy to convince oneself that

$$\sum_{i=0}^{\infty} i^{(1/2)+d}(\|\bar{A}_i\| + \|\bar{B}_i\| + |a_i|) < \infty. \tag{4.11}$$

Note that by (4.8)–(4.10),

$$\sum_{i=0}^{p_n-1} \{\alpha_n^{(i)\tau}[\text{Adj } A(z)]\, B(z) + \beta_n^{(i)\tau}[\det A(z)]\} z^i = \sum_{i=0}^{\infty} h_n^{(i)\tau} z^i, \tag{4.12}$$

$$\sum_{i=0}^{p_n-1} \alpha_n^{(i)\tau}[\text{Adj } A(z)]\, z^i = \sum_{i=0}^{\infty} g_n^{(i)\tau} z^i \tag{4.13}$$

which imply that

$$h_n^{(i)\tau} = \sum_{j=0}^{i} [\alpha_n^{(j)\tau} \bar{B}_{i-j} + \beta_n^{(j)\tau} a_{i-j}], \qquad g_n^{(i)\tau} = \sum_{j=0}^{i} \alpha_n^{(j)\tau} \bar{A}_{i-j}, \qquad \forall i \geq 0, \tag{4.14}$$

where by definition $\alpha_n^{(i)\tau}=0$, $\beta_n^{(i)\tau}=0$, $\forall i\geqslant p_n, n\geqslant 1$. Therefore by (4.6), (4.7),

$$\begin{aligned}\sum_{i=2p_n}^{\infty}\|h_n^{(i)}\|^2 &\leqslant 2\sum_{i=2p_n}^{\infty}\left\{\left(\sum_{j=0}^{p_n-1}\|\alpha_n^{(j)}\|\,\|\bar{B}_{i-j}\|\right)^2+\left(\sum_{j=0}^{p_n-1}\|\beta_n^{(j)}\|\,|a_{i-j}|\right)^2\right\}\\ &\leqslant 2\sum_{i=2p_n}^{\infty}\left\{\sum_{j=0}^{p_n-1}(\|\alpha_n^{(j)}\|^2+\|\beta_n^{(j)}\|^2)\right\}\left\{\sum_{j=0}^{p_n-1}(\|\bar{B}_{i-j}\|^2+|a_{i-j}|^2)\right\}\\ &=2\sum_{i=2p_n}^{\infty}\sum_{j=0}^{p_n-1}(\|\bar{B}_{i-j}\|^2+|a_{i-j}|^2)\leqslant 2\sum_{j=0}^{p_n-1}\sum_{i=p_n+1}^{\infty}(\|\bar{B}_i\|^2+|a_i|^2)\\ &=2p_n\left\{\sum_{i=p_n+1}^{\infty}(\|\bar{B}_i\|+|a_i|)\right\}^2.\end{aligned}$$

Consequently, by (4.11) we know that as $n\longrightarrow\infty$,

$$\sup_{\|x_n\|=1} p_n^{2d}\sum_{i=2p_n}^{\infty}\|h_n^{(i)}\|^2\leqslant 2\left\{\sum_{i=p_n+1}^{\infty}i^{d+(1/2)}(\|\bar{B}_i\|+|a_i|)\right\}^2\longrightarrow 0.$$

Similarly,

$$\sup_{\|x_n\|=1} p_n^{2d}\sum_{i=2p_n}^{\infty}\|g_n^{(i)}\|^2\longrightarrow 0,\qquad \text{as } n\longrightarrow\infty.$$

Hence, for this lemma it suffices to show that

$$\liminf_{n\longrightarrow\infty}\ \inf_{\|x_n\|=1} p_n^{2d}\sum_{i=0}^{\infty}(\|h_n^{(i)}\|^2+\|g_n^{(i)}\|^2)\neq 0,\qquad \text{a.s.} \tag{4.15}$$

If (4.15) were not true, we would find a set $D$ with $P(D)>0$, and a subsequence of $\{n\}$ (without loss of generality assume this sequence is also $\{n\}$), such that

$$p_n^{2d}\sum_{i=0}^{\infty}(\|h_n^{(i)}\|^2+\|g_n^{(i)}\|^2)\longrightarrow 0,\qquad \forall\omega\in D, \text{ as } n\longrightarrow\infty. \tag{4.16}$$

Now, substituting (4.13) into (4.12) we see that

$$\sum_{i=0}^{p_n-1}\beta_n^{(i)\tau}z^i[\det A(z)]=\sum_{i=0}^{\infty}[h_n^{(i)\tau}-g_n^{(i)\tau}B(z)]\,z^i, \tag{4.17}$$

and noting $[\operatorname{Adj} A(z)]\,A(z)=[\det A(z)]\,I$, we have by (4.13),

$$\sum_{i=0}^{p_n-1}\alpha_n^{(i)\tau}z^i[\det A(z)]=\sum_{i=0}^{\infty}g_n^{(i)\tau}z^iA(z). \tag{4.18}$$

Let $e^{j\theta_i}$, $i=1, \ldots, s$, $j^2 \triangleq -1$, $\theta_i \in [0, 2\pi]$, be distinct zeros of $\det A(z)$ on the unit circle $|z|=1$, and let their multiplicities be $d_1, \ldots, d_s$. Then we have $d=\max\{d_1, \ldots, d_s\}$, and

$$\det A(e^{j\theta}) = f(e^{j\theta})(e^{j\theta}-e^{j\theta_1})^{d_1} \cdots (e^{j\theta}-e^{j\theta_s})^{d_s}, \qquad f(e^{j\theta}) \neq 0, \theta \in [0, 2\pi]. \tag{4.19}$$

Without loss of generality assume that

$$\theta_0 \triangleq 0 < \theta_1 < \theta_2 < \cdots < \theta_s < 2\pi \triangleq \theta_{s+1}. \tag{4.20}$$

Then it is easy to convince oneself from (4.19) that there exists a constant $c_1 > 0$ such that

$$\min_{(k \in [0, s])} \min_{(\theta \in [\theta_k + \varepsilon, \theta_{k+1} - \varepsilon])} |\det A(e^{j\theta})|^2 \geqslant c_1 \varepsilon^{2d}, \tag{4.21}$$

holds for all appropriately small $\varepsilon > 0$.

Thus, by (4.18) and (4.21) it follows that for any small $\varepsilon > 0$ ($j^2 \triangleq -1$),

$$\begin{aligned}
\sum_{i=0}^{p_n-1} \|\alpha_n^{(i)}\|^2 &= \frac{1}{2\pi} \int_0^{2\pi} \left\| \sum_{i=0}^{p_n-1} \alpha_n^{(i)\tau} e^{ij\theta} \right\|^2 d\theta \\
&= \frac{1}{2\pi} \sum_{k=0}^{s} \int_{\theta_k + \varepsilon/p_n}^{\theta_{k+1} - \varepsilon/p_n} \left\| \sum_{i=0}^{p_n-1} \alpha_n^{(i)\tau} e^{ij\theta} \right\|^2 d\theta \\
&\quad + \frac{1}{2\pi} \left\{ \int_{\theta_0}^{\theta_0 + \varepsilon/p_n} + \sum_{k=1}^{s} \int_{\theta_k - \varepsilon/p_n}^{\theta_k + \varepsilon/p_n} + \int_{\theta_{s+1} - \varepsilon/p_n}^{\theta_{s+1}} \right\} \\
&\quad \times \left\| \sum_{i=0}^{p_n-1} \alpha_n^{(i)\tau} e^{ij\theta} \right\|^2 d\theta \\
&\leqslant \frac{p_n^{2d}}{2\pi c_1 \varepsilon^{2d}} \sum_{k=0}^{s} \int_{\theta_k + \varepsilon/p_n}^{\theta_{k+1} - \varepsilon/p_n} \left\| \sum_{i=0}^{p_n-1} \alpha_n^{(i)\tau} e^{ij\theta} [\det A(e^{j\theta})] \right\|^2 d\theta \\
&\quad + \frac{1}{2\pi} \left\{ 2(s+1) \frac{\varepsilon}{p_n} \right\} p_n \sum_{i=0}^{p_n-1} \|\alpha_n^{(i)}\|^2 \\
&\leqslant \frac{p_n^{2d}}{2\pi c_1 \varepsilon^{2d}} \int_0^{2\pi} \left\| \sum_{i=0}^{\infty} g_n^{(i)\tau} e^{ij\theta} A(e^{j\theta}) \right\|^2 d\theta + (s+1)\varepsilon/\pi \\
&\leqslant \frac{p_n^{2d} \|A(e^{j\theta})\|_\infty^2}{2\pi c_1 \varepsilon^{2d}} F_0^{2\pi} \left\| \sum_{i=0}^{\infty} g_n^{(i)\tau} e^{ij\theta} \right\|^2 d\theta + (s+1)\varepsilon/\pi \\
&= \frac{\|A(e^{i\theta})\|_\infty^2}{c_1 \varepsilon^{2d}} \left\{ p_n^{2d} \sum_{i=0}^{\infty} \|g_n^{(i)}\|^2 \right\} + (s+1)\varepsilon/\pi.
\end{aligned}$$

Consequently, it follows from (4.16) that

$$\limsup_{n\to\infty} \sum_{i=0}^{p_n-1} \|\alpha_n^{(i)}\|^2 \leqslant (s+1)\,\varepsilon/\pi, \qquad \forall\, \omega \in D$$

and hence the arbitrariness of $\varepsilon$ yields

$$\limsup_{n\to\infty} \sum_{i=0}^{p_n-1} \|\alpha_n^{(i)}\|^2 = 0, \qquad \forall\, \omega \in D. \tag{4.22}$$

Similarly, by (4.16), (4.17), and (4.21), it can be shown that

$$\limsup_{n\to\infty} \sum_{i=0}^{p_n-1} \|\beta_n^{(i)}\|^2 = 0, \qquad \forall\, \omega \in D. \tag{4.23}$$

Finally, combining (4.6), (4.7), (4.22), and (4.23) we get the contradiction

$$1 = \|x_n\|^2 = \sum_{i=0}^{p_n-1} (\|\alpha_n^{(i)}\|^2 + \|\beta_n^{(i)}\|) \xrightarrow[n\to\infty]{} 0, \qquad \forall\, \omega \in D. \tag{4.24}$$

Hence, the assertion (4.15) holds. This completes the proof of the lemma. ∎

LEMMA 4.3. *With $h_n^{(i)}$ and $g_n^{(i)}$ defined by* (4.9),

$$\sum_{i=0}^{n-1} \left\{ \sum_{j=2p_n}^{\infty} [h_n^{(j)\tau} u_{i-j} + g_n^{(j)\tau} w_{i-j}] \right\}^2 = O(\delta_n^0 r_n^0), \qquad a.s.,$$

*where $\delta_n^0$ and $r_n^0$ are given by* (2.28) *and* (2.27), *respectively.*

*Proof.* By (4.10), (4.14), and Lemma 4.1(i) it follows that

$$\begin{aligned}\sum_{i=2p_n}^{\infty} \|g_n^{(i)}\| &\leqslant \sum_{j=0}^{p_n-1} \sum_{i=2p_n}^{\infty} \|\alpha_n^{(j)}\| \, \|\bar{A}_{i-j}\| \leqslant \sum_{j=0}^{p_n-1} \|\alpha_n^{(j)}\| \sum_{i=p_n}^{\infty} \|\bar{A}_i\| \\ &= O\left( [p_n]^{1/2} \sum_{i=[p_n/2^m]-1}^{\infty} \|A_i\| \right),\end{aligned} \tag{4.25}$$

and similarly,

$$\sum_{i=2p_n}^{\infty} \|g_n^{(i)}\| = O\left( [p_n]^{1/2} \sum_{i=[p_n/2^m]-1}^{\infty} [\|A_i\| + \|B_i\|] \right). \tag{4.26}$$

Hence, by (4.26) and the Schwarz inequality,

$$\sum_{i=0}^{n-1}\left\{\sum_{j=2p_n}^{\infty} h_n^{(j)\tau} u_{i-j}\right\}^2$$
$$\leqslant \sum_{i=0}^{n-1}\sum_{j=2p_n}^{\infty} \|h_n^{(j)}\| \sum_{j=2p_n}^{\infty} \|h_n^{(j)}\| \, \|u_{i-j}\|^2$$
$$\leqslant \left\{\sum_{j=2p_n}^{\infty} \|h_n^{(j)}\|\right\}^2 \sum_{i=0}^{n-1} \|u_i\|^2$$
$$= O\left(p_n \sum_{i=[p_n/2^m]-1}^{\infty} [\|A_i\|+\|B_i\|]^2 \sum_{i=0}^{n-1}\|u_i\|^2\right) = O(\delta_n^0 r_n^0)_n. \quad (4.27)$$

Similarly, by (4.25),

$$\sum_{i=0}^{n-1}\left\{\sum_{j=2p_n}^{\infty} g_n^{(j)\tau} w_{i-j}\right\}^2 = O(\delta_n^0 r_n^0). \quad (4.28)$$

Finally, the desired result follows from (4.27) and (4.28). ∎

*Proof of Theorem* 2.3. Let us define, under (2.2) and (4.10),

$$\psi_i(n) \triangleq [\det A(z)]\, \phi_i(n). \quad (4.29)$$

Then we have

$$\lambda_{\min}\left\{\sum_{i=0}^{n-1} \psi_i(n)\,\psi_i^\tau(n)\right\} = \inf_{\|x\|=1} \sum_{i=0}^{n-1} [x^\tau \psi_i(n)]^2$$
$$= \inf_{\|x\|=1} \sum_{i=0}^{n-1}\left[\sum_{j=0}^{\infty} a_j x^\tau \phi_{i-j}(n)\right]^2$$
$$\leqslant \left(\sum_{j=0}^{\infty}|a_j|\right) \inf_{\|x\|=1} \sum_{i=0}^{n-1}\sum_{j=0}^{\infty} |a_j| [x^\tau \phi_{i-j}(n)]^2$$
$$\leqslant \left(\sum_{j=0}^{\infty}|a_j|\right)^2 \inf_{\|x\|=1} \sum_{i=0}^{n-1} [x^\tau \phi_i(n)]^2$$
$$\leqslant \left(\sum_{j=0}^{\infty}|a_j|\right)^2 \lambda_{\min}(n). \quad (4.30)$$

Multiplying $z^i \operatorname{Adj} A(z)$ on both sides of (1.4) and noting (1.7), we see that

$$[\det A(z)]\, y_{n-i} = z^i[\operatorname{Adj} A(z)]\, B(z)\, u_n + z^i[\operatorname{Adj} A(z)]\, w_n, \quad \forall i \geqslant 0, n \geqslant 0. \quad (4.31)$$

Now let $x_n \in R^{(m+l)p_n}$ be the unit eigenvector corresponding to the minimum eigenvalue of the matrix $\sum_{i=0}^{n-1} \psi_i(n)\psi_i^\tau(n)$, and let $x_n$ be written as (4.7), and $h_n^{(i)}$ and $g_n^{(i)}$ be defined via (4.9). Then similar to the finite order ARMAX case (e.g., [2, pp. 863–864]), it is easy to see from (2.2), (4.7)–(4.9), (4.29), and (4.31) that

$$\lambda_{\min}\left\{\sum_{i=0}^{n-1} \psi_i(n)\psi_i^\tau(n)\right\} = \sum_{i=0}^{n-1} [x_n^\tau \psi_i(n)]^2 = \sum_{i=0}^{n-1}\left\{H_n(z)[u_i^\tau, w_i^\tau]^\tau\right\}^2$$
$$= \sum_{i=0}^{n-1}\left\{\sum_{j=0}^{\infty} [h_n^{(j)\tau} u_{i-j} + g_n^{(j)\tau} w_{i-j}]\right\}^2.$$

Consequently, by Lemmas 4.2 and 4.3, and the elementary inequality $(x+y)^2 \geqslant \frac{1}{2}x^2 - y^2$, we know that

$$\begin{aligned}
&\lambda_{\min}\left\{\sum_{i=0}^{n-1} \psi_i(n)\psi_i^\tau(n)\right\}\\
&\geqslant \frac{1}{2}\sum_{i=0}^{n-1}\left\{\sum_{j=0}^{2p_n-1} [h_n^{(j)\tau} u_{i-j} + g_n^{(j)\tau} w_{i-j}]\right\}^2\\
&\quad - \sum_{i=0}^{n-1}\left\{\sum_{j=2p_n}^{\infty} [h_n^{(j)\tau} u_{i-j} + g_n^{(j)\tau} w_{i-j}]\right\}^2\\
&= \frac{1}{2}\sum_{i=0}^{n-1}\{[h(o)\tau_n, ..., h_n^{(2p_n-1)\tau}, g_n^{(o)\tau}, ..., g_n^{(2p_n-1)}]^\tau \phi_i^0(n)\}^2 + O(\delta_n^0 r_n^0)\\
&\geqslant \frac{1}{2}\lambda_{\min}^0(n)\sum_{i=0}^{2p_n-1}(\|h_n^{(i)}\|^2 + \|g_n^{(i)}\|^2) + O(\delta_n^0 r_n^0)\\
&= \frac{1}{2}p_n^{-2d}\lambda_{\min}^0(n)\left\{p_n^{2d}\sum_{i=0}^{2p_n-1}(\|h_n^{(i)}\|^2 + \|g_n^{(i)}\|^2)\right\} + O(\delta_n^0 r_n^0)\\
&\geqslant c_2 p_n^{-2d}\lambda_{\min}^0(n) + O(\delta_n^0 r_n^0), \qquad \text{a.s. as } n \longrightarrow \infty,
\end{aligned}$$

where $c_2 > 0$ is a constant.

Finally, combining the above inequality with (4.30), we see that the assertion of Theorem 2.3 holds with $c_0 = c_2/(\sum_{i=0}^{\infty}|a_i|)^2 > 0$. ∎

We now give the proof of Example 2.1 stated in Section 2.

*Proof* of Example 2.1. By Lemma 3.5 we know that

$$\max_{1\leqslant t\leqslant n}\max_{1\leqslant s\leqslant n}\left\|\sum_{i=0}^{n-1} w_{i-s}u_{i-t}^\tau\right\| = O(\{b^b \log n\}^{1/2}), \qquad \text{a.s.} \tag{4.32}$$

$$\max_{0\leqslant t\leqslant n}\max_{t< s\leqslant n}\left\|\sum_{i=0}^{n-1} w_{i-s}w_{i-t}^\tau\right\| = O(\{n \log n\}^{1/2}), \qquad \text{a.s.} \tag{4.33}$$

and by (2.17) and (2.30) it is not difficult to see that for some random constant $c_w$,

$$\liminf_{n \longrightarrow \infty} \min_{0 \leqslant j \leqslant p_n} \lambda_{\min} \left\{ \frac{1}{n} \sum_{i=0}^{n-1} w_{i-j} w_{i-j}^{\tau} \right\} \geqslant c_w > 0, \qquad \text{a.s.} \qquad (4.34)$$

Then for any vector $[\alpha^\tau, \beta^\tau]^\tau$ with $\|\alpha\|^2 + \|\beta\|^2 = 1$ and $\alpha \triangleq [\alpha_0^\tau, \alpha_1^\tau, ..., \alpha_{2p_n-1}^\tau]^\tau \in R^{2lp_n}$, $\beta \triangleq [\beta_0^\tau, \beta_1^\tau, ..., \beta_{2p_n-1}^\tau]^\tau \in R^{2mp_n}$, we have by (2.23), (2.33), 2.34), and (4.32)–(4.34),

$$\begin{aligned}
&[\alpha^{\tau}, \beta^\tau] \sum_{i=0}^{n-1} \phi_i^0(n) \phi_i^{0\tau}(n) [\alpha^\tau, \beta^\tau]^\tau \\
&= \alpha^\tau \sum_{i=0}^{n-1} \phi_i^1(n) \phi_i^{1\tau}(n) \alpha + 2\alpha^\tau \sum_{i=0}^{n-1} \phi_i^1(n) [w_i^\tau, ..., w_{i-2p_n+1}^\tau] \beta \\
&\quad + \sum_{j=0}^{2p_n-1} \beta_j^\tau \sum_{i=0}^{n-1} w_{i-j} w_{i-j}^\tau \beta_j + 2 \sum_{t=0}^{2p_n-2} \sum_{s=t+1}^{2p_n-1} \beta_t^\tau \sum_{i=0}^{n-1} w_{i-t} w_{i-s}^\tau \beta_s \\
&\geqslant \lambda_{\min}^1(n) \|\alpha\|^2 + O(p_n \{n^b \log n\}^{1/2}) + c_w n \|\beta\|^2 \\
&\quad + O\left( \{n \log\}^{1/2} \left\{ \sum_{j=0}^{2p_n-1} |\beta_j| \right\}^2 \right) \\
&\geqslant \min\{1, c_w\} \min\{n, \lambda_{\min}^1(n)\} + O(p_n \{n^b \log n\}^{1/2}).
\end{aligned}$$

This proves the assertion (2.32). In a similar way, it can be shown that under (2.35)–(2.38), there is a random constant $c_v > 0$ such that

$$\lambda_{\min} \left\{ \sum_{i=0}^{n-1} \phi_i^1(n) \phi_i^{1\tau}(n) \right\} \geqslant c_v n + o(n), \qquad \text{a.s.}$$

which in conjunction with (2.32) yields (2.39). Hence the proof is completed. ▮

## References

[1] Chen, H. F. (1982) Strong consistency and convergence rate of least squares identification. *Sci. Sinica Ser. A* **25** 771–784.

[2] Chen, H. F., and Guo, L. (1987). Optimal adaptive control and consistent parameter estimates for ARMAX model with quadratic cost. *SIAM J. Control Optim.* **25**, No. 4 845–867.

[3] Glover, K. (1984) All optimal Hankel-norm approximations of linear multivariable systems and their $L^\infty$ error bounds. *Internat. J. Control* **39** 1115–1193.

[4] Grenander, U., and Szego, G. (1958). *Toeplitz Forms and Their Applications.* University of California Press, Berkeley, CA.

[5] GUO, L., AND HUANG, D W (1988). Least squares identification for ARMAX models without the positive real condition Submitted for publication.
[6] HANNAN, E J (1987). Rational transfer function approximation. *Statist Sci.* **2**, No. 2 135–161.
[7] HANNAN, E. J., AND DEISTLER, M (1988). *The Statistical Theory of Linear Systems* Wiley, New York.
[8] HUANG, D W (1988). Convergence rate of sample autocorrelations and autocovariances for stationary time series. *Sci Sinica Ser. A* **31**, No. 4 406–424.
[9] HUANG, D W (1988). Estimating the orders and parameters for ARUMA model. Submitted for publication.
[10] LAI, T. L., AND WEI, C Z. (1982). Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *Ann. Statist.* **10**, No. 1 154–166.
[11] LAI, T. L. ROBBINS, H., AND WEI, C. Z (1979). Strong consistency of least squares estimates in multiple regression II. *J. Multivariate Anal* **9** 343–361.
[12] LJUNG, L. (1976) Consistency of least squares identification method. *IEEE Trans. Automat. Control* **AC-21** 779–781.
[13] LJUNG, L. (1977). On positive real functions and the convergence of some recursive schemes *IEEE Trans Automat. Control* **AC-22** 539–551.
[14] LJUNG, L. (1984). *Asymptotic Properties of the Least Squares Method for Estimating Transfer Functions and Disturbance Spectra Report* LiTH-ISY-I-0709. Department of Electrical Engineering, Linkoping University, Sweden.
[15] MOORE, J B (1978). On strong consistency of least squares algorithm *Automatica* **14** 505–509.
[16] MOORE, J. B (1983). Persistency of excitation in extended least squares *IEEE Trans Automat. Control* **AC-28** 60–68.
[17] RISSANEN, J (1978). Modeling by shortest data description. *Automatica* **14** 475–471.
[18] RISSANEN, J (1983). Universal priori for parameters and estimation by minimum description length. *Ann. Statist.* **11** 416–431.
[19] SOLO, V. (1979). The convergence of AML. *IEEE Trans. Automat. Control* **AC-24**, No. 6 958–962
[20] STOUT, W F (1974). *Almost Sure Convergence*. Academic Press, New York.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium