



<b>Title</b>	<b>Stabilization of discrete-time nonlinear uncertain systems by feedback based on LS algorithm</b>
<b>Author(s)</b>	<b>Li, C; Lam, J</b>
<b>Citation</b>	<b>SIAM Journal of Control and Optimization, 2013, v. 51 n. 2, p. 1128-1151</b>
<b>Issued Date</b>	<b>2013</b>
<b>URL</b>	<b><a href="http://hdl.handle.net/10722/189208">http://hdl.handle.net/10722/189208</a></b>
<b>Rights</b>	<b>Creative Commons: Attribution 3.0 Hong Kong License</b>

## STABILIZATION OF DISCRETE-TIME NONLINEAR UNCERTAIN SYSTEMS BY FEEDBACK BASED ON LS ALGORITHM\*

CHANYING LI<sup>†</sup> AND JAMES LAM<sup>‡</sup>

**Abstract.** This paper is concerned with the use of the least squares (LS) algorithm to design a feedback control law to stabilize a basic class of discrete-time nonlinear uncertain systems. The result shows that if a certain polynomial criterion is satisfied, the system can be stabilized by feedback based on the LS algorithm for Gaussian distributed noise and unknown parameters. This result thus provides an answer to the question of what are the fundamental limitations of the discrete-time adaptive nonlinear control. This issue of feedback capability has been an open problem for more than ten years since it was put forward.

**Key words.** feedback capability, LS algorithm, nonlinear systems, uncertainties, stabilization

**AMS subject classifications.** 93D15, 93D20, 93E24

**DOI.** 10.1137/110852875

**1. Introduction.** As is well known, the theory of adaptive control for both continuous- and discrete-time linear systems is thoroughly studied (see, e.g., [1], [2], [6], [7]). There is also a vast literature on nonlinear systems with nonlinearity having linear growth rate (see, e.g., [19] and [21]). However, when one attempts to control systems with output nonlinearity growing faster than linearity, an interesting phenomenon happens. For such nonlinear systems, the similarities of adaptive control between continuous- and discrete-time cases do not remain anymore. A large class of continuous-time nonlinear systems can be globally stabilized by nonlinear damping or a backstepping approach in adaptive control regardless of how fast the growth rate is (see, e.g., [8] and [10]). Unfortunately, these powerful methods in controlling continuous-time systems are no longer effective in the discrete-time case (see [9]). In fact, as pointed out by [4], fundamental difficulties arise for adaptive control of discrete-time parametric nonlinear systems. These difficulties also emerge in the control of discrete-time nonparametric nonlinear systems ([22], [12], [25]), linear stochastic systems with the unknown time-varying parameter process [23], and in the sampled-data control of continuous-time nonlinear systems with a prescribed sampling rate [24], even if the nonlinearity is bounded by a linear growth rate.

Given these difficulties mentioned above, one may be curious to know whether or not such difficulties are caused by the inherent limitations of the feedback principle. If the instability of nonlinear systems stems from the excessive uncertainties, which are beyond the maximal capability of feedback mechanism, then it is impossible to design any feedback control law to stabilize the systems, despite how hard one may try. From this view of point, a natural and important question from both the theoretical perspective and practical applications is, “Does the feedback principle have any

---

\*Received by the editors October 25, 2011; accepted for publication (in revised form) December 28, 2012; published electronically March 21, 2013.

<http://www.siam.org/journals/sicon/51-2/85287.html>

<sup>†</sup>National Center for Mathematics and Interdisciplinary Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (cyli@amss.ac.cn). This author’s research was supported by an Engineering Postdoctoral Fellowship, NNSFC 61203067, and the National Center for Mathematics and Interdisciplinary Sciences, CAS.

<sup>‡</sup>Department of Mechanical Engineering, The University of Hong Kong, Hong Kong (james.lam@hku.hk). This author’s research was supported by GRF 7137/09E.

limitations in dealing with uncertainties?" To investigate this fundamental problem, we have to place ourselves into a framework that is somewhat beyond those of the traditional robust control and adaptive control, because one needs to study the fundamental limitations of the full feedback mechanism which includes all (nonlinear and time-varying) feedback laws, not just restricted to a specific class of feedback laws.

The first effort to answer this question was made by [4], where the following basic model is considered:

$$(1) \quad y_{t+1} = \theta_1 y_t^b + u_t + w_{t+1},$$

where  $\theta_1$  is an unknown parameter,  $\{w_t\}$  is a Gaussian white noise sequence, and exponent  $b > 0$  is the growth rate. It was proved that the system is almost surely globally stabilizable if and only if  $b < 4$  (see [4]). This result can also be derived if the Gaussian noise is replaced by bounded noises (see [14]).

Later on, a negative result of [4] was extended by [20] to systems with multiple unknown parameters :

$$(2) \quad y_{t+1} = \theta_1 y_t^{b_1} + \theta_2 y_t^{b_2} + \cdots + \theta_p y_t^{b_p} + u_t + w_{t+1},$$

where the noises are Gaussian distributed. It proved that system (2) is not almost surely stabilizable by feedback if there is an  $x \in [1, b_1]$  such that  $P(x) < 0$ , where

$$(3) \quad P(x) = x^{p+1} - b_1 x^p + (b_1 - b_2)x^{p-1} + \cdots + b_p.$$

Although this paper provided a polynomial necessity rule to describe the "impossibility theorem," it has not given a complete characterization on feedback capability, since no evidence showed that the polynomial rule is also sufficient for the stabilizability of system (2) in this paper. That is, it has not found any feedback control law to stabilize system (2). This question in fact has not been answered for the more than ten years since [20] was published. Afterwards, the limitations of the feedback mechanism was further studied by [11] for the high order case of multiple parameterized stochastic systems based on [20]. What is the barrier that keeps people from presenting a complete characterization of the feedback capability?

It is the essential difficulties encountered in the analysis of adaptive law that introduces multiplicative nonlinearity and complexity for the multiple parametric systems. Specifically, the feedback controller based on a recursive least squares (LS) algorithm, which stabilized the uncertain systems in [4], turns out to be quite involved theoretically when systems own more than one unknown parameter. As a matter of fact, it is rather challenging to estimate the maximum and minimum eigenvalues of information matrices generated by the LS algorithm for the case of multiple parameters. Meanwhile for work [4] with the scalar parameter, the information matrices degenerated to a series of real numbers and there was no obstacle to dealing with these eigenvalues. The existing techniques could even cause exponentially large error in the eigenvalue estimation for system (2) due to the high nonlinearity and complexity. Without developing any new ideas, many attempts during the last decade trying to prove the sufficient part of this problem inevitably failed.

Not long ago, Li, Xie, and Guo [18] proved that the polynomial rule (3) does serve as a necessary and sufficient condition for global feedback stabilization of system (2), but with bounded multiple unknown parameters and bounded noises. This result was derived by using a somewhat complicated purely deterministic method. Shortly thereafter, by introducing a simple stochastic embedding approach, a new critical theorem on the feedback capability was established for the uncertain systems where an

additional uncertain parameter is inserted on the input channel [15]. This stochastic embedding approach was further developed in [16], and was used to study the capability and limitations of feedback mechanisms in globally stabilizing a basic class of discrete-time nonlinearly parameterized dynamical systems with bounded unknown parameters and bounded noises [17].

As mentioned above, although the research on feedback capability for nonlinear uncertain systems first began under a stochastic framework, it has almost stood still since earlier attempts [20] and [21] due to the analysis difficulties. On the other hand, much progress has been made in the deterministic framework, and the methods to investigate the fundamental limitations of feedback are well developed. However, these methods are not applicable to the stochastic case. Thus, to achieve a parallel theory in the stochastic framework, some new ideas and approaches have to be developed.

In this paper, we will see that the polynomial rule (3) does indeed provide a necessary and sufficient condition of stabilizability for system (2) in the stochastic framework. We will prove that system (2) can be stabilized by a feedback control law based on the recursive LS algorithm. While various excitation conditions (including the persistent excitation condition) are widely used in the majority of existing works when the LS algorithm is performed, we successfully get rid of them by introducing a constructive method which is effective in overcoming the essential difficulties referred to in the foregoing paragraphs, that is, to estimate the maximal and minimum eigenvalues of the information matrices. In particular, this result completes the characterization of feedback limitations for a stochastic system (2) with multiple parameters, which has remained open in this field since work [20] was published.

The rest of the paper is organized as follows. In the next section, we will present the main theorem of the paper, with the proof given in section 3. In section 4, we provide some numerical experiments to illustrate our main result. Finally, the concluding remarks will be given in section 5.

**2. Main results.** Consider the following system

$$(4) \quad y_{t+1} = \theta_1 y_t^{b_1} + \theta_2 y_t^{b_2} + \cdots + \theta_n y_t^{b_n} + u_t + w_{t+1},$$

where  $y_t$ ,  $u_t$ , and  $w_t$  are the system output, input, and noise sequences, respectively,  $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n$  is a random or deterministic unknown parameter, and  $b_1 > 1$ ,  $b_1 > b_2 > \cdots > b_n > 0$  are  $n$  real numbers.

To facilitate the analysis of the above closed-loop control system, we need the following conditions.

- (A1)  $\{w_t\}$  is an independently and identically distributed (i.i.d.) sequence with standard normal distribution  $N(0, 1)$ .
- (A2) The unknown parameter vector  $\theta$  is independent of  $\{w_t\}$  and has a Gaussian distribution  $N(\bar{\theta}, I_n)$ .

To explore the feedback capability of the uncertain system in (4), a standard definition of globally stabilizable (see [20]) is presented below.

**DEFINITION 2.1.** Let  $\sigma\{y_i, 0 \leq i \leq t\}$  be the  $\sigma$  field generated by the observations  $\{y_i, 0 \leq i \leq t\}$ . System (4) is said to be almost surely globally stabilizable, if there exists a feedback control  $u_t \in \mathcal{F}_t^y \triangleq \sigma\{y_i, 0 \leq i \leq t\}$ ,  $t = 0, 1, \dots$  such that for any initial condition  $y_0 \in \mathbb{R}^1$ ,

$$(5) \quad \limsup_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^t y_i^2 < \infty \quad a.s.$$

The main theorem of this paper is presented as follows.

**THEOREM 2.1.** *Under assumptions (A1) and (A2), system (4) with  $b_1 > 1$ ,  $b_1 > b_2 > \dots > b_n > 0$  is globally stabilizable if and only if for any  $x \in (1, b_1)$ ,  $P(x) > 0$ , where  $P(x)$  is defined by*

$$(6) \quad P(x) = x^{n+1} - b_1 x^n + (b_1 - b_2)x^{n-1} + \dots + b_n.$$

**3. The proof of Theorem 2.1.** As introduced in the preceding sections, the necessity of Theorem 2.1 was almost proved by [20]. Thus, this part mainly focuses on the proof of sufficiency in section 3.1. Since the arguments are relatively involved, we outline the overall idea.

To achieve the stabilization of system (4) by feedback based on the LS algorithm under polynomial criterion (6), we first define the recursive LS algorithm below. Suppose the probability space is  $(\Omega, \mathcal{F}, P)$ . The standard LS estimate  $\theta_t$  for  $\theta$  is recursively defined by

$$(7) \quad \theta_{t+1} \triangleq \theta_t + P_{t+1} \phi_t (y_{t+1} - u_t - \phi_t^\tau \theta_t),$$

$$(8) \quad P_{t+1} \triangleq P_t - \frac{P_t \phi_t \phi_t^\tau P_t}{1 + \phi_t^\tau P_t \phi_t}, \quad P_0 = I,$$

$$(9) \quad \phi_t \triangleq (y_t^{b_1}, y_t^{b_2}, \dots, y_t^{b_n})^\tau, \quad t \geq 0,$$

where  $\theta_0, P_0 = I$  are the deterministic initial conditions of the algorithm, and  $\phi_0$  is possibly a random initial vector of the system.

The certainty equivalence adaptive tracking control is defined by

$$(10) \quad u_t = -\theta_t^\tau \phi_t, \quad t \geq 0;$$

substituting this into (4), we have the following closed-loop equation:

$$(11) \quad y_{t+1} = \tilde{\theta}_t^\tau \phi_t + w_{t+1}, \quad t \geq 0,$$

where  $\tilde{\theta}_t \triangleq \theta - \theta_t$ . In fact, we only need to prove that the outputs of (11) satisfy (5) for sufficiently large  $|y_0|$ .

Insight into the problem is as follows: with the help of [3], we could write the average sum of squares of the inputs in (11) by

$$(12) \quad \frac{1}{t} \sum_{i=0}^t y_{i+1}^2 = O \left( \frac{1}{t} \log |P_{t+1}| \frac{|P_{t+1}|}{|P_t|} + \frac{1}{t} \sum_{i=0}^t w_{i+1}^2 \right).$$

Note that assumption (A1) implies the almost sure boundedness of the noise term in the right-hand side (RHS) of (12); therefore, it is the first term in the RHS of (12) that dominates the stabilization. Simple calculation shows that

$$\log |P_{t+1}| = o \left( \sum_{i=0}^t y_{i+1}^2 \right) + O(1);$$

thus, the most important and also difficult task in our proof is to verify that

$$(13) \quad \sup_{t \geq 0} \frac{|P_{t+1}|}{|P_t|} < \infty \quad \text{a.s.}$$

in the presence of the polynomial criterion. This aim is achieved by two steps. It may be somewhat surprising from the proof that the only role polynomial (6) plays

in stabilizing system (4) is to prevent part of the information matrices from growing excessively large:

$$(14) \quad |P_{t+1}^{-1}| < 2^{\log^2 t}, \quad \text{i.o.},$$

which is the main target of the first step. The abbreviation “i.o.” represents “infinitely often.” With property (14), the validity of (13) can be proved in the next step with the essential difficulty of the LS algorithm solved.

Now, we are ready to prove Theorem 2.1 in detail.

**3.1. Sufficiency part of Theorem 2.1.** In the first part of the proof, we proceed to show that (14) holds almost surely. For this, define a random matrix sequence  $\{Q_k^{-1}\}$  by

$$(15) \quad \begin{cases} Q_0^{-1} = I, \\ Q_1^{-1} = Q_0^{-1} + \phi_0 \phi_0^\tau, \\ Q_k^{-1} = Q_{k-1}^{-1} + \phi_{t_{k-1}} \phi_{t_{k-1}}^\tau, \quad k = 1, 2, \dots, \end{cases}$$

where the monotone random subscript  $t_k$  with  $t_0 = 0$  satisfies, for  $k \geq 1$ ,

$$(16) \quad \begin{cases} \phi_{t_k}^\tau Q_k \phi_{t_k} > \phi_{t_{k-1}}^\tau Q_{k-1} \phi_{t_{k-1}}, \\ \phi_t^\tau Q_k \phi_t \leq \phi_{t_{k-1}}^\tau Q_{k-1} \phi_{t_{k-1}} \quad \text{for any } t_{k-1} < t < t_k. \end{cases}$$

From (16) and the fact ([3])

$$(17) \quad \frac{|Q_{k+1}^{-1}|}{|Q_k^{-1}|} = 1 + \phi_{t_k}^\tau Q_k \phi_{t_k},$$

it is easy to see by (15) and (16) that

$$(18) \quad \frac{|Q_{k+1}^{-1}|}{|Q_k^{-1}|} \geq \frac{|Q_k^{-1}|}{|Q_{k-1}^{-1}|} \geq 1, \quad k = 1, 2, \dots$$

and

$$(19) \quad |Q_k^{-1}| \geq |Q_{k-1}^{-1}| \geq 1, \quad k = 1, 2, \dots$$

If  $\{t_k\}$  is a finite subsequence on some set  $G \subset \Omega$  with positive probability, then there is a random  $k_0 \geq 0$  such that for all  $t \geq t_{k_0}$  on  $G$ ,

$$\phi_t^\tau Q_{k_0+1} \phi_t \leq \phi_{t_{k_0}}^\tau Q_{k_0} \phi_{t_{k_0}},$$

which immediately leads to the boundness of  $\|\phi_t\|$  on this set. Then, system (4)–(9) is globally stable on set  $G$  and the remaining task is to prove the stabilization on set  $G^c$ . Without loss of generality, we assume  $\{t_k\}$  is an infinite random sequence everywhere. This is because the whole space mentioned in the following can be viewed as the probability space restricted on  $G^c$ .

**LEMMA 3.1.** For any  $t_{k-1} < t \leq t_k$ , where  $t_k$  is defined by (16),  $\frac{|P_t^{-1}|}{|P_{t-1}^{-1}|} \leq \frac{|Q_k^{-1}|}{|Q_{k-1}^{-1}|}$ .

*Proof.* If  $t-1 > t_{k-1}$ , then  $t-2 \geq t_{k-1}$ . From  $P_{t-1}^{-1} = \sum_{i=0}^{t-2} \phi_i \phi_i^\tau$ , from (15) we have  $P_{t-1}^{-1} \geq Q_k^{-1}$ , and hence  $P_{t-1} \leq Q_k$  by the Milliken–Akdeniz theorem [13]. Consequently,

$$(20) \quad \frac{|P_t^{-1}|}{|P_{t-1}^{-1}|} = 1 + \phi_{t-1}^\tau P_{t-1} \phi_{t-1} \leq 1 + \phi_{t-1}^\tau Q_k \phi_{t-1}.$$

Note that  $t_{k-1} < t - 1 < t_k$ ; by (16) we have

$$\phi_{t-1}^\tau Q_k \phi_{t-1} \leq \phi_{t_{k-1}}^\tau Q_{k-1} \phi_{t_{k-1}},$$

which together with (17) yields the lemma for the case  $t - 1 > t_{k-1}$ .

For  $t - 1 = t_{k-1}$ , we have  $P_{t-1} = P_{t_{k-1}} \leq Q_{k-1}$ . Since  $\phi_{t-1} = \phi_{t_{k-1}}$ , we have

$$\frac{|P_t^{-1}|}{|P_{t-1}^{-1}|} \leq 1 + \phi_{t_{k-1}}^\tau Q_{k-1} \phi_{t_{k-1}} = \frac{|Q_k^{-1}|}{|Q_{k-1}^{-1}|}.$$

The proof is thus completed.  $\square$

LEMMA 3.2. *Let  $P(x)$  be the polynomial defined by (6) and  $P(x) > 0$  whenever  $x \in (1, b_1)$ . Then,*

$$P \left\{ |Q_{k+1}^{-1}| < \left( 2^{\log^2 t_k} t_k^{-2(n+\sum_{i=1}^n b_i)} \right)^{\frac{1}{2\sum_{i=1}^n b_i}} \text{ i.o.} \right\} = 1.$$

*Proof.* Define  $t_{-1} \triangleq -1$ ,  $y_{-1} \triangleq 1$ , and for any given  $k \geq n - 1$  define

$$\mathcal{H}_k \triangleq \{(h_1, h_2, \dots, h_n) : h_i \in \{t_{-1}, t_0, \dots, t_k\}, \\ 1 \leq i \leq n; h_i \neq h_j \text{ if } i \neq j \text{ except } h_j = t_{-1}\}$$

and  $\mathcal{S} \triangleq \pi(1, 2, \dots, n)$ , where  $\pi(\cdot)$  denotes the class of all permutations of some countable sequence. By (15), we have

$$(21) \quad Q_{k+1}^{-1} = \sum_{i=0}^k \phi_{t_i} \phi_{t_i}^\tau + I;$$

then, for  $k \geq n - 1$ ,  $|Q_{k+1}^{-1}|$  is the summation of the following general form (see [20])

$$y_{h_1}^{b_1+b_{s_1}} y_{h_2}^{b_2+b_{s_2}} \dots y_{h_n}^{b_n+b_{s_n}},$$

where  $(h_1, h_2, \dots, h_n) \in \mathcal{H}_k$  and  $(s_1, s_2, \dots, s_n) \in \mathcal{S}$ . It is easy to see that there are  $(k + 2)^n n!$  such terms; hence for any  $k \geq n - 1$ ,

$$(22) \quad |Q_{k+1}^{-1}| \leq [(k + 2)^n n!] \prod_{i=1}^n y_{m_i}^{2b_i},$$

where  $m_i, i = 1, \dots, n$  are the subscripts of the largest  $n$  numbers of  $\{|y_{t_j}|, j = -1, \dots, k\}$  with the order

$$|y_{m_1}| \geq |y_{m_2}| \geq \dots \geq |y_{m_n}|, \quad m_i \neq m_j.$$

Now, let  $\alpha_t \triangleq (1 + \phi_t^\tau P_t \phi_t)^{-1} (\tilde{\theta}_t^\tau \phi_t)^2$ ; we have by (11) that

$$(23) \quad y_t^2 \leq 2\alpha_{t-1} \frac{|P_t^{-1}|}{|P_{t-1}^{-1}|} + 2w_t^2, \quad t \geq 1.$$

If we define  $\alpha_{-1} = \alpha_{-2} \triangleq \frac{1}{2} \max\{1, y_0^2\}$ ,  $w_{-1} = w_0 \triangleq 0$ , and  $P_{-1}^{-1} = P_{-2}^{-1} \triangleq I$ , then (23) holds for all  $t \geq -1$ . Therefore,

$$(24) \quad y_{m_i}^2 \leq 2\alpha_{m_i-1} \frac{|P_{m_i}^{-1}|}{|P_{m_i-1}^{-1}|} + 2w_{m_i}^2, \quad i = 1, 2, \dots, n.$$

Since for any  $i \in [1, n]$ ,  $m_i = t_{k_i}$  for some  $k_i \in [-1, k]$ , by Lemma 3.1, we have

$$(25) \quad \frac{|P_{m_i}^{-1}|}{|P_{m_i-1}^{-1}|} \leq \frac{|Q_{k_i}^{-1}|}{|Q_{k_i-1}^{-1}|},$$

where  $Q_{-1}^{-1} = Q_{-2}^{-1} \triangleq I$ . Thus, by (18), for  $1 \leq t \leq m_i$ ,

$$(26) \quad \frac{|P_t^{-1}|}{|P_{t-1}^{-1}|} \leq \frac{|Q_{k_i}^{-1}|}{|Q_{k_i-1}^{-1}|}.$$

Consequently, we have by (19) that

$$(27) \quad \begin{aligned} |P_{m_i}^{-1}| &= \frac{|P_{m_i}^{-1}|}{|P_{m_i-1}^{-1}|} \frac{|P_{m_i-1}^{-1}|}{|P_{m_i-2}^{-1}|} \cdots \frac{|P_{-1}^{-1}|}{|P_{-2}^{-1}|} |P_{-2}^{-1}| \\ &\leq \left( \frac{|Q_{k_i}^{-1}|}{|Q_{k_i-1}^{-1}|} \right)^{|m_i|} \frac{|P_0^{-1}|}{|P_{-1}^{-1}|} \frac{|P_{-1}^{-1}|}{|P_{-2}^{-1}|} |P_{-2}^{-1}| \\ &\leq |Q_{k_i}^{-1}|^{|m_i|}. \end{aligned}$$

From [3, Corollary 2.1(i)], (27), and the definitions of  $\alpha_{-j}$ ,  $j = -1, -2$ , we have

$$(28) \quad \alpha_{m_i-1} = O(\log |P_{m_i}^{-1}|) + O(1) = O(|m_i| \log |Q_{k_i}^{-1}|) + O(1) \quad \text{a.s.}$$

Observe that by (A1),  $w_t^2 = O(t)$  a.s. for  $t \geq 1$ ; then given  $1 \leq i \leq n$ , by (19), (24), (25), and (28) we have for sufficiently large  $k$  that

$$(29) \quad \begin{aligned} y_{m_i}^2 &= O\left( (\alpha_{m_i-1} + |m_i|) \frac{|P_{m_i}^{-1}|}{|P_{m_i-1}^{-1}|} \right) \\ &= O\left( (|m_i| \log |Q_{k_i}^{-1}| + |m_i| + 1) \frac{|Q_{k_i}^{-1}|}{|Q_{k_i-1}^{-1}|} \right) \\ &= O\left( (t_k \log |Q_k^{-1}|) \frac{|Q_{k_i}^{-1}|}{|Q_{k_i-1}^{-1}|} \right) \quad \text{a.s.}, \end{aligned}$$

where the last inequality follows from the assumption  $|Q_k^{-1}| \rightarrow \infty$  as  $k \rightarrow \infty$ .

Note that by definition  $m_i \neq m_j$  for  $i \neq j$ , then  $k_i \neq k_j$ . By (18), (22), and (29), for any  $k$  sufficiently large, we have

$$(30) \quad |Q_{k+1}^{-1}| = O\left( t_k^{\sum_{j=1}^n b_j + n} (\log |Q_k^{-1}|)^{\sum_{j=1}^n b_j} \prod_{i=1}^n \left( \frac{|Q_{k-i+1}^{-1}|}{|Q_{k-i}^{-1}|} \right)^{b_i} \right) \quad \text{a.s.}$$

Therefore, note that  $|Q_k^{-1}| \rightarrow \infty$  as  $k \rightarrow \infty$ , for any  $\epsilon > 0$ , by taking the logarithm on both sides of (30), it yields

$$(31) \quad \begin{aligned} (1 - \epsilon) \log |Q_{k+1}^{-1}| &\leq \sum_{i=1}^n b_i (\log |Q_{k-i+1}^{-1}| - \log |Q_{k-i}^{-1}|) \\ &\quad + \left( \sum_{j=1}^n b_j + n \right) \log t_k + O(1). \end{aligned}$$



Suppose there is a positive random variable  $k'$  and a set  $D \subset \Omega$  with  $P\{D\} > 0$  such that for all  $k \geq k'$  large enough,

$$(32) \quad |Q_{k+1}^{-1}| \geq \left(2^{\log^2 t_k} t_k^{-2(n+\sum_{i=1}^n b_i)}\right)^{\frac{1}{2\sum_{i=1}^n b_i}} \quad \text{on } D.$$

Then,  $\log t_k = o(\log |Q_{k+1}^{-1}|)$  and hence by (31), on the set  $D$  for any  $k \geq k'$ ,

$$(33) \quad (1 - 2\epsilon) \log |Q_{k+1}^{-1}| \leq \sum_{i=1}^n b_i (\log |Q_{k-i+1}^{-1}| - \log |Q_{k-i}^{-1}|) \quad \text{a.s.}$$

Now, let  $a_k = \log |Q_k^{-1}|$  and rewrite (33) by

$$(34) \quad (1 - 2\epsilon)a_{k+1} \leq \sum_{i=1}^n b_i (a_{k-i+1} - a_{k-i}) \quad \text{a.s.,}$$

where  $a_k \nearrow \infty$  a.s. as we assumed earlier. Define  $z_k \triangleq \frac{a_k}{a_{k-1}}$  and denote random  $z \triangleq \lim_{k \rightarrow \infty} z_k$ . Obviously,  $z \geq 1$  according to (18). Rewrite (34) by dividing by  $a_{k+1}$ ; we then have

$$(35) \quad (1 - 2\epsilon) + \sum_{i=1}^{n-1} (b_i - b_{i+1}) \frac{1}{\prod_{j=0}^i z_{k+1-j}} + b_n \frac{1}{\prod_{j=0}^n z_{k+1-j}} \leq b_1 \frac{1}{z_{k+1}} \quad \text{a.s.}$$

Taking limit inferior on both sides of the above inequality, we obtain that on set  $D$ , for any  $\epsilon > 0$ ,

$$(1 - 2\epsilon) + \sum_{i=1}^{n-1} (b_i - b_{i+1}) \frac{1}{z^{i+1}} + b_n \frac{1}{z^{n+1}} \leq b_1 \frac{1}{z} \quad \text{a.s.}$$

Letting  $\epsilon \rightarrow 0$ , we immediately deduce that the limit  $z \neq \infty$  and  $P(z) \leq 0$  a.s. on  $D$ . Observe that  $P(x) > 0$  for any  $x \in (1, b_1)$  implies  $P(x) > 0$  for any  $x \geq 1$ , which leads to a contradiction since  $z \geq 1$ . Thus, (32) cannot be true and hence the lemma is proved.  $\square$

LEMMA 3.3. *Under the conditions of Lemma 3.2, (14) holds with probability 1.*

*Proof.* Similar to (22), we have for  $t \geq n$  that

$$(36) \quad |P_t^{-1}| \leq [(t+1)^n n!] \prod_{i=1}^n y_{m'_i}^{2b_i},$$

where  $m'_i, i = 1, \dots, n$  are the subscripts of the largest  $n$  numbers of  $\{|y_j|, j = -1, \dots, t-1\}$ . As the arguments for (24), we can obtain that

$$y_{m'_i}^2 \leq 2\alpha_{m'_i-1} \frac{|P_{m'_i}^{-1}|}{|P_{m'_i-1}^{-1}|} + 2w_{m'_i}^2, \quad i = 1, 2, \dots, n.$$

Hence, by  $\alpha_i = O(\log |P_i^{-1}|) + O(1)$  a.s. and (36), for sufficiently large  $t$ ,

$$(37) \quad \begin{aligned} |P_t^{-1}| &= O\left(t^n \prod_{i=1}^n \left[ (\alpha_{m'_i-1} + |m'_i|) \frac{|P_{m'_i}^{-1}|}{|P_{m'_i-1}^{-1}|} \right]^{b_i}\right) \\ &= O\left(t^{n+\sum_{i=1}^n b_i} (\log |P_t^{-1}|)^{\sum_{i=1}^n b_i} \prod_{i=1}^n \left(\frac{|P_{m'_i}^{-1}|}{|P_{m'_i-1}^{-1}|}\right)^{b_i}\right) \quad \text{a.s.,} \end{aligned}$$

where the last inequality follows from

$$\alpha_{m'_i-1} + |m'_i| = O(\log |P_t^{-1}| + 1 + (t-1)) = O(t \log |P_t^{-1}|) \quad \text{a.s.},$$

since  $t = O(\sum_{i=0}^{t-1} y_i^2) = O(|P_t^{-1}|) \rightarrow \infty$  by [4, Remark 2.7].

Note that for any  $t \geq 1$ , there is a  $t_k$  such that  $t_k < t-1 \leq t_{k+1}$ . Since  $m'_i \leq t-1$ , by (26), we have

$$\frac{|P_{m'_i}^{-1}|}{|P_{m'_i-1}^{-1}|} \leq \frac{|Q_{k+1}^{-1}|}{|Q_k^{-1}|} \leq |Q_{k+1}^{-1}|, \quad 1 \leq i \leq n.$$

Then, from (37) we obtain that

$$|P_t^{-1}| = O\left(t^{n+\sum_{i=1}^n b_i} (\log |P_t^{-1}|)^{\sum_{i=1}^n b_i} |Q_{k+1}^{-1}|^{\sum_i b_i}\right) \quad \text{a.s.},$$

and hence for large enough  $t$ ,

$$|P_t^{-1}|^{\frac{1}{2}} \leq t^{n+\sum_{i=1}^n b_i} |Q_{k+1}^{-1}|^{\sum_i b_i} \quad \text{a.s.},$$

which immediately gives by Lemma 3.2 and  $t_k < t$  that

$$|P_t^{-1}| \leq t^{2(n+\sum_{i=1}^n b_i)} |Q_{k+1}^{-1}|^{2(\sum_i b_i)} < 2^{\log^2 t} \quad \text{i.o.},$$

i.e., (14) holds with probability 1.  $\square$

In the second part, the arguments are devoted to proving  $\sup_{t \geq 0} \frac{|P_{t+1}^{-1}|}{|P_t^{-1}|} < \infty$  a.s.

The key idea is to estimate the minimal eigenvalue of  $P_t^{-1}$ , whose growth rate turns out to be faster than  $t$ . This is the most difficult part in our proof and solved by introducing a stochastic complex function. Since the analysis is quite involved, we put the estimation of  $\lambda_{\min}(t)$ , that is, the minimal eigenvalue of  $P_t^{-1}$  in Appendix A. By the fact that  $t = O(\lambda_{\min}(t))$  as  $t \rightarrow \infty$  from Appendix A, we have the following conclusion.

LEMMA 3.4. *There is a random time  $t_1$  such that for any  $t \geq t_1$ ,  $\frac{|P_{t+1}^{-1}|}{|P_t^{-1}|} < 2$  a.s.*

*Proof.* To prove the result, we first estimate the outputs  $y_t$  in terms of  $\frac{|P_t^{-1}|}{|P_{t-1}^{-1}|}$ ,  $t \geq 1$ . From the inequality

$$(38) \quad 1 - \Phi(y) \leq \frac{1}{y} \phi(y) \quad \forall y > 0,$$

where  $\Phi(y)$  and  $\phi(y)$  are the standard normal distribution function and density, respectively, defined before, we have by (56) and (38) that

$$(39) \quad \begin{aligned} P\left\{\frac{y_t}{\sigma_{t-1}} \geq \log t\right\} &\leq \frac{1}{\sqrt{2\pi}} e^{-\frac{\log^2 t}{2}} \frac{1}{\log t} \\ &= O\left(\frac{1}{t^2}\right). \end{aligned}$$

Consequently,

$$\begin{aligned} P\left\{\frac{y_t^2}{\sigma_{t-1}^2} \geq \log^2 t\right\} &= P\left\{\frac{y_t}{\sigma_{t-1}} \geq \log t\right\} + P\left\{\frac{y_t}{\sigma_{t-1}} \leq -\log t\right\} \\ &= 2P\left\{\frac{y_t}{\sigma_{t-1}} \geq \log t\right\} = O\left(\frac{1}{t^2}\right), \end{aligned}$$

which immediately yields

$$\sum_{t=1}^{\infty} P \left\{ \frac{y_t^2}{\sigma_{t-1}^2} \geq \log^2 t \right\} < \infty.$$

Thus, by the Borel–Cantelli lemma,

$$(40) \quad y_t^2 = O(\sigma_{t-1}^2 \log^2 t) = O \left( \log^2 t \frac{|P_t^{-1}|}{|P_{t-1}^{-1}|} \right) \quad \text{a.s.}$$

Now, we assert

$$(41) \quad P \left\{ \frac{|P_{t+1}^{-1}|}{|P_t^{-1}|} < 2 \quad \text{i.o.} \right\} = 1.$$

Otherwise,  $2^t = O(|P_t^{-1}|)$  with positive probability, which contradicts Lemma 3.3. Hence there is a sufficiently large random  $t_1 > t_0$  such that

$$(42) \quad \frac{|P_{t_1+1}^{-1}|}{|P_{t_1}^{-1}|} < 2 \quad \text{a.s.}$$

Now, suppose  $\frac{|P_{k+1}^{-1}|}{|P_k^{-1}|} < 2$  a.s. for some random  $k \geq t_1 > t_0$ . Then, from (40), for sufficiently large random  $t_1$ ,

$$(43) \quad \begin{aligned} y_{k+1}^2 &= O \left( \log^2(k+1) \frac{|P_{k+1}^{-1}|}{|P_k^{-1}|} \right) \\ &= O(\log^2(k+1)) \quad \text{a.s.} \end{aligned}$$

Note that  $\lambda_{\min}(k+1) \geq Ck$ ; by Lemma A.8 and

$$\frac{|P_{k+2}^{-1}|}{|P_{k+1}^{-1}|} = 1 + \phi_{k+1}^T P_{k+1} \phi_{k+1},$$

we have that for sufficiently large random  $t_1$ ,

$$\begin{aligned} \frac{|P_{k+2}^{-1}|}{|P_{k+1}^{-1}|} &\leq 1 + \frac{\|\phi_{k+1}\|^2}{\lambda_{\min}(k+1)} \leq 1 + \frac{\sum_{i=1}^n y_{k+1}^{2b_i}}{Ck} \\ &= 1 + O \left( \frac{\log^{2b_1}(k+1)}{k} \right) \\ &< 2 \quad \text{a.s.,} \end{aligned}$$

which implies the lemma is true by induction.  $\square$

*Proof of the sufficiency of Theorem 2.1.* Since by Lemma 3.1,  $\lim_{k \rightarrow \infty} |Q_k^{-1}| < \infty$  trivially gives

$$\sup_t \frac{|P_{t+1}^{-1}|}{|P_t^{-1}|} < \infty,$$

from Lemma 3.4, we know that the above inequality always holds almost surely. Therefore, by (11), we have

$$\sum_{i=0}^t (y_{i+1} - w_{i+1})^2 = \sum_{i=0}^t \alpha_i \frac{|P_{i+1}^{-1}|}{|P_i^{-1}|} = O(\log r_t) \quad \text{a.s.}$$

By (61) of [4], we obtain  $\sum_{i=0}^t y_i^2 = O(t)$ .  $\square$

**3.2. Necessity part of Theorem 2.1.** The sufficiency of Theorem 2.1 is proved in the above section. So, the remaining issue is to verify the necessity. In fact, the necessity of the theorem is almost proved by [20]. This paper shows that system (4) is not globally stabilizable whenever  $P(x) < 0$  for some  $x \in (1, b_1)$ . Now, we will prove that system (4) is also not globally stabilizable if there is some  $x \in (1, b_1)$  such that  $P(x) = 0$ . This assertion can be checked directly by the following simple lemma.

LEMMA 3.5. *Let  $\{a_t\}$  be a nonnegative sequence satisfying*

$$a_{t+1} \geq \frac{(a_0^{-x-n} P(x))^{x^t}}{\sqrt{3}(t+1)^{5/4}} a_t^x,$$

where  $x \in (1, b_1)$  is a constant such that  $P(x) = 0$ . Then,  $a_t$  diverges for sufficiently large  $a_0$ .

*Proof.* Since  $P(x) = 0$ , we immediately obtain that

$$(44) \quad a_{t+1} \geq \frac{a_t^x}{\sqrt{3}(t+1)^{5/4}}.$$

Let

$$(45) \quad c_t \triangleq \left(3(t+1)^{\frac{5}{4}}\right)^{\frac{1}{(x-1)}}.$$

Note that for sufficiently large  $t$ ,

$$(46) \quad \frac{c_{t+1}}{c_t} = \left(\frac{t+2}{t+1}\right)^{\frac{5}{4(x-1)}} < \sqrt{3};$$

then, by (44)–(46),

$$\begin{aligned} \frac{a_{t+1}}{c_{t+1}} &\geq \frac{1}{c_{t+1}} \left( \frac{a_t^x}{\sqrt{3}(t+1)^{5/4}} \right) \\ &= \frac{\sqrt{3}c_t}{c_{t+1}} \left( \frac{a_t}{c_t} \right)^x \\ &\geq \left( \frac{a_t}{c_t} \right)^x. \end{aligned}$$

Take  $a_0$  large enough such that  $\frac{a_0}{c_0} > 1$ . Since  $x > 1$ , we immediately have  $\frac{a_t}{c_t} \rightarrow \infty$  as  $t \rightarrow \infty$ . Note that by (45)  $c_t > 1$ , the divergence of  $a_t$  then follows at once.  $\square$

*Proof of Theorem 2.1.* We first prove that system (4) is not globally stabilizable if there is some  $x \in (1, b_1)$  with  $P(x) = 0$ . Now, let  $x$  be a constant in  $(1, b_1)$  such

that  $P(x) = 0$ . By [20, p. 1780], we know that there is a set  $D_0 \subset \Omega$  with  $P(D_0) > 0$  such that

$$|y_{t+1}| \geq \frac{(|y_0|^{-x-n} P(x))^{x^t}}{\sqrt{3}(t+1)^{5/4}} |y_t|^x,$$

when  $|y_0|$  is large enough. By Lemma 3.5, we have  $|y_t|$  tends to infinity. According to [20, Theorem 1], the necessity is then proved. Since the sufficiency is already verified in section 2, we complete the proof of Theorem 2.1.  $\square$

**4. Simulation.** In this section, we demonstrate the stabilization of system (4) with feedback controller based on the recursive LS algorithm (7)–(9) via simulation. First, we characterize the necessary and sufficient condition of Theorem 2.1 graphically by a stabilizability region of vector  $(b_1, b_2, \dots, b_n)$  in the Euclidean space. As we know, when dimension  $n = 1$ , the polynomial criterion degenerates to  $b_1 < 4$ , which gives a very clear description on the necessary and sufficient condition of stabilizability for system (4). In fact, for the cases  $n = 2$  and  $n = 3$ , the polynomial criterion can also be visualized with the help of the stabilizability region of vector  $(b_1, b_2, \dots, b_n)$ ,  $n = 2$  or 3, in the corresponding Euclidean space; see Figures 4.1 and 4.2. System (4) with growth rate description vector  $(b_1, b_2, \dots, b_n)$  falling in the stabilizability region, that is, the part within the shaded area in Figure 4.1 (for  $n = 2$ ) or 4.2 (for  $n = 3$ ), can be determined directly to be stabilizable by Theorem 2.1.

Now, for a point  $(b_1, b_2, \dots, b_n)$  falling in the stabilizability region, we present two examples for dimensions  $n = 2$  and  $n = 3$ , respectively, to illustrate our result.

*Example 4.1.* Let  $n = 2$ . Consider system (4) with  $b_1 = 3$  and  $b_2 = 1$ . Let noise sequence  $\{w_t\}$  and unknown parameter vector  $\theta$  satisfy assumptions (A1) and (A2), where the mean value of  $\theta$  is given by  $(2, 3)$ . It is easy to check that the two local extrema of curve

$$P(x) = x^3 - 3x^2 + 2x + 1$$

occur in  $(0, 1)$  and  $(1, \infty)$ , respectively. This means for any  $x \in (0, 1)$ ,

$$P(x) \geq \min\{P(0), P(1)\} > 0.$$

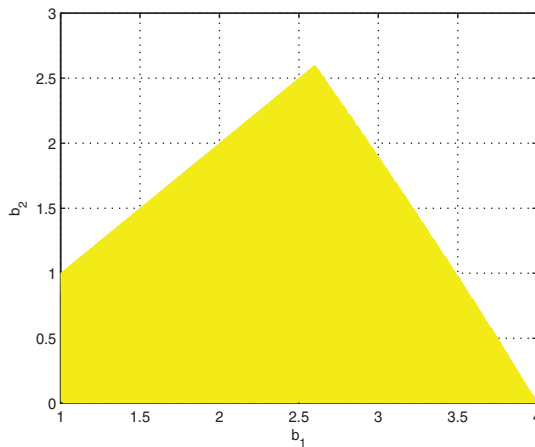


FIG. 4.1. Stabilizability region of  $(b_1, b_2)$  for  $n = 2$ .

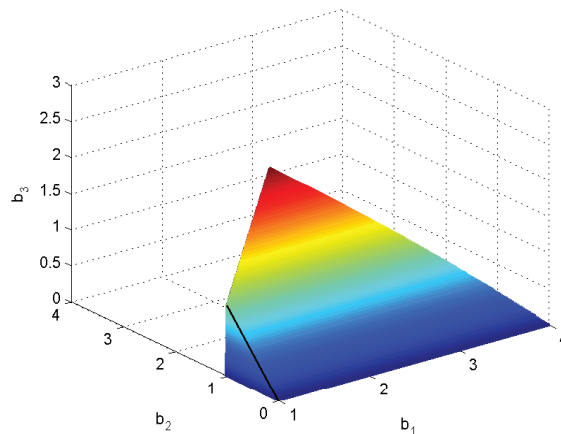


FIG. 4.2. Stabilizability region of  $(b_1, b_2, b_3)$  for  $n = 3$ .

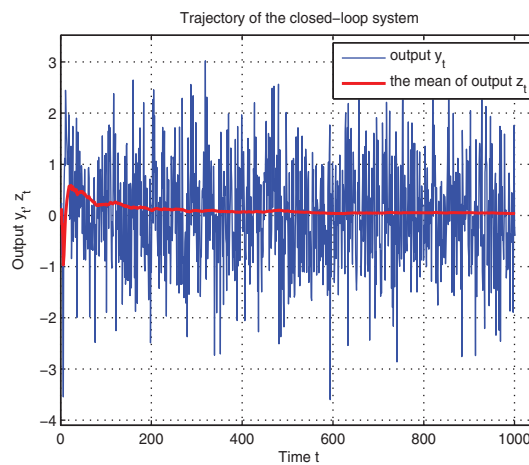


FIG. 4.3. Trajectory of closed-loop system for  $n = 2$ . Process  $y_t$  is the output with initial value  $y_0 = 0$ , and  $z_t = Ey_t$  is the mean of the output. Noise  $\{w_t\}$  takes its value from an i.i.d sequence with standard normal distribution  $N(0, 1)$ .

Thus, by Theorem 2.1, system (4) is stabilizable. As a matter of fact, system (4) indeed can be stabilized by a feedback controller based on the recursive LS algorithm (7)–(9). A trajectory of the closed-loop system (4) and (7)–(9) for this case is shown in Figure 4.3.

*Example 4.2.* Let  $n = 3$ . Consider system (4) with  $b_1 = \frac{7}{3}$ ,  $b_2 = 2$ , and  $b_3 = 1$ . Let noise sequence  $\{w_t\}$  and unknown parameter vector  $\theta$  satisfy assumptions (A1) and (A2), where the mean value of  $\theta$  is given by  $(0, 0, 0)$ . Polynomial  $P(x)$  defined by (6) is thus

$$P(x) = x^4 - \frac{7}{3}x^3 + \frac{1}{3}x^2 + x + 1.$$

By some simple calculations, it also can be verified that  $P(x) > 0$  for any  $x \in (0, 1)$ . Again, Theorem 2.1 implies the stabilizability of system (4). Figure 4.4 shows a trajectory of the closed-loop system (4) and (7)–(9) in the current case.

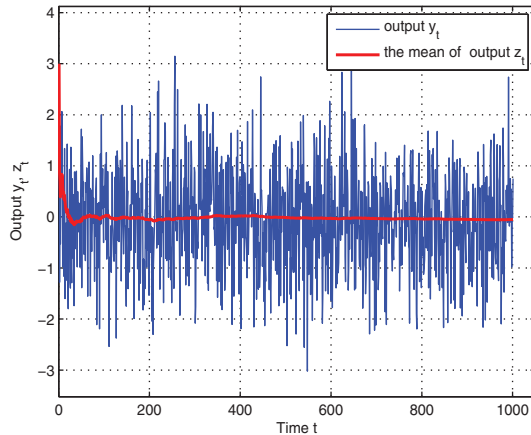


FIG. 4.4. Trajectory of closed-loop system for  $n = 3$ . Process  $y_t$  is the output with initial value  $y_0 = 3$ , and  $z_t = Ey_t$  is the mean of the output. Noise  $\{w_t\}$  takes its value from an i.i.d sequence with standard normal distribution  $N(0, 1)$ .

**5. Concluding remarks.** In this paper, we have studied a basic class of stochastic nonlinear discrete-time systems with multiple unknown parameters. We proved that the systems can be stabilized by a feedback controller based on a recursive LS algorithm. This result together with [20] gives a full characterization of feedback limitations for this basic class of uncertain systems.

**Appendix A.** We prove the key fact that the minimal eigenvalue of  $P_t^{-1}$  has the same growth rate as  $t$ . Now, for any  $x = (x_1, \dots, x_n) \in \mathbb{C}^n$  with  $\|x\| = 1$ , define

$$(47) \quad f_{x,b}(y) \triangleq \sum_{i=1}^n x_i y^{b_i},$$

where  $b = (b_1, \dots, b_n) \in \mathbb{R}^n$  with  $b_1 > b_2 > \dots > b_n > 0$  defined as before. Note that if  $b_i = \frac{q}{2p}$ ,  $p, q \in \mathbb{Z}^+$  or  $b_i$  is irrational for some  $i$ , system (4) has no meaning for initial value  $y_0 < 0$ . So, we only consider the case where all  $b_i = \frac{q_i}{p_i}$  for some odd  $p_i > 0$  and integer  $q_i > 0$ .

Given rational vector  $b = (\frac{q_1}{p_1}, \dots, \frac{q_n}{p_n})$ , let  $d_i = \frac{q_i h}{p_i}$  and  $h = \prod_{i=1}^n p_i$ , then  $f_{x,b}(y)$  can be rewritten by

$$(48) \quad f_{x,b}(y) = \sum_{i=1}^n x_i y^{\frac{d_i}{h}} = \sum_{i=1}^n x_i z^{d_i} \triangleq f_{x,d}(z),$$

where  $z = y^{\frac{1}{h}}$  and  $d = (d_1, \dots, d_n)$ . Let  $\bar{i}$  be the smallest subscript of  $x_i$  such that  $x_i \neq 0$ . Define  $\bar{x}_1 \triangleq x_{\bar{i}}$  and  $\bar{d}_1 \triangleq d_{\bar{i}}$ . By the fundamental theorem of algebra, the polynomial function  $f_{x,d}(z) = 0$  has  $\bar{d}_1$  zeros (counting multiplicity) in  $\mathbb{C}$ , and let  $\nu_{x,1}, \dots, \nu_{x,\bar{d}_1}$  be the  $\bar{d}_1$  roots. Denote  $\text{Re}(\nu_{x,j})$  and  $\text{Im}(\nu_{x,j})$  as the real part and imaginary part of  $\nu_{x,j}$ ,  $j = 1, 2, \dots, \bar{d}_1$ , respectively. Let  $\mu_{x,j} \triangleq |\text{Re}(\nu_{x,j})|$ ,

$$(49) \quad U_{x,j} \triangleq \begin{cases} \left( -\infty, \left( \frac{\mu_{x,j} - 1}{2} \right)^h \right] \cup \left( \left( \frac{3\mu_{x,j} + 1}{2} \right)^h, \infty \right), & \text{Re}(\nu_{x,j}) \geq 0, \\ \left( -\infty, -\left( \frac{3\mu_{x,j} + 1}{2} \right)^h \right] \cup \left( -\left( \frac{\mu_{x,j} - 1}{2} \right)^h, \infty \right), & \text{Re}(\nu_{x,j}) < 0, \end{cases}$$

and

$$(50) \quad U_x \triangleq \left( \bigcap_{j=1}^{\bar{d}_1} U_{x,j} \right).$$

LEMMA A.1. For any  $y \in U_x$ ,  $|f_{x,b}(y)| \geq \frac{1}{\sqrt{n}2^{2\bar{d}_1}}$ .

*Proof.* Since  $f_{x,d}(z)$  is a polynomial with roots  $\nu_{x,j}$ ,  $j = 1, 2, \dots, \bar{d}_1$ , we have

$$(51) \quad f_{x,d}(z) = x_1 \prod_{j=1}^{\bar{d}_1} (z - \nu_{x,j}).$$

Note that  $h$  is an odd number; by (49), it can be checked that for any  $y \in U_{x,j} \subset \mathbb{R}$ ,

$$\begin{aligned} |z - \nu_{x,j}| &= \sqrt{\left(y^{\frac{1}{h}} - \operatorname{Re}(\nu_{x,j})\right)^2 + \operatorname{Im}^2(\nu_{x,j})} \\ &\geq \sqrt{\left(\frac{|\operatorname{Re}(\nu_{x,j})| + 1}{2}\right)^2 + \operatorname{Im}^2(\nu_{x,j})} \\ &\geq \frac{\sqrt{|\nu_{x,j}|^2 + 1}}{2}. \end{aligned}$$

As a result, by (51), we have for any  $y \in U_x$  that

$$(52) \quad |f_{x,d}(z)| \geq \frac{|\bar{x}_1|}{2^{\bar{d}_1}} \prod_{j=1}^{\bar{d}_1} \max\{|\nu_{x,j}|, 1\}.$$

Now, we estimate  $\bar{x}_1$  in terms of  $\nu_{x,j}$ ,  $1 \leq j \leq \bar{d}_1$ . Observe that there at least exists an  $x_i$  with  $i \geq \bar{i}$  and  $|x_i| \geq \frac{1}{\sqrt{n}}$ . If  $i > \bar{i}$ , from the relationships between roots and coefficients (Viète's formula), we have

$$\left| \sum_{1 \leq k_1 < \dots < k_{(\bar{d}_1 - d_i)} \leq \bar{d}_1} \prod_{j=1}^{\bar{d}_1 - d_i} \nu_{x,k_j} \right| = \left| \frac{x_i}{\bar{x}_1} \right|.$$

Consequently, by the fact  $C_{\bar{d}_1}^{\bar{d}_1 - d_i} \leq 2^{\bar{d}_1}$  and  $|x_i| \geq \frac{1}{\sqrt{n}}$ ,

$$|x_1| \geq \frac{1}{\sqrt{n}2^{\bar{d}_1} \max_{1 \leq k_1 < \dots < k_{(\bar{d}_1 - d_i)} \leq \bar{d}_1} \left| \prod_{j=1}^{\bar{d}_1 - d_i} \nu_{x,k_j} \right|}.$$

Then, by (52),

$$(53) \quad \begin{aligned} |f_{x,d}(z)| &\geq \frac{\prod_{j=1}^{\bar{d}_1} \max\{|\nu_{x,j}|, 1\}}{\sqrt{n}2^{2\bar{d}_1} \max_{1 \leq k_1 < \dots < k_{(\bar{d}_1 - d_i)} \leq \bar{d}_1} \left| \prod_{j=1}^{\bar{d}_1 - d_i} \nu_{x,k_j} \right|} \\ &\geq \frac{1}{\sqrt{n}2^{2\bar{d}_1}}. \end{aligned}$$

On the other hand, if  $i = 1$ , that is,  $|\bar{x}_1| \geq \frac{1}{\sqrt{n}}$ , it is straightforward by (52) to deduce (53). Note that  $\bar{d}_1 \leq d_1$ , therefore,  $|f_{x,b}(y)| \geq \frac{1}{\sqrt{n}2^{2\bar{d}_1}}$  by (48) and (53) if  $y \in U_x$ .  $\square$



Next, we will show that there is a constant  $0 < M < 1$  such that the number of outputs  $y_i, 1 \leq i \leq t$  with

$$|f_{x,b}(y_i)| \geq \frac{1}{\sqrt{n}2^{2d_1}},$$

is larger than  $\frac{Mt}{2}$  for any  $x \in \mathbb{C}^n$  with  $\|x\| = 1$ , if  $t$  is large enough. For this, define the random process  $g_x$  by

$$(54) \quad g_x(i) \triangleq I_{\{y_i \in U_x\}} - P(y_i \in U_x | \mathcal{F}_{i-1}^y), \quad i \geq 1.$$

LEMMA A.2. For any  $\epsilon > 0$ , there is a class  $\mathcal{G}_\epsilon$  such that

(i) each element of  $\mathcal{G}_\epsilon$ , denoted by  $g_\epsilon$ , is a random series  $\{g_\epsilon(i)\}$  with the form

$$(55) \quad g_\epsilon(i) = I_{\{y_i \in U_{\epsilon,i-1}\}} - P(y_i \in U_{\epsilon,i-1} | \mathcal{F}_{i-1}^y) - \epsilon, \quad i \geq 1,$$

where  $U_{\epsilon,i-1} \subset \mathbb{R}$  is some random set;

(ii)  $\mathcal{G}_\epsilon$  contains a lower process  $g_\epsilon$  to each  $g_x$  in the sense that

$$g_\epsilon(i) \leq g_x(i) \quad \forall i \geq 1;$$

(iii) for each  $\omega \in \Omega$  and  $t \geq 1$ , the number of distinct sequence  $\{g_\epsilon(i), 1 \leq i \leq t\}$  is finite.

*Proof.* (i) First, we construct  $\mathcal{G}_\epsilon$ . Observe that both  $\theta$  and  $\{w_t\}$  are Gaussian distributed, and  $\theta_i = E(\theta | \mathcal{F}_i^y)$ , then given  $\mathcal{F}_{i-1}^y$ , the conditional distribution of  $y_i$  is Gaussian with conditional mean and variance, respectively, given by

$$(56) \quad E(y_i | \mathcal{F}_{i-1}^y) = 0 \quad \text{and} \quad \text{Var}(y_i | \mathcal{F}_{i-1}^y) = \frac{|P_i^{-1}|}{|P_{i-1}^{-1}|} = \sigma_{i-1}^2.$$

Let  $\Phi(y)$  and  $\phi(y)$  denote the distribution function and density of the standard normal random variable, respectively, that is,

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{\tau^2}{2}} d\tau \quad \text{and} \quad \phi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}.$$

Now, given  $\epsilon > 0$ , there is a constant  $A_\epsilon > 0$  such that

$$(57) \quad \Phi(-A_\epsilon) < \frac{\epsilon}{4}.$$

Define random scalar  $B_{\epsilon,i} \triangleq A_\epsilon \max_{0 \leq k \leq i} \sigma_k^2$  for any  $i \geq 0$ . Then, by (56), (57), and the symmetry of the standard normal distribution, we have for each  $i \geq 1$ ,

$$(58) \quad \begin{aligned} P(|y_i| \geq B_{\epsilon,i-1} | \mathcal{F}_{i-1}^y) &\leq 2P\left(\frac{y_i}{\sigma_{i-1}} \leq -A_\epsilon \mid \mathcal{F}_{i-1}^y\right) \\ &= 2\Phi(-A_\epsilon) < \frac{\epsilon}{2} \quad a.s. \end{aligned}$$

Further, observe that there is a constant  $\Delta_\epsilon > 0$  such that

$$(59) \quad \frac{1}{\sqrt{2\pi}} \int_{-d_1 \Delta_\epsilon}^{d_1 \Delta_\epsilon} e^{-\frac{\tau^2}{2}} d\tau < \frac{\epsilon}{2};$$

we equally divide the real line  $\mathbb{R}$  into a series of left-open and right-closed intervals  $\mathcal{I}_{\mathbb{R}} = \{I_s\}$  such that  $I_s \cap I_k = \emptyset$  for  $s \neq k$ ,  $\bigcup_s I_s = \mathbb{R}$ , and each  $I_s$  has length  $\Delta_\epsilon$ . Let  $U_\epsilon$  be a union of these intervals (finite or infinitely many) and let

$$(60) \quad U_{\epsilon,i} \triangleq U_\epsilon \cap (-B_{\epsilon,i}, B_{\epsilon,i}], \quad i \geq 0.$$

Then, for any fixed  $U_\epsilon$ , we can define a random process  $g_\epsilon$  by

$$g_\epsilon(i) \triangleq I_{\{y_i \in U_{\epsilon,i-1}\}} - P(y_i \in U_{\epsilon,i-1} | \mathcal{F}_{i-1}^y) - \epsilon \quad \text{for all } i \geq 1.$$

This in fact means that each  $g_\epsilon$  is determined by a  $U_\epsilon$ . Denote  $\mathcal{G}_\epsilon$  as the class of all  $g_\epsilon$ ; then, property (i) is satisfied.

(ii) By the definition of  $U_x$  in (50), it is easy to see that  $U_x$  is a union of at most  $d_1 + 1$  intervals (bounded or semiunbounded). Then, for any  $g_x$ , there is a  $U_\epsilon$  such that  $U_\epsilon \subset U_x$  and  $U_x - U_\epsilon$  falls into a union set of at most  $2d_1$  intervals  $\{I_{s_j}\} \subset \mathcal{I}_{\mathbb{R}}$ . Denote this union set by  $\Delta U_{x,\epsilon}$  for the given  $g_x$ . Consequently, for all  $i \geq 1$ , we have by (60) that

$$(61) \quad U_{\epsilon,i-1} \subset U_x \cap (-B_{\epsilon,i-1}, B_{\epsilon,i-1}]$$

and

$$(62) \quad (U_x \cap (-B_{\epsilon,i-1}, B_{\epsilon,i-1}]) - U_{\epsilon,i-1} = (U_x - U_\epsilon) \cap (-B_{\epsilon,i-1}, B_{\epsilon,i-1}] \subset \Delta U_{x,\epsilon}.$$

Now, for any  $S = \bigcup_j (a_j, b_j]$  and  $\sigma > 0$ , let  $\frac{S}{\sigma} \triangleq \bigcup_j (\frac{a_j}{\sigma}, \frac{b_j}{\sigma}]$  and  $|S| \triangleq \sum_j |b_j - a_j|$ . Note that  $\sigma_{i-1} \geq 1$  and the number of intervals  $\{I_{s_j}\}$  with length  $\Delta_\epsilon$  in  $\Delta U_{x,\epsilon}$  is no more than  $2d_1$ ; it is easy to see that

$$\left| \frac{\Delta U_{x,\epsilon}}{\sigma_{i-1}} \right| \leq 2d_1 \Delta_\epsilon,$$

which together with (59) yields

$$(63) \quad \begin{aligned} P\left(\frac{y_i}{\sigma_{i-1}} \in \frac{\Delta U_{x,\epsilon}}{\sigma_{i-1}} \middle| \mathcal{F}_{i-1}^y\right) &= \frac{1}{\sqrt{2\pi}} \int_{\frac{\Delta U_{x,\epsilon}}{\sigma_{i-1}}} e^{-\frac{\tau^2}{2}} d\tau \\ &\leq \frac{1}{\sqrt{2\pi}} \int_{-d_1 \Delta_\epsilon}^{d_1 \Delta_\epsilon} e^{-\frac{\tau^2}{2}} d\tau \\ &< \frac{\epsilon}{2}. \end{aligned}$$

Since the set  $(U_x \cap (-B_{\epsilon,i-1}, B_{\epsilon,i-1}]) - U_{\epsilon,i-1}$  is a union of some left-open and right-closed intervals, by (62),

$$\frac{(U_x \cap (-B_{\epsilon,i-1}, B_{\epsilon,i-1}]) - U_{\epsilon,i-1}}{\sigma_{i-1}} \subset \frac{\Delta U_{x,\epsilon}}{\sigma_{i-1}},$$

then, from (63), we have

$$(64) \quad \begin{aligned} &P(y_i \in (U_x \cap (-B_{\epsilon,i-1}, B_{\epsilon,i-1}]) - U_{\epsilon,i-1} | \mathcal{F}_{i-1}^y) \\ &= P\left(\frac{y_i}{\sigma_{i-1}} \in \frac{(U_x \cap (-B_{\epsilon,i-1}, B_{\epsilon,i-1}]) - U_{\epsilon,i-1}}{\sigma_{i-1}} \middle| \mathcal{F}_{i-1}^y\right) \\ &\leq P\left(\frac{y_i}{\sigma_{i-1}} \in \frac{\Delta U_{x,\epsilon}}{\sigma_{i-1}} \middle| \mathcal{F}_{i-1}^y\right) < \frac{\epsilon}{2}. \end{aligned}$$

Further, according to (61),

$$(65) \quad U_x = U_{\epsilon, i-1} \cup (U_x \cap (-B_{\epsilon, i-1}, B_{\epsilon, i-1}] - U_{\epsilon, i-1}) \cup (U_x \cap (-B_{\epsilon, i-1}, B_{\epsilon, i-1}]^c);$$

as a result, by (58) and (64), for any  $g_x$  we have

$$\begin{aligned} g_x(i) &= I_{\{y_i \in U_x\}} - P(y_i \in U_x | \mathcal{F}_{i-1}^y) \\ &= I_{\{y_i \in U_x\}} - P(y_i \in U_{\epsilon, i-1} | \mathcal{F}_{i-1}^y) \\ &\quad - P(y_i \in (U_x \cap (-B_{\epsilon, i-1}, B_{\epsilon, i-1}]) - U_{\epsilon, i-1} | \mathcal{F}_{i-1}^y) \\ &\quad - P(y_i \in U_x \cap (-B_{\epsilon, i-1}, B_{\epsilon, i-1}]^c | \mathcal{F}_{i-1}^y) \\ &\geq g_\epsilon(i) \quad \forall i \geq 1. \end{aligned}$$

Therefore, property (ii) also holds.

(iii) Fix  $\omega \in \Omega$  and  $t \geq 1$ . Obviously,  $\max_{0 \leq k \leq t-1} \sigma_k^2$  is finite and so is  $B_{\epsilon, t-1} = A_\epsilon \max_{0 \leq k \leq t-1} \sigma_k^2$ . Since there are at most  $\lfloor \frac{2B_{\epsilon, t-1}}{\Delta_\epsilon} \rfloor$  intervals  $I_s \in \mathcal{I}_\mathbb{R}$  belonging to  $(-B_{\epsilon, t-1}, B_{\epsilon, t-1}]$ , by the definitions of  $U_\epsilon$  and  $U_{\epsilon, t-1}$ , the number of distinct  $U_{\epsilon, t-1}$  must be finite. Note that for any  $1 \leq i \leq t$ ,

$$(-B_{\epsilon, i-1}, B_{\epsilon, i-1}] \subset (-B_{\epsilon, t-1}, B_{\epsilon, t-1}];$$

by (60), we immediately have

$$U_{\epsilon, i-1} = U_{\epsilon, t-1} \cap (-B_{\epsilon, i-1}, B_{\epsilon, i-1}].$$

Hence, by (55), for all  $1 \leq i \leq t$ ,

$$g_\epsilon(i) = I_{\{y_i \in (U_{\epsilon, t-1} \cap (-B_{\epsilon, i-1}, B_{\epsilon, i-1}])\}} - P(y_i \in (U_{\epsilon, t-1} \cap (-B_{\epsilon, i-1}, B_{\epsilon, i-1}]) | \mathcal{F}_{i-1}^y) - \epsilon,$$

which asserts that the numbers of distinct  $\{g_\epsilon(i), 1 \leq i \leq t\}$  and  $U_{\epsilon, t-1}$  are equal.  $\square$

LEMMA A.3 ([5, Corollary 3]). *Let  $\{(X_{ti}, \mathcal{F}_{ti}), i \geq 1\}$  be a sequence of martingale differences such that*

$$\sup_{t,i} E(X_{ti}^2 | \mathcal{F}_{t,i-1}^y) \leq K \quad \text{a.s.}$$

*for some constant  $K$ . Let  $c > 0, p, q \in \mathbb{R}$  be given numbers with  $p > |q|$ , and let  $\{b_{ti}\}$  be an array of real numbers satisfying*

$$(66) \quad |b_{ti}| \leq ct^{-p}i^{-q}, \quad 1 \leq i \leq t, \quad b_{ti} = 0, \quad i > t,$$

*and*

$$(67) \quad \sum_{i=1}^t b_{ti}^2 = O(t^{-\alpha}) \quad \text{for some } \alpha > 0.$$

*If there is a random variable  $W$  such that  $E|W|^{\frac{2}{p+q}} < \infty$  and*

$$\sup_{t,i} P\{|X_{ti}| \geq x\} \leq CP\{|W| \geq x\} \quad \forall x > 0,$$

for some finite constant  $C$ , then,

$$\sum_{t=1}^{\infty} P \left\{ \sup_{1 \leq k \leq t} \left| \sum_{i=1}^k b_{ti} X_{ti} \right| > \varepsilon \right\} < \infty \quad \forall \varepsilon > 0.$$

LEMMA A.4. Given  $\epsilon > 0$ , let  $\{V_{\epsilon,i}^t, 0 \leq i \leq t, t \geq 0\}$  be a class of random sets and  $\{g_{\epsilon,t}, t \geq 1\}$  be a series of random processes with the form

$$(68) \quad g_{\epsilon,t}(i) = I_{\{y_i \in V_{\epsilon,i-1}^{t-1}\}} - P(y_i \in V_{\epsilon,i-1}^{t-1} | \mathcal{F}_{i-1}^y) - \epsilon, \quad 1 \leq i \leq t.$$

Then,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^t g_{\epsilon,t}(i) + \epsilon = 0 \quad a.s.$$

*Proof.* Fix  $\epsilon > 0$ . By (68), it is easy to see that for any  $t \geq 1$ ,

$$\begin{aligned} E(g_{\epsilon,t}(i) + \epsilon | \mathcal{F}_{i-1}^y) \\ &= E \left( I_{\{y_i \in V_{\epsilon,i-1}^{t-1}\}} - P(y_i \in V_{\epsilon,i-1}^{t-1} | \mathcal{F}_{i-1}^y) | \mathcal{F}_{i-1}^y \right) \\ &= 0, \quad 1 \leq i \leq t, \end{aligned}$$

which means  $\{(g_{\epsilon,t}(i) + \epsilon, \mathcal{F}_i^y), 1 \leq i \leq t\}$  is a martingale difference sequence. Thus,  $\{(g_{\epsilon,t}(i) + \epsilon, \mathcal{F}_i^y), 1 \leq i \leq t, t \geq 1\}$  forms a triangular array of martingale differences.

Note that  $|g_{\epsilon,t}(i) + \epsilon| \leq 2$ ; then

$$\sup_{t,i} E((g_{\epsilon,t}(i) + \epsilon)^2 | \mathcal{F}_{i-1}^y) \leq 4.$$

If we let  $c = 1$ ,  $p = 1$ , and  $q = 0$  in Lemma A.3, and define

$$b_{ti} \triangleq \begin{cases} \frac{1}{t}, & i \leq t, \\ 0, & i > t, \end{cases}$$

the sequence  $\{b_{ti}\}$  satisfies (66) and (67) obviously. Let  $W = 3$ ; hence  $E|W|^2 < \infty$  and for any  $x > 0$ ,

$$\sup_{t,i} P\{|g_{\epsilon,t}(i) + \epsilon| \geq x\} \leq P\{|W| \geq x\}.$$

Therefore, by Lemma A.3, we immediately obtain that  $\frac{1}{t} \sum_{i=1}^t (g_{\epsilon,t}(i) + \epsilon)$  converges completely to 0, and hence with probability 1,

$$\frac{1}{t} \sum_{i=1}^t g_{\epsilon,t}(i) + \epsilon \rightarrow 0.$$

The lemma is thus proved.  $\square$

LEMMA A.5. Let  $g_x$  be defined by (54); then

$$\liminf_{t \rightarrow \infty} \inf_x \frac{1}{t} \sum_{i=1}^t g_x(i) \geq 0 \quad a.s.$$

*Proof.* Let  $\mathcal{G}_\epsilon$  be the class of processes  $g_\epsilon$  defined in Lemma A.2. By Lemma A.2 (i)–(ii), for each  $x \in \mathbb{C}^n$  with  $\|x\| = 1$ , there is a  $g_\epsilon^x \in \mathcal{G}_\epsilon$  with form (55) such that  $g_x(i) \geq g_\epsilon^x(i)$  for all  $i \geq 1$ . Note that for each sample point  $\omega \in \Omega$  and each  $t \geq 1$ , the number of distinct sequences  $\{g_\epsilon(i), 1 \leq i \leq t\}$  is finite by Lemma A.2 (iii); then, there exists a random process  $g_{\epsilon,t}$  such that for any  $g_\epsilon \in \mathcal{G}_\epsilon$ ,

$$(69) \quad \frac{1}{t} \sum_{i=1}^t g_{\epsilon,t}(i) \leq \frac{1}{t} \sum_{i=1}^t g_\epsilon(i) \quad \forall t \geq 1.$$

The random array  $\{(g_{\epsilon,t}(i), 1 \leq i \leq t, t \geq 1)\}$  is picked by the form

$$g_{\epsilon,t}(i) = I_{\{y_i \in V_{\epsilon,i-1}^{t-1}\}} - P(y_i \in V_{\epsilon,i-1}^{t-1} | \mathcal{F}_{i-1}^y) - \epsilon,$$

where  $V_{\epsilon,i-1}^{t-1} = U_{\epsilon,i-1}, 1 \leq i \leq t$  on each  $\omega \in \Omega$  for some sequence  $\{U_{\epsilon,i-1}, 1 \leq i \leq t\}$ . However, the random series  $\{V_{\epsilon,i-1}^{t-1}\}$  does not necessarily equal any sequence of random sets  $\{U_{\epsilon,i-1}\}$ . This is because each  $\{U_{\epsilon,i-1}\}$  is determined by a unique  $U_\epsilon$  according to (60), while  $\{V_{\epsilon,i-1}^{t-1}\}$  corresponds to different  $U_\epsilon$  as  $\omega$  and  $t$  vary. Consequently, by (69) and Lemma A.4,

$$\begin{aligned} \liminf_{t \rightarrow \infty} \inf_x \frac{1}{t} \sum_{i=1}^t g_x &\geq \liminf_{t \rightarrow \infty} \inf_x \frac{1}{t} \sum_{i=1}^t g_\epsilon^x(i) \\ &\geq \liminf_{t \rightarrow \infty} \frac{1}{t} \sum_{i=1}^t g_{\epsilon,t}(i) \\ &= \lim_{t \rightarrow \infty} \left( \frac{1}{t} \sum_{i=1}^t g_{\epsilon,t}(i) + \epsilon \right) - \epsilon \\ &= -\epsilon \quad \text{a.s.} \end{aligned}$$

By taking each real numbers  $\epsilon$ , the above inequality implies

$$\liminf_{t \rightarrow \infty} \inf_x \frac{1}{t} \sum_{i=1}^t g_x(i) \geq 0 \quad \text{a.s.,}$$

which completes the proof.  $\square$

LEMMA A.6. *There is a constant  $0 < M < 1$  such that for any  $x \in \mathbb{C}^n$  with  $\|x\| = 1$ ,*

$$\frac{1}{t} \sum_{i=1}^t P(y_i \in U_x | \mathcal{F}_{i-1}^y) \geq M.$$

*Proof.* From (49), for any  $1 \leq j \leq \bar{d}_1$ , we have

$$(70) \quad U_{x,j}^c = \begin{cases} \left[ \left( \left( \frac{\mu_{x,j} - 1}{2} \right)^h, \left( \frac{3\mu_{x,j} + 1}{2} \right)^h \right) \right], & \text{Re}(\nu_{x,j}) \geq 0, \\ \left[ \left( - \left( \frac{3\mu_{x,j} + 1}{2} \right)^h, - \left( \frac{\mu_{x,j} - 1}{2} \right)^h \right) \right], & \text{Re}(\nu_{x,j}) < 0. \end{cases}$$

Note that if  $\mu_{x,j} \geq 3$ ,

$$(71) \quad \left(\frac{3\mu_{x,j} + 1}{\mu_{x,j} - 1}\right)^h = \left(3 + \frac{4}{\mu_{x,j} - 1}\right)^h \leq 5^h.$$

This illustrates that for any  $i \geq 0$  and  $\mu_{x,j} \geq 3$ ,  $\frac{U_{x,j}^c}{\sigma_i}$  belongs to the set

$$(72) \quad \frac{\bar{U}_{x,j}^c}{\sigma_i} \triangleq \begin{cases} \left[\frac{1}{\sigma_i} \left(\frac{\mu_{x,j} - 1}{2}\right)^h, \frac{5^h}{\sigma_i} \left(\frac{\mu_{x,j} - 1}{2}\right)^h\right], & \text{Re}(\nu_{x,j}) \geq 3, \\ \left[-\frac{5^h}{\sigma_i} \left(\frac{\mu_{x,j} - 1}{2}\right)^h, -\frac{1}{\sigma_i} \left(\frac{\mu_{x,j} - 1}{2}\right)^h\right], & \text{Re}(\nu_{x,j}) \leq -3. \end{cases}$$

Let  $\beta \triangleq \max\left\{\sqrt{\frac{2h \log 5}{5^{2h} - 1}}, 1\right\}$ . It can be checked that when  $|y| \geq \beta$ ,

$$\frac{d}{dy} \int_y^{5^h y} e^{-\frac{\tau^2}{2}} d\tau = 5^h e^{-\frac{5^{2h} y^2}{2}} - e^{-\frac{y^2}{2}} \leq 0,$$

which implies by symmetry that  $\int_y^{5^h y} e^{-\frac{\tau^2}{2}} d\tau$  is a nonincreasing function with respect to  $y$  whenever  $|y| \geq \beta$ . Define the random index sets for any  $i \geq 0$  by

$$J_i \triangleq \left\{j : \frac{1}{\sigma_i} \left(\frac{\mu_{x,j} - 1}{2}\right)^h \geq \beta, 1 \leq j \leq \bar{d}_1\right\} \quad \text{and} \quad J_i^c \triangleq \{1, 2, \dots, \bar{d}_1\} - J_i.$$

Note that given  $i \geq 0$ ,  $\sigma_i \geq 1$ , then, for any  $j \in J_i$ ,

$$\mu_{x,j} \geq 2(\sigma_i \beta)^{\frac{1}{h}} + 1 \geq 3,$$

which yields that for any  $i \geq 0$ ,

$$(73) \quad \frac{U_{x,j}^c}{\sigma_i} \subset \frac{\bar{U}_{x,j}^c}{\sigma_i}, \quad j \in J_i,$$

where  $\frac{\bar{U}_{x,j}^c}{\sigma_i}$  is only defined for  $\mu_{x,j} \geq 3$  by (72).

On the other hand, for any  $i \geq 0$ , if  $j \in J_i^c$ ,

$$\frac{1}{\sigma_i} \left(\frac{\mu_{x,j} - 1}{2}\right)^h < \beta,$$

which yields

$$\mu_{x,j} < 2(\sigma_i \beta)^{\frac{1}{h}} + 1$$

and hence

$$(74) \quad \frac{1}{\sigma_i} \left(\frac{3\mu_{x,j} + 1}{2}\right)^h < \frac{1}{\sigma_i} \left(3(\sigma_i \beta)^{\frac{1}{h}} + 2\right)^h \leq 5^h \beta.$$

Since it is easy to check that

$$\frac{U_{x,j}^c}{\sigma_i} \subset \left[-\frac{1}{\sigma_i} \left(\frac{3\mu_{x,j} + 1}{2}\right)^h, \frac{1}{\sigma_i} \left(\frac{3\mu_{x,j} + 1}{2}\right)^h\right], \quad i \geq 0,$$

by (74), for any  $i \geq 0$ ,

$$\bigcup_{j \in J_i^c} \frac{U_{x,j}^c}{\sigma_i} \subset (-5^h \beta, 5^h \beta).$$

As a result,

$$(75) \quad \bigcup_{1 \leq j \leq d_1} \frac{U_{x,j}^c}{\sigma_i} \subset \left( \bigcup_{j \in J_i} \frac{U_{x,j}^c}{\sigma_i} \cap (-5^h \beta, 5^h \beta)^c \right) \cup (-5^h \beta, 5^h \beta).$$

Since the cardinality of  $J_i$  is at most  $d_1$ , by the fact  $\int_y^{5^h y} e^{-\frac{\tau^2}{2}} d\tau$  is nonincreasing with respect to  $y$  when  $|y| \geq 5^h \beta \geq \beta$ , we immediately conclude by (73) that for any  $x \in \mathbb{C}^n$  with  $\|x\| = 1$  and  $i \geq 0$ ,

$$(76) \quad \begin{aligned} & \int_{\left(\bigcup_{j \in J_i} \frac{U_{x,j}^c}{\sigma_i}\right) \cup (-5^h \beta, 5^h \beta)^c} e^{-\frac{\tau^2}{2}} d\tau \\ & \leq \int_{\left(\bigcup_{j \in J_i} \frac{\bar{U}_{x,j}^c}{\sigma_i}\right) \cup (-5^h \beta, 5^h \beta)^c} e^{-\frac{\tau^2}{2}} d\tau \\ & \leq \sum_{j=1}^{d_1} \left( \int_{5^{hj} \beta}^{5^{h(j+1)} \beta} e^{-\frac{\tau^2}{2}} d\tau + \int_{-5^{h(j+1)} \beta}^{-5^{hj} \beta} e^{-\frac{\tau^2}{2}} d\tau \right) \\ & \leq \int_{5^h \beta}^{5^{h(d_1+1)} \beta} e^{-\frac{\tau^2}{2}} d\tau + \int_{-5^{h(d_1+1)} \beta}^{-5^h \beta} e^{-\frac{\tau^2}{2}} d\tau. \end{aligned}$$

Consequently, for any  $i \geq 1$  we have by (75) and (76) that

$$(77) \quad \begin{aligned} P \left( y_i \in \bigcup_{1 \leq j \leq d_1} U_{x,j}^c \mid \mathcal{F}_{i-1}^y \right) &= P \left( \frac{y_i}{\sigma_{i-1}} \in \bigcup_{1 \leq j \leq d_1} \frac{U_{x,j}^c}{\sigma_{i-1}} \mid \mathcal{F}_{i-1}^y \right) \\ &\leq P \left( \frac{y_i}{\sigma_{i-1}} \in \left( \bigcup_{j \in J_{i-1}} \frac{U_{x,j}^c}{\sigma_{i-1}} \right) \cap (-5^h \beta, 5^h \beta)^c \mid \mathcal{F}_{i-1}^y \right) \\ &\quad + P \left( \frac{y_i}{\sigma_{i-1}} \in (-5^h \beta, 5^h \beta) \mid \mathcal{F}_{i-1}^y \right) \\ &\leq \frac{1}{\sqrt{2\pi}} \int_{5^h \beta}^{5^{h(d_1+1)} \beta} e^{-\frac{\tau^2}{2}} d\tau \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_{-5^{h(d_1+1)} \beta}^{-5^h \beta} e^{-\frac{\tau^2}{2}} d\tau + \frac{1}{\sqrt{2\pi}} \int_{-5^h \beta}^{5^h \beta} e^{-\frac{\tau^2}{2}} d\tau \\ &= \frac{1}{\sqrt{2\pi}} \int_{-5^{h(d_1+1)} \beta}^{5^{h(d_1+1)} \beta} e^{-\frac{\tau^2}{2}} d\tau. \end{aligned}$$

Let  $M \triangleq 1 - \frac{1}{\sqrt{2\pi}} \int_{-5^{h(d_1+1)} \beta}^{5^{h(d_1+1)} \beta} e^{-\frac{\tau^2}{2}} d\tau$ ; obviously,  $M > 0$ . Then, by (50) and (77), for any  $i \geq 1$ ,

$$(78) \quad P(y_i \in U_x \mid \mathcal{F}_{i-1}^y) = 1 - \left( y_i \in \bigcup_{1 \leq j \leq d_1} U_{x,j}^c \mid \mathcal{F}_{i-1}^y \right) \geq M,$$

and hence

$$\frac{1}{t} \sum_{i=1}^t P(y_i \in U_x | \mathcal{F}_{i-1}^y) \geq M.$$

The proof is completed.  $\square$

LEMMA A.7. For any  $x \in \mathbb{C}^n$  with  $\|x\| = 1$ , there is a random  $t_0$ , which is independent of  $x$ , such that whenever  $t > t_0$ , the number of

$$|f_{x,b}(y_i)| \geq \frac{1}{\sqrt{n}2^{2d_1}}, \quad 1 \leq i \leq t,$$

is at least  $\frac{Mt}{2}$ , where  $M$  is defined by Lemma A.6.

*Proof.* By Lemma A.5, there is a  $t_0 > 0$  such that for all  $t > t_0$ ,

$$\inf_x \frac{1}{t} \sum_{i=1}^t g_x(i) \geq -\frac{M}{2}.$$

Hence, for any  $x \in \mathbb{C}^n$  with  $\|x\| = 1$ ,

$$\frac{1}{t} \sum_{i=1}^t I_{\{y_i \in U_x\}} \geq \frac{1}{t} \sum_{i=1}^t P(y_i \in U_x | \mathcal{F}_{i-1}^y) - \frac{M}{2}.$$

By Lemma A.6, we then have for all  $t \geq t_0$  that

$$\frac{1}{t} \sum_{i=1}^t I_{\{y_i \in U_x\}} \geq \frac{M}{2},$$

which yields

$$\sum_{i=1}^t I_{\{y_i \in U_x\}} \geq \frac{Mt}{2}.$$

This means the number of  $y_i \in U_x$  is at least  $\frac{Mt}{2}$ . Therefore, by Lemma A.1, we immediately deduce the result.  $\square$

Now, we can draw our main conclusion of Appendix A.

LEMMA A.8. Let  $\lambda_{\min}(t+1)$  denote the minimal eigenvalue of  $P_{t+1}^{-1}$ . Then, there is a random constant  $C > 0$  such that for all  $t \geq t_0$ ,

$$\lambda_{\min}(t+1) \geq Ct,$$

where  $t_0$  is defined by Lemma A.7.

*Proof.* By the Rayleigh-Ritz theorem,

$$\lambda_{\min}(t+1) \geq \lambda_{\min} \left( \sum_{i=1}^t \phi_i \phi_i^\tau \right) = \min_{|x|=1, x \in \mathbb{C}^n} x^* \left( \sum_{i=1}^t \phi_i \phi_i^\tau \right) x.$$

Note that  $\phi_i^\tau x = f_{x,b}(y_i)$ ; by Lemma A.7, we have for all  $t \geq t_0$  that

$$\begin{aligned} \lambda_{\min}(t+1) &\geq \min_{|x|=1, x \in \mathbb{C}^n} \sum_{i=1}^t |f_{x,b}(y_i)|^2 \\ &\geq \frac{Mt}{\sqrt{n}2^{2d_1+1}}, \end{aligned}$$

which yields the claim if we take the positive constant  $C = \frac{M}{\sqrt{n}2^{2d_1+1}}$ .  $\square$



**Acknowledgments.** The authors wish to thank Professors Lei Guo and Liang-Liang Xie for their valuable comments and Dr. Tiantian Jiang for her kind help with the simulations.

## REFERENCES

- [1] K. J. ÅSTRÖM AND B. WITTENMARK, *Adaptive Control*, 2nd ed., Addison-Wesley, Reading, MA, 1995.
- [2] H. CHEN AND L. GUO, *Identification and Stochastic Adaptive Control*, Birkhäuser, Boston, MA, 1991.
- [3] L. GUO, *Convergence and logarithm laws of self-tuning regulators*, Automatica J. IFAC, 31 (1995), pp. 435–450.
- [4] L. GUO, *On critical stability of discrete-time adaptive nonlinear control*, IEEE Trans. Automat. Control, 42 (1997), pp. 1488–1499.
- [5] S. GHOSAL AND T. K. CHANDRA, *Complete convergence of martingale arrays*, J. Theoret. Probab., 11 (1998), pp. 621–631.
- [6] G. C. GOODWIN AND K. S. SIN, *Adaptive Filtering, Prediction and Control*, Prentice-Hall, Englewood Cliffs, NJ, 1984.
- [7] P. A. IOANNOU AND J. SUN, *Robust Adaptive Control*, Prentice-Hall, Englewood Cliffs, NJ, 1996.
- [8] A. KOJIC AND A. M. ANNASWAMY, *Adaptive control of nonlinearly parametrized systems with a triangular structure*, Automatica J. IFAC, 38 (2002), pp. 115–123.
- [9] I. KANELAKOPOULOS, *A discrete-time adaptive nonlinear system*, IEEE Trans. Automat. Control, 39 (1994), pp. 2362–2364.
- [10] M. KRSTIĆ, I. KANELAKOPOULOS, AND P. V. KOKOTOVIĆ, *Non-linear and Adaptive Control Design*, Wiley, New York, 1995.
- [11] H. B. MA, *An “impossibility” theorem on a class of high-order discrete-time nonlinear control systems*, Systems Control Lett., 57 (2008), pp. 497–504.
- [12] H. B. MA, *Further results on limitations to the capability of feedback*, Internat. J. Control, 81 (2008), pp. 21–42.
- [13] G. A. MILLIKEN AND F. AKDENIZ, *A theorem on the difference of the generalized inverses of two nonnegative definite matrices*, Commun. Statist. Theory Methods, 6 (1977), pp. 73–79.
- [14] C. LI AND L.-L. XIE, *On robust stability of discrete-time adaptive nonlinear control*, Systems Control Lett., 55 (2006), pp. 452–458.
- [15] C. LI AND L. GUO, *A new critical theorem for adaptive nonlinear stabilization*, Automatica J. IFAC, 46 (2010), pp. 999–1007.
- [16] C. LI AND L. GUO, *A dynamical inequality for the output of uncertain nonlinear systems*, Science China, 56 (2013), pp. 1–9.
- [17] C. LI AND L. GUO, *On feedback capability in a class of nonlinearly parameterized uncertain systems*, IEEE Trans. Automat. Control, 56 (2001), pp. 2946–2951.
- [18] C. LI, L.-L. XIE, AND L. GUO, *A polynomial criterion for adaptive stabilizability of discrete-time nonlinear systems*, Commun. Inf. Systems, 6 (2006), pp. 273–298.
- [19] G. TAO AND P. V. KOKOTOVIĆ, *Adaptive Control of Systems with Actuator and Sensor Non-linearities*, Wiley, New York, 1996.
- [20] L.-L. XIE AND L. GUO, *Fundamental limitations of discrete-time adaptive nonlinear control*, IEEE Trans. Automat. Control, 44 (1999), pp. 1777–1782.
- [21] L.-L. XIE AND L. GUO, *Adaptive control of discrete-time nonlinear systems with structural uncertainties*, in Lectures on Systems, Control, and Information, AMS/IP Studies in Adv. Math. 17, AMS, Providence, RI, 2000, pp. 49–90.
- [22] L.-L. XIE AND L. GUO, *How much uncertainty can be dealt with by feedback?*, IEEE Trans. Automat. Control, 45 (2000), pp. 2203–2217.
- [23] F. XUE, L. GUO, AND M. Y. HUANG, *Towards understanding the capability of adaptation for time-varying systems*, Automatica J. IFAC, 37 (2001), pp. 1551–1560.
- [24] F. XUE AND L. GUO, *On limitations of the sampled-data feedback for nonparametric dynamical systems*, J. Systems Sci. Complexity, 15 (2002), pp. 225–25.
- [25] Y. X. ZHANG AND L. GUO, *A limit to the capability of feedback*, IEEE Trans. Automat. Control, 47 (2002), pp. 687–692.