Global Stabilizability Theorems on Discrete-Time Nonlinear Uncertain Systems

Zhaobo Liu and Chanying Li *Member, IEEE*

***Abstract*—This paper focuses on the stabilizability problem for a basic class of discrete-time nonlinear systems with multiple unknown parameters. We claim that such a system is stabilizable if its nonlinear growth rate is dominated by a polynomial rule. This rule cannot be relaxed in general since it becomes a necessary and sufficient condition when the system has a polynomial form [10]. We further prove that the concerned stabilizable system is possible to grow exponentially fast. Meanwhile, optimality and closed-loop identification are also discussed herein.**

***Index Terms*—stochastic adaptive control, stabilizability, nonlinear systems, discrete-time, least squares**

I. Introduction

Adaptive control of linear systems ([1], [3]) and nonlinear systems growing linearly ([17], [19]) has been a mature topic for decades, both in continuous time and discrete time. But when it comes to systems whose output nonlinearities are faster than linearities, the similarities of adaptive control between continuous- and discrete-time systems disappear. Most continuous-time nonlinear systems can be globally stabilized by employing nonlinear damping or back-stepping techniques ([6], [7]), however, its discrete-time counterpart is not that favored by fortune. It was found in [5] that even for the following basic discrete-time stochastic system

$$y_{t+1} = \theta y_t^b + u_t + w_{t+1}, \tag{1}$$

where θ is a scalar parameter and y_t, u_t, w_t are the output, input and noise signals, respectively, the stabilizability holds if and only if $b < 4$. This fundamental difficulty in discrete-time control was further confirmed by [18], where system (1) is extended to the multi-parameter case:

$$y_{t+1} = \theta_1 y_t^{b_1} + \theta_2 y_t^{b_2} + \cdots + \theta_n y_t^{b_n} + u_t + w_{t+1}. \tag{2}$$

An "impossibility theorem" was established therein by providing a polynomial rule, which was proved, a decade later by [10], to be a necessary and sufficient condition of the stabilizability for system (2). This polynomial rule is characterized by $b_1, \ldots, b_n$, which shows that the nonlinearity plays a key role in determining the stabilizability. Analogous phenomena arise in adaptive control of various discrete-time nonlinear systems. We refer the readers to [11], [13], [14], [15], [20], [21], [22].

We are now intended to extend the results of [10] and [12] to the following class of uncertain systems:

$$y_{t+1} = \theta_1 f_1(y_t) + \theta_2 f_2(y_t) + \cdots + \theta_n f_n(y_t) + u_t + w_{t+1}. \tag{3}$$

It is easy to understand that stabilizability depends on closed-loop identifiability. In addition, the more different the functions $f_i(x)$ and $f_j(x)$ are, the easier the parameters θ_i and θ_j can be distinguished (we cannot distinguish θ_i from θ_j if $f_i(x) \equiv f_j(x)$). However, $f_i(x), i = 1, \ldots, n$ here may be very close to each other or even identical infinitely often. This is an essential difficulty arising in the closed-loop identification, compared with [10] where $f_i(x) = x^{b_i}$, $b_1 > \cdots > b_n > 0$ are quite different when x becomes large. We point out that a trivial extension of the existing methods has trouble coping with the current situation. To overcome this barrier, we establish an inequality on the minimal eigenvalue of the data matrix and the conditional variances of the outputs in Proposition 1. It sets up a bridge between stabilizability and closed-loop identifiability. Based on it, we prove that system (3) is stabilizable if the nonlinear growth rates of $f_1(x), \ldots, f_n(x)$ are dominated by some power functions $x^{b_1}, \ldots, x^{b_n}$ respectively, where $b_1, \ldots, b_n$ satisfy the polynomial rule referred.

Evidently, more unknown parameters means greater uncertainty in system (3). Observe that if the system has a huge number of unknown parameters, the polynomial rule requires the growth rate of the system close to linearity for the sake of the stabilizability. That is, $b_1 \approx 1$ (see [10]). But this is not necessary. We next show that, no matter how large n is, system (3) with exponential growth rate is always possible to be stabilized by discrete-time feedback control. Recall that $n = 1$ in [12], it in fact advances [12] by considering any high-dimensional parameter vector in the system.

The paper is built up as follows. Section II presents the main stabilizability theorems and Section III discusses the control design as well as the corresponding closed-loop identification. The proofs of the main results are included in Sections IV–V.

For ease of reference, Table I below provides a summary of the symbols used in the article.

II. Global Stabilizability

We study the following discrete-time nonlinear system with multiple unknown parameters:

$$y_{t+1} = \theta^T \phi(y_t) + u_t + w_{t+1}, \qquad t \geq 0, \tag{4}$$

This work was supported by the National Natural Science Foundation of China under Grants 11925109 and U21B6001.

Z. Liu is with the College of Computer Science and Software Engineering of Shenzhen University, Shenzhen, P. R. China.

C. Li is with the Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, P. R. China. She is also with the School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, P. R. China.

Corresponding author: Chanying Li (cyli@amss.ac.cn).

TABLE I
SUMMARY OF NOTATION CONVENTIONS USED IN THE ARTICLE

e.g.	Exempli gratia.
i.o.	Infinitely often.
$\sigma\{y_i, 0 \leqslant i \leqslant t\}$	The σ-field generated by $\{y_i, 0 \leqslant i \leqslant t\}$.
$I_{\{\cdot\}}$	The indicator function.
$\lvert A\rvert$	The determinant of matrix A.
$\lVert x\rVert$	The Euclidean norm of vector x.
$f(x) = O(g(x))$ $(x \to \infty)$	There are two constants $\zeta_1, \zeta_2 > 0$ such that for $\lvert x\rvert > \zeta_2$, $\lVert f(x)\rVert \leq \zeta_1 \cdot \lVert g(x)\rVert$.

where the unknown parameters $\theta = (\theta_1, \dots, \theta_n)^T$ takes values in $\mathbb{R}^n, n \geq 2$, y_t, u_t, w_t are the output, input and noise signals, respectively. Here y_0 is a scalar random variable independent of $\{w_t\}$ (y_0 can also be taken deterministic). Assume that $\phi = (f_1, \dots, f_n)^T : \mathbb{R} \to \mathbb{R}^n$ is a known measurable vector function, where $f_j \in C^n(E)$, $1 \leq j \leq n$, and E is an open set in $\mathbb{R}$. We rewrite (4) as

$$y_{t+1} = \sum_{j=1}^{n} \theta_j f_j(y_t) + u_t + w_{t+1}, \qquad t \geq 0, \tag{5}$$

and introduce

Definition 1. *System (5) is said to be globally stabilizable, if there exits a feedback control law*

$$u_t \in \mathcal{F}_t^y \triangleq \sigma\{y_i, 0 \leqslant i \leqslant t\},\ t = 0, 1, \dots \tag{6}$$

such that for any initial y_0,

$$\sup_{t\geq 1} \frac{1}{t} \sum_{i=1}^{t} y_i^2 < +\infty, \quad a.s.. \tag{7}$$

We analyze our problem in some standard assumptions.

- A1 The noise $\{w_t\}$ is an i.i.d random sequence with $w_1 \sim N(0, \sigma_w^2)$.
- A2 Parameter $\theta \sim N(\theta_0, I_n)$ is independent of $\{w_t\}$.
- A3 $f_1, \dots, f_n$ are linearly independent on E.

Remark 1. *Assumption A3 is a natural condition. Consider a typical case where $E = \mathbb{R}$. If $f_j \equiv 0$ for all $j \in [1, n]$, system (5) degenerates to $y_{t+1} = w_{t+1}$. Otherwise, without loss of generality, let $f_1, \dots, f_k$ be linearly independent on $\mathbb{R}$, $1 \leq k \leq n$, such that every $f_l, l \in [k+1, n]$, is a linear combination of $f_1, \dots, f_k$. Consequently, there are $n-k$ unit vectors $(x_{1,l}, \dots, x_{k,l})^T$ satisfying $f_l(y) = \sum_{j=1}^{k} x_{j,l} f_j(y)$, $l \in [k+1, n]$. Therefore, by letting $\theta_j' \triangleq \theta_j + \sum_{l=k}^{n} x_{j,l}\theta_l$, $1 \leq j \leq k$, system (5) becomes*

$$y_{t+1} = \sum_{j=1}^{k} \theta_j' f_j(y_t) + u_t + w_{t+1}, \qquad t \geq 0. \tag{8}$$

This means it suffices to discuss system (8), where all f_i are linearly independent.

Our first theorem extends the result of [10] to a more general situation in the following sense.

Theorem 1. *Under Assumptions A1–A3, let system (5) satisfy*

$$f_j(x) = O(|x|^{b_j}) + O(1), \quad 1 \leq j \leq n, \tag{9}$$

where $b_1 > b_2 > \cdots > b_n > 0$ and $b_1 > 1$. The system is globally stabilizable if for $\forall x \in (1, b_1)$,

$$P(x) = x^{n+1} - b_1 x^n + (b_1 - b_2)x^{n-1} + \cdots + b_n > 0. \tag{10}$$

A simple stabilizing control design for Theorem 1 (including Theorem 2 below) is provided in Section III.

Example 1. *Under Assumptions A1–A2, let*

$$f_1(x) = x^2 \cos x \quad \textit{and} \quad f_2(x) = x\sin x$$

in system (5). Since the images of f_1 and f_2 intersect each other infinitely many times, the stabilizability issue cannot be covered by the existing theory. Now, applying Theorem 1 with $b_1 = 2$ and $b_2 = 1$, we immediately conclude that the system is stablizable. Please see Section III for the control design.

Remark 2. *The polynomial rule (10) becomes a necessary and sufficient condition of the stabilizability when $f_i(x) = x^{b_i}, i = 1, \dots, n$ in the system, as claimed by [10]. So this rule cannot be relaxed in general. But it is not necessary for any system (4). Recall that [12] finds system (4) with $n = 1$ is possible to be stabilized when growing exponentially. The next theorem confirms the phenomenon for the multi-parameter case.*

Theorem 2. *Under Assumptions A1–A3, system (4) is globally stabilizable if*
(i) for some $k_1, k_2 > 0$,

$$\|\phi(x)\| \leqslant k_1 e^{k_2|x|}, \quad \forall x \in \mathbb{R}; \tag{11}$$

(ii) there exists $L > 0$ such that

$$\liminf_{l\to+\infty} \frac{\ell(S_L \cap [-l, l])}{l} > 0, \tag{12}$$

where $S_L \triangleq \{x : \|\phi(x)\| \leq L\}$ and ℓ is the Lebesgue measure.

Clearly, $p_L \triangleq \liminf_{l\to+\infty} \frac{\ell(S_L\cap[-l,l])}{l}$ describes the density of S_L in $\mathbb{R}$. Since $p_L > 0$ can be taken as small as one likes, Theorem 2 shows that a stabilizable system may possess a very sparse S_L. We give an extreme example to illustrate it.

Example 2. *Under Assumptions A1–A2, let*

$$\begin{cases} f_1(x) = 1 + e^x \cdot I_{\{sinx > -0.999\}} \\ f_2(x) = e^{2x} \cdot I_{\{sinx > -0.999\}} \end{cases}$$

in system (5). Clearly, (12) holds for $L = 1$. The system is thus stabilizable by virtue of Theorem 2. We remark that this system grows exponentially on the most part of the real line.

III. Adaptive Control Design and Closed-loop Identification

In order to achieve the stabilization, we employ the self-tuning regulator based on the least-squares algorithm (LS-STR). The standard LS estimates $\hat{\theta}_t$ for parameter θ can be recursively defined by

$$\begin{cases} \hat{\theta}_{t+1} = \hat{\theta}_t + \sigma_w^{-2} P_{t+1}\phi_t(y_{t+1} - u_t - \phi_t^T\hat{\theta}_t) \\ P_{t+1} = P_t - (\sigma_w^2 + \phi_t^T P_t \phi_t)^{-1} P_t\phi_t\phi_t^T P_t,\ P_0 = I_n, \\ \phi_t \triangleq \phi(y_t), \quad t \geq 0 \end{cases} \tag{13}$$

where initial $\hat{\theta}_0 = \theta_0$ is deterministic. In light of the "certainty equivalence principle", the controller is designed as follows:

$$u_t = -\hat{\theta}_t^T \phi_t, \quad t \geq 0. \tag{14}$$

We shall show in the next two sections that the LS-STR (13)–(14) is the desired stabilizing controller for both Theorems 1 and 2. Besides, during the control process, parameter θ can be identified simultaneously.

Theorem 3. *Under the conditions of Theorem 1, the LS estimator is strongly consistent in the closed-loop system (5), (13) and (14). More precisely, $\|\hat{\theta}_t - \theta\|^2 = O\left(\frac{\log t}{t}\right)$ almost surely as $t \to +\infty$.*

Theorem 4. *Under the conditions of Theorem 2, the LS estimator is strongly consistent in the closed-loop system (4), (13) and (14). More precisely, $\|\hat{\theta}_t - \theta\|^2 = O(t^{-\frac{1}{2}})$ almost surely as $t \to +\infty$.*

It is worth mentioning that the strong consistency of the LS estimates in the closed loop can be guaranteed by the stability of the system. To see this, let $\lambda_{\min}(t+1)$ be the minimal eigenvalue of P_{t+1}^{-1} defined in (13). Under Assumptions A1–A2, [16] implies

$$\left\{\lim_{t\to+\infty} \lambda_{\min}(t+1) = +\infty\right\} = \left\{\lim_{t\to+\infty} \hat{\theta}_t = \theta\right\}. \tag{15}$$

Furthermore, [4] and [8] point out that

$$\|\hat{\theta}_{t+1} - \theta\|^2 = O\left(\frac{\log\left(1 + \sum_{i=0}^t \|\phi(y_i)\|^2\right)}{\lambda_{\min}(t+1)}\right), \text{ a.s.}, \tag{16}$$

which offers a way to estimate the convergence rate of $\hat{\theta}_t$. Give (15) and (16), we first derive an important proposition on $\lambda_{\min}(t+1)$, whose proof is included in the appendix.

Proposition 1. *Under Assumptions A1–A3, there is a constant $M > 0$ such that in the closed-loop system (4), (13), (14),*

$$\liminf_{t\to+\infty} \frac{\lambda_{\min}(t+1)}{t} \geq M \liminf_{t\to+\infty} \frac{1}{t}\sum_{i=1}^t \frac{1}{\sigma_{i-1}} \quad a.s.,$$

where $\sigma_{i-1}^2 \triangleq Var(y_i|\mathcal{F}_{i-1}^y)$.

Remark 3. *General speaking, identification brings benifit to stabilization. However, as the system tends to be stable, the output signal will not change remarkably, which conversely reduces the excitation energy used for the identification. This dilemma is solved by Proposition 1. It establishes a tie between $\lambda_{\min}(t+1)$ and $\sum_{i=1}^t \sigma_{i-1}^{-1}$, revealing the stability of the self-tuning system (4), (13), (14) would not destroy the estimation here. In fact, when the conditional variances of the outputs are minor, a sufficiently good estimation performance can still be obtained to control σ_t^2 in the next step, so that a virtuous cycle is created.*

The lemma below is a consequence of Proposition 1, with the proof stated in the appendix.

Lemma 1. *Under Assumptions A1–A2, let the closed-loop system (4), (13), (14) satisfy*

$$\|\phi(x)\| = O(|x|^b) + O(1), \quad b > 0. \tag{17}$$

Then, there is a set D with $P(D) = 0$ such that

$$\begin{aligned}&\left\{\sum_{i=1}^t y_i^2 = O(t)\right\} \backslash D \\ \subseteq\ &\left\{\liminf_{t\to+\infty} \frac{\lambda_{\min}(t+1)}{t} > 0\right\} \cap \left\{\lim_{t\to+\infty} \frac{1}{t}\sum_{i=1}^t y_i^2 = \sigma_w^2\right\},\end{aligned} \tag{18}$$

and with probability 1,

$$\sum_{i=1}^t y_i^2 = O(t) \text{ implies } \|\hat{\theta}_{t+1} - \theta\|^2 = O\left(\frac{\log t}{t}\right). \tag{19}$$

Remark 4. *By virtue of (18) in Lemma 1,*

$$\sum_{i=1}^t y_i^2 = O(t) \text{ is equivalent to } \lim_{t\to+\infty} \frac{1}{t}\sum_{i=1}^t y_i^2 = \sigma_w^2 \quad a.s.,$$

which is referred to as optimality (see [4]). Meanwhile, (19) suggests that Theorem 3 is a direct result of Theorem 1.

For systems growing exponentially, we remark that the proof of Theorem 2 indicates (see (42))

$$\liminf_{t\to+\infty} \frac{\lambda_{\min}(t+1)}{t} > 0, \quad \text{a.s..} \tag{20}$$

So by $\sum_{i=0}^t \|\phi(y_i)\|^2 = O(t \cdot \exp\{\max_{0\leq i\leq t} |y_i|\})$ and (16),

$$\|\hat{\theta}_{t+1} - \theta\|^2 = O\left(\frac{\log t}{t}\right) + O\left(\frac{\sqrt{\sum_{i=1}^t y_i^2}}{t}\right) \quad \text{a.s..}$$

Consequently, it is straightforward that

Lemma 2. *Under Assumptions A1–A2, if (11) holds, then in the closed-loop system (4), (13), (14),*

$$\sum_{i=1}^t y_i^2 = O(t) \quad \text{implies} \quad \|\hat{\theta}_t - \theta\|^2 = O(t^{-\frac{1}{2}}), \quad a.s..$$

We close this section by emphasizing that the closed-loop identification can be achieved in the presence of the stability in system (4), (13), (14). In addition, if the growth rate of the system satisfies (17), the stability also implies the optimality.

IV. Proof of Theorem 1

For the closed-loop system (5), (13) and (14), one has

$$P_{t+1}^{-1} = I_n + \frac{1}{\sigma_w^2}\sum_{i=0}^t \phi_i\phi_i^T, \quad y_{t+1} = \tilde{\theta}_t f(y_t) + w_{t+1}, \tag{21}$$

where $\tilde{\theta}_t \triangleq \theta - \hat{\theta}_t$, $t \geq 0$. Since the LS algorithm (13) is exactly the standard Kalman filter in our case, it yields that $\hat{\theta}_t = E[\theta|\mathcal{F}_t^y]$ and $P_t = E[\tilde{\theta}_t\tilde{\theta}_t^T|\mathcal{F}_t^y]$. Hence, for each $t \geq 0$, y_{t+1} possesses a conditional Gaussian distribution given $\mathcal{F}_t^y$. The conditional mean and variance are

$$m_t = E[y_{t+1}|\mathcal{F}_t^y] = u_t + \hat{\theta}_t^T \phi_t = 0, \tag{22}$$

$$\sigma_t^2 = Var(y_{t+1}|\mathcal{F}_t^y) = \sigma_w^2 + \phi_t^T P_t \phi_t = \sigma_w^2 \cdot \frac{|P_{t+1}^{-1}|}{|P_t^{-1}|}. \tag{23}$$

So, we shall make efforts to prove $\sup_t \sigma_t < +\infty$ almost surely. To this end, we provide several relevant lemmas.

Lemma 3. *Under the conditions of Theorem 1, if $P(D) > 0$ with $D^c = \{|P_t^{-1}| < (1+\sigma_w^{-2})^t, i.o.\}$, then*

$$\sup_t \sigma_t < +\infty, \quad a.s. \quad on\ D. \tag{24}$$

Proof. Clearly, there is a sequence of random subscript t_k with $t_0 = 0$ such that we can define the following random matrices:

$$\begin{cases} Q_0^{-1} = I_n \\ Q_k^{-1} = Q_{k-1}^{-1} + \frac{1}{\sigma_w^2} \cdot \phi_{t_{k-1}} \phi_{t_{k-1}}^T, \quad k \geq 1 \end{cases}$$

with the property

$$\begin{cases} \phi_{t_k}^T Q_k \phi_{t_k} > \phi_{t_{k-1}}^T Q_{k-1} \phi_{t_{k-1}} \\ \phi_t^T Q_k \phi_t \leq \phi_{t_{k-1}}^T Q_{k-1} \phi_{t_{k-1}}, \qquad t_{k-1} < t < t_k \end{cases}, \quad k \geq 1.$$

If $\{t_k\}$ is a finite sequence, then there is an integer k' such that $\phi_t^T Q_{k'+1} \phi_t \leq \phi_{t_{k'}}^T Q_{k'} \phi_{t_{k'}}$, $\forall t > t_{k'}$. Consequently,

$$\begin{aligned} \sigma_t^2 &= \sigma_w^2 \cdot \frac{|P_{t+1}^{-1}|}{|P_t^{-1}|} = \sigma_w^2 + \phi_t^T P_t \phi_t \\ &\leq \sigma_w^2 + \phi_t^T Q_{k'+1} \phi_t \leq \sigma_w^2 + \phi_{t_{k'}}^T Q_{k'} \phi_{t_{k'}}, \quad \forall\ t > t_{k'}, \end{aligned}$$

which leads to $\sup_t \sigma_t < +\infty$.

Now, assume that there exists a set $D' \subset D$ with $P(D') > 0$ such that $\{t_k\}$ is infinite on D'. Clearly,

$$\begin{aligned} \frac{|Q_k^{-1}|}{|Q_{k-1}^{-1}|} &= \sigma_w^2 + \phi_{t_{k-1}}^T Q_{k-1} \phi_{t_{k-1}} \\ &< \sigma_w^2 + \phi_{t_k}^T Q_k \phi_{t_k} = \frac{|Q_{k+1}^{-1}|}{|Q_k^{-1}|}, \quad k \geq 1. \end{aligned} \tag{25}$$

Similarly to [10, Lemma 3.1], we can prove that for any $t \in (t_{k-1}, t_k]$, $\frac{|P_t^{-1}|}{|P_{t-1}^{-1}|} \leq \frac{|Q_k^{-1}|}{|Q_{k-1}^{-1}|}$. On the other hand, since

$$\begin{aligned} \sum_{t=1}^{+\infty} P(|y_t| > \sigma_{t-1} \log t | \mathcal{F}_{t-1}^y) &= \frac{1}{\sqrt{2\pi}} \sum_{t=1}^{+\infty} \int_{|x| \geq \log t} e^{-\frac{x^2}{2}}\, dx \\ &< +\infty, \end{aligned}$$

by *Borel-Cantelli-Levy* theorem, with probability 1, the events $\{|y_t| > \sigma_{t-1} \log t\}$ occur only finite many times for $t \geq 1$. That is, for all sufficiently large t,

$$|y_t|^2 \leq \sigma_{t-1}^2 \log^2 t = \sigma_w^2 \cdot \frac{|P_t^{-1}|}{|P_{t-1}^{-1}|} \log^2 t, \quad \text{a.s.}, \tag{26}$$

which infers that there exists a random number $\gamma > \sigma_w^{-2}$ such that for all $t \geq 0$,

$$\max\{1, |y_t|^2\} \leq \gamma \cdot \sigma_{t-1}^2 \log^2(t+3), \quad \text{a.s.}, \tag{27}$$

where $\sigma_{-1} \triangleq \sigma_w$ and $P_{-1} \triangleq I_n$.

Next, for $d \geq 1$, we define $\alpha_d(-1) \triangleq e_d$ and $\alpha_d(j) = \sigma_w^{-2} f(y_{t_j}) f_d(y_{t_j})$, $j \geq 0$, where e_d is the dth column of the identity matrix I_n. Then,

$$\begin{aligned} |Q_{k+1}^{-1}| &= \det\left(\sum_{j=-1}^{k} \alpha_1(j), \ldots, \sum_{j=-1}^{k} \alpha_n(j)\right) \\ &= \sum_{s_1, \ldots, s_n = -1}^{k} \det\left(\alpha_1(s_1), \ldots, \alpha_n(s_n)\right). \end{aligned} \tag{28}$$

If there exist two subscripts $d \neq d'$ such that $s_d = s_{d'} \neq -1$, we obtain

$$\det\left(\alpha_1(s_1), \ldots, \alpha_n(s_n)\right) = 0,$$

and hence

$$|Q_{k+1}^{-1}| = \sum_{(s_1, \ldots, s_n) \in \mathcal{W}(k)} \det\left(\alpha_1(s_1), \ldots, \alpha_n(s_n)\right). \tag{29}$$

Here, given integer $k \geq 1$ and positive integers $l_1 < \cdots < l_m$, $\mathcal{W}(k) \triangleq \{(l_1, \ldots, l_n) : l_i \in \{-1, 0, \ldots, k\}, i \in [1, n]; l_i \neq l_{i'} \text{ if } i \neq i',\ l_{i'} \neq -1\}$, $\mathcal{H}_k^{(l_1, \ldots, l_m)} \triangleq \{(i_1, \ldots, i_k) : i_j \in \{l_1, \ldots, l_m\}, 1 \leq j \leq k; i_r \neq i_s \text{ if } r \neq s\}$. Now, for any $(s_1, \ldots, s_n) \in \mathcal{W}(k)$, by (9) and (27),

$$\begin{aligned} & \det\left(\alpha_1(s_1), \ldots, \alpha_n(s_n)\right) \\ \leq & \sum_{(l_1, \ldots, l_n) \in \mathcal{H}_n^{(1, \ldots, n)}} \prod_{i \in [1,n], s_i \neq -1} \frac{1}{\sigma_w^2} \cdot |f_{l_i}(y_{t_{s_i}}) f_i(y_{t_{s_i}})| \\ \leq & \sum_{(l_1, \ldots, l_n) \in \mathcal{H}_n^{(1, \ldots, n)}} \prod_{i \in [1,n], s_i \neq -1} \frac{1}{\sigma_w^2} \cdot (L_1 + L_2 |y_{s_i}|^{b_{l_i}}) \\ & \cdot (L_1 + L_2 |y_{s_i}|^{b_i}) \\ \leq & (L_1 + L_2)^{2n} \cdot \sum_{(l_1, \ldots, l_n) \in \mathcal{H}_n^{(1, \ldots, n)}} \prod_{i \in [1,n], s_i \neq -1} \\ & \frac{1}{\sigma_w^2} \cdot \left(\gamma \cdot \log^2(t_{s_i} + 3) \cdot \sigma_w^2 \cdot \frac{|P_{t_{s_i}}^{-1}|}{|P_{t_{s_i}-1}^{-1}|}\right)^{\frac{b_{l_i} + b_i}{2}} \\ \leq & (L_1 + L_2)^{2n} \cdot \sum_{(l_1, \ldots, l_n) \in \mathcal{H}_n^{(1, \ldots, n)}} \prod_{i \in [1,n], s_i \neq -1} \\ & \frac{1}{\sigma_w^2} \cdot \left(\gamma \cdot \log^2(t_{s_i} + 3) \cdot \sigma_w^2 \cdot \frac{|Q_{s_i}^{-1}|}{|Q_{s_i-1}^{-1}|}\right)^{\frac{b_{l_i} + b_i}{2}} \\ \leq & (L_1 + L_2)^{2n} \cdot (1 + \sigma_w^{2b_1 - 2} + \sigma_w^{2b_n - 2})^n \cdot n! \\ & \cdot (\gamma \cdot \log^2(t_k + 3))^{(b_1 + \cdots + b_n)} \cdot \prod_{i=1}^{n} \left(\frac{|Q_{k+1-i}^{-1}|}{|Q_{k-i}^{-1}|}\right)^{b_i}, \end{aligned} \tag{30}$$

where $Q_{-1} \triangleq I_n$, and $L_1, L_2 > 1$ are two numbers satisfying $|f_j(x)| \leq L_1 + L_2 |x|^{b_j}$, $\forall x \in \mathbb{R}$, $j \in [1, n]$. By combining (29) and (30), we conclude

$$\begin{aligned} |Q_{k+1}^{-1}| \leq & (k+2)^n \cdot (L_1 + L_2)^{2n} \cdot (1 + \sigma_w^{2b_1 - 2} + \sigma_w^{2b_n - 2})^n \\ & \cdot n! \cdot (\gamma \cdot \log^2(t_k + 3))^{\sum_{i=1}^n b_i} \cdot \prod_{i=1}^{n} \left(\frac{|Q_{k+1-i}^{-1}|}{|Q_{k-i}^{-1}|}\right)^{b_i}. \end{aligned}$$

As a consequence, if $|Q_{k+1}^{-1}| > t_k^{\sqrt{\log t_k}}$ for all sufficiently large k, then there must exist a random number t'_ϵ for any given $\epsilon > 0$ such that for $k \geq t'_\epsilon$,

$$\begin{aligned} (1 - \epsilon) \log |Q_{k+1}^{-1}| &\leq \sum_{i=1}^{n} b_i (\log |Q_{k+1-i}^{-1}| - \log |Q_{k-i}^{-1}|) \\ &= b_1 \log |Q_k^{-1}| - \sum_{i=1}^{n-1} (b_i - b_{i+1}) \log |Q_{k-i}^{-1}| - b_n \log |Q_{k-n}^{-1}|. \end{aligned}$$

Define $z_k \triangleq \log |Q_{k+1}^{-1}| / \log |Q_k^{-1}|$ for $k \geq 1$ and set

$$z \triangleq \liminf_{k \to +\infty} z_k \geq 1.$$

Then, the above inequality reduces to

$$1-\epsilon+\sum_{i=1}^{n-1}(b_i-b_{i+1})\frac{1}{\prod_{j=0}^{i} z_{k-j}}+b_n\frac{1}{\prod_{j=0}^{n} z_{k-j}}\le b_1\frac{1}{z_k}.$$

Taking limit superior on both sides of the above inequality, we have $1-\epsilon+\sum_{i=1}^{n-1}(b_i-b_{i+1})\frac{1}{z^{i+1}}+b_n\frac{1}{z^n}\le b_1\frac{1}{z}$. Letting $\epsilon\to+\infty$ shows that $P(z)\le 0$ and $z>1$. This contradicts to the definition of $P(x)$. Hence, we immediately deduce that

$$|Q_{k+1}^{-1}|\le t_k^{\sqrt{\log t_k}} \quad \text{i.o.} \quad \text{a.s. on } D'. \tag{31}$$

Similarly to (28)–(30), when k is sufficiently large and satisfies $|Q_{k+1}^{-1}|\le t_k^{\sqrt{\log t_k}}$, for any $t\in(t_k+1,t_{k+1}+1]$,

$$\begin{aligned}
&|P_t^{-1}|\le (L_1+L_2)^{2n}\sum_{(s_1,\dots,s_n)\in\mathcal{W}(t-1)}\ \sum_{(l_1,\dots,l_n)\in\mathcal{H}_n^{(1,\dots,n)}}\\
&\prod_{i\in[1,n],s_i\neq-1}\frac{1}{\sigma_w^2}\left(\gamma\cdot\log^2(s_i+3)\cdot\sigma_w^2\cdot\frac{|P_{s_i}^{-1}|}{|P_{s_i-1}^{-1}|}\right)^{\frac{b_{l_i}+b_i}{2}}\\
\le\ & (L_1+L_2)^{2n}\cdot(1+\sigma_w^{2b_1-2}+\sigma_w^{2b_n-2})^n\\
&\cdot(\gamma\cdot\log^2(t+2))^{\sum_{i=1}^n b_i}(t+1)^n\cdot n!\cdot\left(\frac{|Q_{k+1}^{-1}|}{|Q_k^{-1}|}\right)^{\sum_{i=1}^n b_i}\\
\le\ & (L_1+L_2)^{2n}\cdot(1+\sigma_w^{2b_1-2}+\sigma_w^{2b_n-2})^n\\
&\cdot(\gamma\cdot\log^2(t+2))^{\sum_{i=1}^{m_1} b_i}(t+1)^n\cdot n!\cdot t_k^{\sqrt{\log t_k}\cdot\sum_{i=1}^n b_i}\\
<\ & (1+\sigma_w^{-2})^t.
\end{aligned}$$

This together with (31) leads to $|P_t^{-1}|<(1+\sigma_w^{-2})^t$ i.o. almost surely on D'. However, according to the assumption, events $\{|P_t^{-1}|<(1+\sigma_w^{-2})^t\}_{t\ge1}$ occur finitely on D, which arises a contradiction. Hence $\{t_k\}$ is finite on D almost surely, and (24) follows. ■

Lemma 4. *Under Assumptions A1–A3, assume (17) holds and there is a set D with $P(D)>0$ such that*

$$|P_t^{-1}|<(1+\sigma_w^{-2})^t \quad \text{i.o.} \quad \text{a.s. on } D. \tag{32}$$

Then,

$$\sup_t \sigma_t<+\infty, \quad \text{a.s.} \quad \text{on } D. \tag{33}$$

Proof. Denote

$$F\triangleq\{t\ge0:|P_t^{-1}|\ge(1+\sigma_w^{-2})^t,|P_{t+1}^{-1}|<(1+\sigma_w^{-2})^{t+1}\}.$$

Firstly, by (32), for all sufficiently large t, $|P_t^{-1}|<(1+\sigma_w^{-2})^t$ a.s. on $\{|F|<+\infty\}\cap D$. Let $\varepsilon=\frac{M}{3}\cdot\min\{(1+\sigma_w)^{-1},\sigma_w\cdot(1+\sigma_w^2)^{-1}\}$. In view of Proposition 1, there is a random integer t_1 such that for all $t>t_1$,

$$\begin{aligned}
\lambda_{\min}(t+1)\ &\ge\ M\sum_{i=1}^{t}\frac{1}{\sigma_{i-1}}-\varepsilon t\\
&\ge\ M\cdot t\cdot\left(\frac{1}{\prod_{i=1}^t\sigma_{i-1}}\right)^{\frac{1}{t}}-\varepsilon t\\
&=\ M\sigma_w^{-1}\cdot t\cdot\left(\frac{1}{|P_t^{-1}|}\right)^{\frac{1}{2t}}-\varepsilon t\\
&>\ M\sigma_w^{-1}\cdot t\cdot\left(\frac{1}{(1+\sigma_w^{-2})^t}\right)^{\frac{1}{2t}}-\varepsilon t\\
&>\ \varepsilon t, \quad \text{a.s.} \quad \text{on } \{|F|<+\infty\}\cap D.
\end{aligned}$$

In addition, for some integer $N>0$,

$$\varepsilon t>K_1+K_2\log^{2b}t\cdot(1+\sigma_w)^{2b}, \quad \forall t\ge N, \tag{34}$$

where constants K_1,K_2 satisfy $\|\phi(x)\|\le K_1+K_2|x|^b$ for $x\in\mathbb{R}$. Clearly, there exists a random integer $t_2>t_1+N$ such that $\sigma_{t_2}<1+\sigma_w$ and (26) holds for all $t\ge t_2$. Next, we show $\sigma_t<1+\sigma_w$ for all $t\ge t_2$ by induction on set $\{|F|<+\infty\}\cap D$.

Suppose that $\sigma_k<1+\sigma_w$ for some $k\ge t_2$, then (34) gives

$$\begin{aligned}
\sigma_{k+1}^2\ &=\ \sigma_w^2\cdot\frac{|P_{k+2}^{-1}|}{|P_{k+1}^{-1}|}=\sigma_w^2+\phi_{k+1}^TP_{k+1}\phi_{k+1}\\
&\le\ \sigma_w^2+\frac{\|\phi(y_{k+1})\|^2}{\lambda_{\min}(k+1)}\\
&\le\ \sigma_w^2+\frac{K_1+K_2|y_{k+1}|^{2b}}{\varepsilon k}\\
&\le\ \sigma_w^2+\frac{K_1+K_2\log^{2b}k\cdot\sigma_k^{2b}}{\varepsilon k}\\
&<\ (1+\sigma_w)^2, \quad \text{a.s.} \quad \text{on } \{|F|<+\infty\}\cap D.
\end{aligned}$$

By induction, we know $\sigma_t<1+\sigma_w$ for all $t\ge t_2$. This means $\sup_t\sigma_t<+\infty$ almost surely on $\{|F|<+\infty\}\cap D$.

So the remainder of the argument is focused on set $\{|F|=+\infty\}\cap D$. By Proposition 1 again, there exists a random integer t_1' such that

$$\lambda_{\min}(t+1)\ \ge\ M\sum_{i=1}^{t}\frac{1}{\sigma_{i-1}}-\varepsilon t, \quad t>t_1', \quad \text{a.s.}. \tag{35}$$

On $\{|F|=+\infty\}\cap D$, select an random integer $k'\in F$ such that $k'>t_1'+N+2$. we now prove that for all $t\ge k'$, on $\{|F|=+\infty\}\cap D$,

$$\lambda_{\min}(t+1)>\varepsilon t \quad \text{and} \quad \sigma_t<1+\sigma_w \quad \text{a.s.}. \tag{36}$$

As a matter of fact, for $k'\in F$, we have $\sigma_{k'}^2=\sigma_w^2\cdot\frac{|P_{k'+1}^{-1}|}{|P_{k'}^{-1}|}<\sigma_w^2\cdot(1+\sigma_w^{-2})^{k'+1-k'}<(1+\sigma_w)^2$. Consequently, (35) yields

$$\begin{aligned}
\lambda_{\min}(k'+1)\ &\ge\ M\sum_{i=1}^{k'}\frac{1}{\sigma_{i-1}}-\varepsilon k'\\
&\ge\ M\cdot\sigma_w^{-1}\cdot k'\cdot\left(\frac{1}{|P_{k'}^{-1}|}\right)^{\frac{1}{2k'}}-\varepsilon k'\\
&\ge\ M\cdot\sigma_w^{-1}\cdot k'\cdot\left(\frac{1}{|P_{k'+1}^{-1}|}\right)^{\frac{1}{2k'}}-\varepsilon k'\\
&>\ M\cdot\sigma_w^{-1}\cdot k'\cdot\left(\frac{1}{(1+\sigma_w^{-2})^{k'+1}}\right)^{\frac{1}{2k'}}-\varepsilon k'\\
&\ge\ \varepsilon k'.
\end{aligned}$$

Assume that for some $j \geq k'$, (36) holds for all $t \in [k', j]$. Then it follows that

$$\begin{aligned}\sigma_{j+1}^2 &= \sigma_w^2 \cdot \frac{|P_{j+2}^{-1}|}{|P_{j+1}^{-1}|} \leq \sigma_w^2 + \frac{\|\phi(y_{j+1})\|^2}{\lambda_{\min}(j+1)} \\ &\leq \sigma_w^2 + \frac{K_1 + K_2|y_{j+1}|^{2b}}{\varepsilon j} \\ &\leq \sigma_w^2 + \frac{K_1 + K_2 \log^{2b} j \cdot \sigma_j^{2b}}{\varepsilon j} \\ &< (1+\sigma_w)^2 \quad \text{a.s.} \quad \text{on } \{|F| = +\infty\} \cap D.\end{aligned}$$

As a result,

$$\begin{aligned}&\lambda_{\min}(j+2) \\ \geq\ & M \sum_{i=1}^{j+1} \frac{1}{\sigma_{i-1}} - \varepsilon(j+1) \\ =\ & M \sum_{i=1}^{k'} \frac{1}{\sigma_{i-1}} + M \sum_{k' < i \leq j+1} \frac{1}{\sigma_{i-1}} - \varepsilon(j+1) \\ >\ & M \cdot \sigma_w^{-1} \cdot k' \cdot \left(\frac{1}{(1+\sigma_w^{-2})^{k'+1}} \right)^{\frac{1}{2k'}} \\ & + \frac{M}{1+\sigma_w}(j-k+1) - \varepsilon(j+1) \\ \geq\ & \varepsilon(j+1) \quad \text{a.s.} \quad \text{on } \{|F| = +\infty\} \cap D.\end{aligned}$$

Hence (36) is true for $t = j+1$, and the induction is completed. So $\sup_t \sigma_t < +\infty$ almost surely on $\{|F| = +\infty\} \cap D$.

To sum up, (33) holds as desired. ∎

Lemma 5. *Under Assumptions A1–A2, if (11) holds and* $\sup_t \sigma_t < +\infty$ *a.s., then* $\sup_{t \geq 1} \frac{1}{t} \sum_{i=1}^t y_i^2 < +\infty$, *a.s..*

Proof. Recall from [4, Lemma 3.1] that

$$\sum_{i=0}^t \alpha_i = O\left(\log\left(1 + \sum_{i=0}^t \phi_i^T \phi_i \right) \right), \text{ a.s.,}$$

where $\alpha_i \triangleq (1 + \phi_i^T P_i \phi_i)^{-1} (\tilde{\theta}_i^T \phi_i)^2$, $i \geq 0$. Therefore,

$$\begin{aligned}&\frac{1}{2} \sum_{i=1}^t y_i^2 - \sum_{i=0}^t w_{i+1}^2 \\ \leq\ & \sum_{i=0}^t (y_{i+1} - w_{i+1})^2 = \sum_{i=0}^t \alpha_i \frac{|P_{i+1}^{-1}|}{|P_i^{-1}|} = O\left(\sum_{i=0}^t \alpha_i \right) \\ =\ & O\left(\log\left(1 + \sum_{i=0}^t \phi_i^T \phi_i \right) \right) \\ \leq\ & O\left(\log\left(1 + k_1^2 \sum_{i=0}^t e^{2k_2|y_i|} \right) \right) \\ \leq\ & O\left(\log\left(1 + k_1^2 e^{2k_2|y_0|} + k_1^2 t \cdot e^{2k_2 \cdot \max_{1 \leq i \leq t} |y_i|} \right) \right) \\ =\ & O(1) + O(\log t) + O\left(\left(\sum_{i=1}^t y_i^2 \right)^{\frac{1}{2}} \right), \quad \text{a.s..}\end{aligned}$$

Observe that $\sum_{i=0}^t w_{i+1}^2 = O(t)$ as $t \to +\infty$, then

$$\frac{1}{2} \sum_{i=1}^t y_i^2 \leq O(t) + O\left(\left(\sum_{i=1}^t y_i^2 \right)^{\frac{1}{2}} \right),$$

which implies $\sum_{i=1}^t y_i^2 = O(t)$ almost surely. ∎

Theorem 1 is now straightforward.

Proof of Theorem 1. Taking account of Lemmas 3 and 4, we have $\sup_t \sigma_t < +\infty$ a.s., which leads to Theorem 1 directly by Lemma 5. ∎

V. Proof of Theorem 2

The proof is based on two lemmas below.

Lemma 6. *If* $\liminf_{t \to +\infty} \frac{\lambda_{\min}(t+1)}{t} > 0$ *and* $\sup_t \sigma_t = +\infty$ *hold almost surely on a set* D *with* $P(D) > 0$, *then* $\lim_{t \to +\infty} \sigma_t = +\infty$ *a.s. on* D.

Proof. Define $\Omega_{k+1} \triangleq \{\sigma_k^2 \leqslant z, \sigma_{k+1}^2 \geqslant z\}, k \geq 0$ for some given $z > \sigma_w^2$. Therefore,

$$\begin{aligned}&P(\Omega_{k+1} | \mathcal{F}_k^y) \\ =\ & P\left(\sigma_w^2 \cdot \frac{|P_{k+2}^{-1}|}{|P_{k+1}^{-1}|} \geqslant z, \sigma_k^2 \leqslant z \middle| \mathcal{F}_k^y \right) \\ \leq\ & P\left(\sigma_w^2 + \frac{\|\phi(y_{k+1})\|^2}{\lambda_{\min}(k+1)} \geqslant z, \sigma_k^2 \leqslant z \middle| \mathcal{F}_k^y \right) \\ \leq\ & E\left\{ I_{\left\{|y_{k+1}| \geqslant \frac{1}{k_2} \log\left(\frac{(z-\sigma_w^2)\lambda_{\min}(k+1)}{k_1} \right)\right\}} \cdot I_{\{\sigma_k^2 \leqslant z\}} \middle| \mathcal{F}_k^y \right\} \\ =\ & I_{\{\sigma_k^2 \leqslant z\}} \cdot E\left\{ I_{\left\{|y_{k+1}| \geqslant \frac{1}{k_2} \log\left(\frac{(z-\sigma_w^2)\lambda_{\min}(k+1)}{k_1} \right)\right\}} \middle| \mathcal{F}_k^y \right\} \\ =\ & I_{\{\sigma_k^2 \leqslant z\}} \cdot \frac{1}{\sqrt{2\pi}} \int_{|x \cdot \sigma_k| \geqslant \frac{1}{k_2} \log\left(\frac{(z-\sigma_w^2)\lambda_{\min}(k+1)}{k_1} \right)} e^{-\frac{x^2}{2}}\, dx \\ \leq\ & \frac{1}{\sqrt{2\pi}} \int_{|x \cdot \sqrt{z}| \geqslant \frac{1}{k_2} \log\left(\frac{(z-\sigma_w^2)\lambda_{\min}(k+1)}{k_1} \right)} e^{-\frac{x^2}{2}}\, dx \\ =\ & \frac{1}{\sqrt{2\pi}} \int_{|x| \geq X_k} e^{-\frac{x^2}{2}}\, dx, \end{aligned} \tag{37}$$

where

$$X_k \triangleq \frac{1}{\sqrt{z}} \frac{1}{k_2} \log\left(\frac{(z - \sigma_w^2)\lambda_{\min}(k+1)}{k_1} \right).$$

Since $\liminf_{k \to +\infty} \frac{\lambda_{\min}(k+1)}{k} > 0$ implies

$$\liminf_{k \to +\infty} \frac{X_k}{\log k} > 0,$$

(37) yields

$$\sum_{k=1}^{+\infty} P(\Omega_{k+1} | \mathcal{F}_k^y) \leq \sum_{k=1}^{+\infty} \frac{1}{\sqrt{2\pi}} \int_{|x| \geq X_k} e^{-\frac{x^2}{2}}\, dx < +\infty.$$

Taking account of *Borel-Cantelli-Levy* theorem, $\{\Omega_k\}$ occur only finite times almost surely. The rest of the proof is as the same of that for [12, Lemma 3.5]. ∎

Lemma 7. *Let* $\{A_k\}_{k \geq 1}$ *be a sequence of events that* $A_k \triangleq \{y_k \in S_L\}$. *Then, there exists a constant* $c > 0$, *which only depends on* $f_1, \ldots, f_n$, *such that for all sufficiently large* t,

$$\sum_{k=1}^t I_{A_k} \geq ct, \quad \textit{a.s..} \tag{38}$$

Proof. Recall that (12) means there are two numbers $q_1, q_2 > 0$ such that

$$\frac{\ell(S_L \cap [-l,l])}{l} > q_2 \quad \text{for} \quad l \geq q_1. \tag{39}$$

Since y_{i+1} is conditional Gaussian with the conditional mean $m_i = 0$ and variance σ_i^2 from (22) and (23), we compute

$$\begin{aligned} P\left(A_{i+1}|\mathcal{F}_i^y\right) &= \frac{1}{\sqrt{2\pi}} \int_{|x\sigma_i|\in S_L} e^{-\frac{x^2}{2}}\, dx \\ &\geq \frac{1}{\sqrt{2\pi}} \int_{|x\sigma_i|\in S_L, |x|\leqslant \frac{q_1}{\sigma_w}} e^{-\frac{x^2}{2}}\, dx \\ &\geq \ell\left(x : |x\sigma_i| \in S_L, |x| \leqslant \frac{q_1}{\sigma_w}\right) \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{q_1^2}{2\sigma_w^2}} \\ &= \frac{\ell(S_L \cap [-q_1\sigma_i\sigma_w^{-1}, q_1\sigma_i\sigma_w^{-1}])}{\sigma_i} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{q_1^2}{2\sigma_w^2}}. \end{aligned}$$

Owing to (39) and $\sigma_i \geq \sigma_w$, we immediately deduce

$$P(A_{i+1}|\mathcal{F}_i^y) > \frac{q_1 q_2}{\sqrt{2\pi}\sigma_w} e^{-\frac{q_1^2}{2\sigma_w^2}}.$$

Further, by applying the strong law of large numbers for the martingale differences, we have

$$\frac{\sum_{k=1}^t (I_{A_k} - P(A_k|\mathcal{F}_{k-1}^y))}{t} = o(1), \quad \text{a.s.}.$$

Then for all sufficiently large t,

$$\begin{aligned} \sum_{k=1}^t I_{A_k} &\geq \sum_{k=1}^t P(A_k|\mathcal{F}_{k-1}^y) - \frac{q_1 q_2}{2\sqrt{2\pi}\sigma_w} e^{-\frac{q_1^2}{2\sigma_w^2}} t \\ &\geq \frac{q_1 q_2}{2\sqrt{2\pi}\sigma_w} e^{-\frac{q_1^2}{2\sigma_w^2}} t. \end{aligned}$$

So Lemma 7 follows by letting $c = \frac{q_1 q_2}{2\sqrt{2\pi}\sigma_w} e^{-\frac{q_1^2}{2\sigma_w^2}}$. ∎

Proof of Theorem 2. Same as previous, we are going to show

$$\sup_t \sigma_t < +\infty, \quad \text{a.s.}. \tag{40}$$

Assume there is a set D with $P(D) > 0$ such that $\sup_t \sigma_t = +\infty$ on D. Note that $y_k \in S_L$ infers $\sigma_k^2 = \frac{|P_{k+1}^{-1}|}{|P_k^{-1}|} \leq \sigma_w^2 + \frac{\|\phi(y_k)\|^2}{\lambda_{\min}(k)} \leq \sigma_w^2 + L^2$, then for any $t \geq 1$,

$$\sum_{k=1}^t I_{\{\sigma_k \leq \sqrt{\sigma_w^2 + L^2}\}} \geq \sum_{k=1}^t I_{A_k}, \tag{41}$$

where $I_{A_k}, k \in [1,t]$ are defined in Lemma 7. Take $\varepsilon = \frac{1}{4}(\sigma_w^2 + L^2)^{-\frac{1}{2}} Mc$. By Proposition 1, Lemma 7 and (41), for all sufficiently large t, we have

$$\begin{aligned} \lambda_{\min}(t+2) &\geq M \sum_{i=1}^{t+1} \frac{1}{\sigma_{i-1}} - \varepsilon(t+1) \\ &\geq M \cdot \frac{ct}{\sqrt{\sigma_w^2 + L^2}} - \frac{1}{4}(\sigma_w^2 + L^2)^{-\frac{1}{2}} Mc(t+1) \\ &> \frac{Mct}{2\sqrt{\sigma_w^2 + L^2}}, \quad \text{a.s.}. \end{aligned} \tag{42}$$

This means $\liminf_{t\to+\infty} \frac{\lambda_{\min}(t+1)}{t} > 0$ almost surely. So, Lemma 6 gives

$$\lim_{t\to+\infty} \sigma_t = +\infty, \quad \text{a.s.} \quad \text{on } D.$$

According to (41), it turns out that

$$\begin{aligned} \limsup_{t\to+\infty} \frac{\sum_{k=1}^t I_{A_k}}{t} &\leq \limsup_{t\to+\infty} \frac{\sum_{k=1}^t I_{\{\sigma_k \leq \sqrt{\sigma_w^2 + L^2}\}}}{t} \\ &= 0, \quad \text{a.s.} \quad \text{on } D, \end{aligned}$$

which contradicts to (38). We conclude (40) and thus Lemma 5 applies. ∎

VI. Simulations

This section provides the numerical simulations for Examples 1–2 to illustrate the main results. Let the output sequence $\{y_t; t = 0, \ldots, N\}, N = 10000$ is generated by model (5), where $\theta \sim N(\theta_0, I_2)$ with $\theta_0 = (1,2)^\tau$, $\{w_t; t = 1, \ldots, N\}$ are i.i.d with distribution $N(0,1)$ and $y_0 \sim N(0,1)$. Besdies, $\theta, \{w_t\}, y_0$ are set independent. Select the system functions as those in Examples 1 and 2 respectively.

Now, given algorithm (13)–(14), we calculate the average mean squared outputs $\frac{1}{t}\sum_{i=1}^t y_i^2$, outputs y_t and squared estimation errors $\|\hat{\theta}_t - \theta\|^2$. Figures 1–2 show the results of Examples 1–2, respectively. We notice several points.
(i) In both Examples 1 and 2, $\frac{1}{t}\sum_{i=1}^t y_i^2$ converge to $\sigma_w^2 = 1$, which is consistent with our theorems. Although the exponential growth of the system in Example 2 causes the average outputs quite large in the initial stage, the closed-loop system still finally tends to be stable over time.
(ii) Owing to the exponential growth of the system in Example 2, the error curve $\|\hat{\theta}_t - \theta\|^2$ fluctuates wildly. Here it tends to zero fast, yet its image is not as regular as that in Example 1.

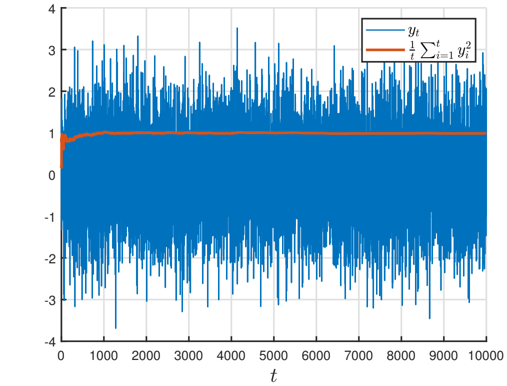
VII. Concluding remarks

This paper derives two stabilizability theorems for a basic class of discrete-time nonlinear uncertain systems, by establishing a fundamental inequality on the minimum eigenvalue of the information matrix. As shown both theoretically and numerically, the concerned self-tuning system has a fast convergence rate, in spite of the high nonlinearity. We point out that the input gain is fully known here. When the input gain is unknown, our technique analysis requires essential improvement. In this regard, [2] and [4] solved it for ARMA models and [9] for nonlinear systems with two unknown parameters. But for more general settings, there still remains a lot of problems which need further research.

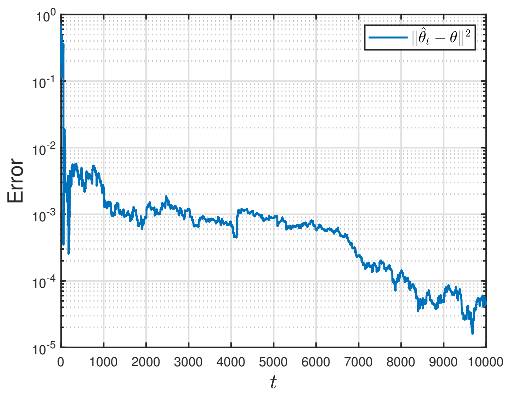
Appendix

To show Lemma 1, we need a simple fact.

Lemma 8. *Assume the conditions of Lemma 1 hold, then* $\sup_{t\geq 1} \frac{1}{t}\sum_{i=1}^t y_i^2 < +\infty$ *equals to* $\sup_t \sigma_t < +\infty$ *a.s..*

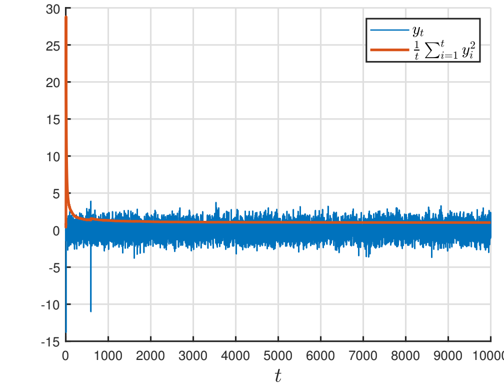


(a) The outputs of the closed-loop system

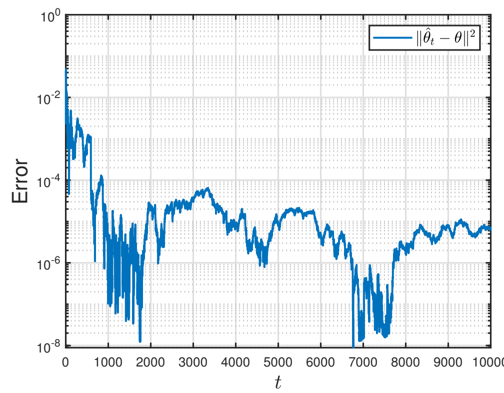


(b) The squared estimation error

Fig. 1. Simulation results on the stability and estimation in Example 1.



(a) The outputs of the closed-loop system



(b) The squared estimation error

Fig. 2. Simulation results on the stability and estimation in Example 2.

Proof. Denote $W \triangleq \left\{\sup_{t\geq 1}\frac{1}{t}\sum_{i=1}^t y_i^2 < +\infty\right\}$. Then, for all sufficiently large t, *Hadamard* inequality yields

$$\begin{aligned}
|P_t^{-1}| &\leq \prod_{j=1}^{n}\left(1+\frac{1}{\sigma_k^2}\sum_{i=0}^{t-1} f_j^2(y_i)\right)\\
&= \prod_{j=1}^{n}\left(1+O(t)+O\left(\sum_{i=0}^{t-1} y_i^{2b}\right)\right)\\
&\leq \prod_{j=1}^{n}\left(1+O(t)+O\left(\left(\sum_{i=0}^{t-1} y_i^{2}\right)^b\right)\right)\\
&= O\left(t^{(b+1)n}\right), \quad \text{on } W. \qquad (43)
\end{aligned}$$

So, applying Lemma 4 infers $\sup_t \sigma_t < +\infty$ a.s. on W. The lemma is true as the converse part is verified by Lemma 5. ■

Proof of Lemma 1. Let W be defined in Lemma 8 and $\varepsilon = \frac{M}{2\sigma_w}$. By Proposition 1,

$$\begin{aligned}
\lambda_{\min}(t+1) &\geq M\sum_{i=1}^{t}\frac{1}{\sigma_{i-1}} - \frac{M}{2\sigma_w}t\\
&\geq M\sigma_w^{-1}t\left(\frac{1}{|P_t^{-1}|}\right)^{\frac{1}{2t}} - \frac{M}{2\sigma_w}t, \quad \text{a.s.} \quad \text{on } W, \qquad (44)
\end{aligned}$$

where t is sufficiently large. Combining (43) and (44) gives

$$\liminf_{t\to+\infty}\frac{\lambda_{\min}(t+1)}{t} > 0, \quad \text{a.s.} \quad \text{on } W. \qquad (45)$$

Define a martingale difference sequence $Z_i = \frac{1}{i}(y_i^2 - E(y_i^2|\mathcal{F}_{i-1}^y)) = \frac{1}{i}(y_i^2 - \sigma_{i-1}^2)$, $i \geq 1$. In view of Lemma 8, $\sum_{i=1}^{+\infty} E(Z_i^2|\mathcal{F}_{i-1}^y) = \sum_{i=1}^{+\infty}\frac{2\sigma_{i-1}^4}{i^2} < +\infty$, a.s. on W. By [2, Theorem 2.7], we deduce that $\sum_{i=1}^{+\infty} Z_i$ converges almost surely on W. This, together with *Kronecker Lemma*, leads to

$$\sum_{i=1}^{t}(y_i^2 - \sigma_{i-1}^2) = o(t), \quad \text{a.s.} \quad \text{on } W. \qquad (46)$$

By (17), (26), (45) and Lemma 8, for sufficiently large t,

$$\begin{aligned}
\sigma_{t+1}^2 &\leq \sigma_w^2 + \frac{\|\phi(y_{t+1})\|^2}{\lambda_{\min}(t+1)} = \sigma_w^2 + O\left(\frac{y_{t+1}^{2b}}{t}\right)\\
&= \sigma_w^2 + O\left(\frac{(\sigma_t\log(t+1))^{2b}}{t}\right)\\
&= \sigma_w^2 + o(1), \quad \text{a.s.} \quad \text{on } W.
\end{aligned}$$

Recall that $\sigma_t \geq \sigma_w$, therefore

$$\lim_{t\to\infty}\sigma_t = \sigma_w, \quad \text{a.s.} \quad \text{on } W. \qquad (47)$$

According to (46) and (47), we derive $\lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^t y_i^2 = \lim_{t\to\infty}\frac{1}{t}\sum_{i=1}^t \sigma_{i-1}^2 = \sigma_w^2$, a.s. on W, which together with (45) gives Lemma 1. ■

Borrowing the idea of [10], the proof of Proposition 1 will be completed in the following three subsections.
Section A: Observe that

$$\lambda_{\min}(t+1) = \min_{\|x\|=1} x^T\left(I_n + \sum_{i=1}^{t}\phi_i\phi_i^T\right)x = 1 + \min_{\|x\|=1}\sum_{i=1}^{t}(\phi_i^T x)^2,$$

so for any unit vector $x \in \mathbb{R}^n$, we shall construct a bounded set $U_x \subset \mathbb{R}^n$ such that $\inf_{y\in U_x}|\phi^T(y)x| \geq \delta$ for some $\delta > 0$.
Section B: We shall analyze the properties of U_x and derive a key technique result for our problem in Lemma 19.
Section C: This section is intended to prove Proposition 1 by estimating the frequency of $\{y_t\}_{t\geq 1}$ falling into U_x.

A. *Construction of* U_x

The important set U_x is constructed from a finite family of disjoint open intervals $\{S^j(q)\}$ defined below.

1) Open Intervals $S^j(q)$: We claim that there exists a finite family of disjoint open intervals $\{S^j(q)\}_{j=1}^p$ for some $q \in \mathbb{N}^+$ fulfilling:
(i) $\phi \in C^n$ in $\bigcup_{j=1}^p S^j(q)$;
(ii) $\bigcup_{j=1}^p S^j(q)$ has no points in Z_s^2 defined later in (59);
(iii) For every unit vector $x \in \mathbb{R}^n$,

$$\ell\left(\left\{y \in \bigcup_{j=1}^p S^j(q) : |\phi^T(y)x| > 0\right\}\right) > 0. \tag{48}$$

We preface the proof with several auxiliary lemmas.

Lemma 9. *Let $\{U_j\}_{j\geq 1}$ be a series of open sets in E satisfying*

$$U_1 \subset U_2 \cdots \subset U_j \subset \cdots \quad \textit{and} \quad \lim_{j\to+\infty} U_j = U, \tag{49}$$

where U is a non-empty open set that $\ell(\{y \in U : |\phi^T(y)x| > 0\}) > 0$, $\forall x \in \mathbb{R}^n, \|x\| = 1$. Then, there is an integer j such that $\ell(\{y \in U_j : |\phi^T(y)x| > 0\}) > 0$, $\forall x \in \mathbb{R}^n, \|x\| = 1$.

Proof. If the assertion is not true, then by the continuity of ϕ, for each $j \geq 1$, there is an $x^j \in \mathbb{R}^n$ with $\|x^j\| = 1$ such that

$$\phi^T(y)x^j = 0, \quad \forall y \in U_j. \tag{50}$$

Then there is a subsequence $\{x^{n_i}\}_{i\geq 1}$ of $\{x^j\}_{j\geq 1}$ satisfying

$$\lim_{i\to+\infty} x^{n_i} = x^\infty \tag{51}$$

with $\|x^\infty\| = 1$. On the other hand, $\ell(\{y \in U : |\phi^T(y)x^\infty| > 0\}) > 0$, so there is a $y^* \in U$ such that

$$|\phi^T(y^*)x^\infty| > 0. \tag{52}$$

By (49), there is an integer $h' \geq 1$ such that $y^* \in U_j$ for all $j \geq h'$, and hence (50)–(52) yield $0 < |\phi^T(y^*)x^\infty| = \lim_{i\to+\infty}|\phi^T(y^*)x^{n_i}| = 0$, which leads to a contradiction. ■

Remark 5. *Since every open $E \subset \mathbb{R}$, is a countable union of disjoint open intervals, Lemma 9 implies that there is an open set $E' \subset E$ such that E' consists of a finite number of disjoint open intervals and $\ell(\{y \in E' : |\phi^T(y)x| > 0\}) > 0$, $\forall x \in \mathbb{R}^n$, $\|x\| = 1$. So, without loss of generality, assume E in the sequel is a finite union of disjoint open intervals.*

Now, we introduce a series of operators. Denote D as the differential operator, then for any sufficiently smooth functions $\{g_l\}_{l\geq 1}$, recursively define

$$\begin{cases} \Lambda_1(g_1) \triangleq g_1 \\ \Lambda_{l+1}(g_1, \ldots, g_{l+1}) \triangleq \Lambda_l\left(\frac{Dg_1}{Dg_{l+1}}, \ldots, \frac{Dg_l}{Dg_{l+1}}\right), l \geq 1 \end{cases}. \tag{53}$$

These operators $\{\Lambda_l\}_{l\geq 1}$ have the following property:

Lemma 10. *Let functions $\{g_i\}_{i=1}^{l+1}$, $l \in \mathbb{N}^+$ be sufficiently smooth, then*

$$\Lambda_{l+1}(g_1, \ldots, g_{l+1}) = \frac{D(\Lambda_l(g_1, g_3, \ldots, g_{l+1}))}{D(\Lambda_l(g_2, g_3, \ldots, g_{l+1}))}. \tag{54}$$

Proof. We use the induction method to show this lemma. By the definition of Λ_2, it is easy to check $\Lambda_2(g_1, g_2) = \Lambda_1\left(\dfrac{Dg_1}{Dg_2}\right) = \frac{Dg_1}{Dg_2} = \frac{D(\Lambda_1(g_1))}{D(\Lambda_1(g_2))}$. Let $k \geq 2$. Suppose (54) holds for any functions $\{g_i\}_{i=1}^{l+1}$, $l = k-1$, then

$$\Lambda_k\left(\frac{Dg_1}{Dg_{k+1}}, \ldots, \frac{Dg_k}{Dg_{k+1}}\right) = \frac{D\left(\Lambda_{k-1}\left(\frac{Dg_1}{Dg_{k+1}}, \frac{Dg_3}{Dg_{k+1}}, \ldots, \frac{Dg_k}{Dg_{k+1}}\right)\right)}{D\left(\Lambda_{k-1}\left(\frac{Dg_2}{Dg_{k+1}}, \frac{Dg_3}{Dg_{k+1}}, \ldots, \frac{Dg_k}{Dg_{k+1}}\right)\right)},$$

and hence by (53),

$$\begin{aligned}\Lambda_{k+1}(g_1, g_2, \ldots, g_{k+1}) &= \Lambda_k\left(\frac{Dg_1}{Dg_{k+1}}, \ldots, \frac{Dg_k}{Dg_{k+1}}\right) \\ &= \frac{D\left(\Lambda_{k-1}\left(\frac{Dg_1}{Dg_{k+1}}, \frac{Dg_3}{Dg_{k+1}}, \ldots, \frac{Dg_k}{Dg_{k+1}}\right)\right)}{D\left(\Lambda_{k-1}\left(\frac{Dg_2}{Dg_{k+1}}, \frac{Dg_3}{Dg_{k+1}}, \ldots, \frac{Dg_k}{Dg_{k+1}}\right)\right)} \\ &= \frac{D\left(\Lambda_k\left(g_1, g_3, \ldots, g_{k+1}\right)\right)}{D\left(\Lambda_k\left(g_2, g_3, \ldots, g_{k+1}\right)\right)},\end{aligned}$$

which completes the induction. ■

Before proceeding to the next lemma, we define some notations. Let $l_1 < \cdots < l_s$ be s positive integers. For each $k \in [1, s]$, recall $\mathcal{H}_k^{(l_1,\ldots,l_s)}$ as the k-permutations of $\{l_1, \ldots, l_s\}$. Now, for each $(i_1, \ldots, i_k) \in \mathcal{H}_k^{(1,\ldots,n)}$, $k \in [1, n]$, define

$$\Gamma_{(i_1,\ldots,i_k)} \triangleq \Lambda_k\left(f_{i_1}, \ldots, f_{i_k}\right), \quad \bar{\Gamma}_s \triangleq D\Gamma_s, \tag{55}$$

and for any $s \in \mathcal{H} \triangleq \bigcup_{k=1}^n \mathcal{H}_k^{(1,\ldots,n)}$, define

$$W_s \triangleq \left\{y : \bar{\Gamma}_s(y) \text{ is well-defined}\right\}.$$

Given function g, denote $A(g) \triangleq \{x : g(x) = 0\}$. In addition, for any two sets $\mathcal{X}_1, \mathcal{X}_2 \subset \mathbb{R}$, we say that $\mathcal{X}_1$ is *locally dense* in $\mathcal{X}_2$, if $\mathcal{X}_1$ is not nowhere dense in $\mathcal{X}_2$. That is, there exists a nonempty open interval $\mathcal{X}_3 \subset \mathcal{X}_2$ such that $\mathcal{X}_3 \subset \overline{\mathcal{X}_1}$. With the above pre-definitions, we assert

Lemma 11. *Let integers $k \in [2, n]$ and array $s^* \in \mathcal{H}_k^{(1,\ldots,n)}$. There is a set $H_k \subset \bigcup_{j<k} H_j^{(1,\ldots,n)}$ such that*

$$W_{s^*}^c \cap E = \bigcup_{s\in H_k}(A(\bar{\Gamma}_s) \cap E). \tag{56}$$

Moreover, let $U \subset E$ be a non-empty set with

$$U \subset \overline{W_{s^*}^c} \quad \textit{and} \quad \textit{int}(W_{s^*}^c \cap U) = \emptyset, \tag{57}$$

then we can find some $j < k$ and $s' \in \mathcal{H}_j^{(1,\dots,n)}$ such that $A(\bar{\Gamma}_{s'})$ is locally dense in U and int$(A(\bar{\Gamma}_{s'}) \cap U) = \emptyset$.

Proof. We first prove (56) for the given k. Let $s_{k,1} = s^*$, then for each $j = k, \dots, 2$, Lemma 10 and (55) indicate that there exist some indices $s_{j-1,1}, s_{j-1,2} \in \mathcal{H}$ such that

$$\Gamma_{s_{j,1}} = \frac{\bar{\Gamma}_{s_{j-1,1}}}{\bar{\Gamma}_{s_{j-1,2}}}. \tag{58}$$

Denote $H_k \triangleq \{s_{j,2}, j = 1, \dots, k-1\}$.

Note that by (53), (55), it is easy to see

$$\begin{aligned} &\{y \in E : \Gamma_{s^*}(y) \text{ is well-defined}\} \\ = \ &\{y \in E : D\Gamma_{s^*}(y) \text{ is well-defined}\}. \end{aligned}$$

In addition, Lemma 10 infers that for each $j = 1, \dots, k$,

$$\begin{aligned} &\{y \in E : \Gamma_{s_{j-1,1}}(y) \text{ is well-defined}\} \\ = \ &\{y \in E : \Gamma_{s_{j-1,2}}(y) \text{ is well-defined}\}. \end{aligned}$$

Then, by (58),

$$\begin{aligned} &W_{s^*}^c \cap E = \{y \in E : \Gamma_{s^*}(y) \text{ is undefined}\} \\ = \ &\{y \in E : \bar{\Gamma}_{s_{k-1,1}}(y) \text{ is undefined}\} \cup A(\bar{\Gamma}_{s_{k-1,2}}) \\ = \ &\cdots = \bigcup\nolimits_{s \in H_k} (A(\bar{\Gamma}_s) \cap E), \end{aligned}$$

which is exactly (56). So, if (57) holds, for every $s \in H_k$, int$(A(\bar{\Gamma}_s) \cap U) = \emptyset$. Finally, we show that for some $s' \in H_k$, $A(\bar{\Gamma}_{s'})$ is locally dense in U. Otherwise, $A(\bar{\Gamma}_s)$ is nowhere dense in U for every $s \in H_k$. This means there are a series of nonempty open intervals $U_1 \subset \cdots \subset U_{k-1} \subset U$ such that $U_j \cap \overline{A(\bar{\Gamma}_{s_{l,2}})} = \emptyset$ for all $l = j, \dots, k-1$. As a consequence, by (56), $U_1 \cap \overline{W_{s^*}^c} = U_1 \cap \left(\bigcup_{j=1}^{k-1} \overline{A(\bar{\Gamma}_{s_{j,2}})}\right) = \emptyset$, which contradicts to (57) due to $U_1 \subset U$. ■

Now, we are ready to construct $\{S^j(q)\}_{j=1}^p$. For this, we classify the sets $\overline{A(\bar{\Gamma}_s)}$, $s \in \mathcal{H}$, into three types:

$$\begin{cases} Z_s^1 = \text{int}(A(\bar{\Gamma}_s)) \\ Z_s^2 = d(A(\bar{\Gamma}_s)) \backslash Z_s^1 \\ Z_s^3 = A(\bar{\Gamma}_s) \backslash d(A(\bar{\Gamma}_s)) \end{cases}, \tag{59}$$

where $d(A)$ denotes the derived set of A. Observe that Z_s^1 can be expressed by a countable union of disjoint open intervals and Z_s^3 is in fact the set of the isolated points of $A(\bar{\Gamma}_s)$. Both the two sets have good topological properties. However, the structure of Z_s^2 is not that clear. Therefore, we define the following sets to exclude Z_s^2:

$$S \triangleq E \Big\backslash \left(\bigcup\nolimits_{s \in \mathcal{H}} Z_s^2\right),$$

which are clearly some open sets.

The key idea of the construction of $\{S^j(q)\}_{j=1}^p$ is to find a proper subset of S. To begin with, we prove

Lemma 12. *Under Assumption A3, for any unit $x \in \mathbb{R}^n$,*

$$\ell\left(\{y \in S : |\phi^T(y)x| > 0\}\right) > 0. \tag{60}$$

Proof. We show the lemma in a way of reduction to absurdity. Suppose there exists some $x \in \mathbb{R}^n$ with $\|x\| = 1$ such that

$$\ell\left(\{y \in S : |\phi^T(y)x| > 0\}\right) = 0. \tag{61}$$

As $\phi(\cdot)$ is continuous on open set $S \subset E$, then

$$\phi^T(y)x = 0, \quad \forall y \in S \tag{62}$$

Note that Assumption A3 yields

$$\ell\left(\{y \in E : |\phi^T(y)x| > 0\}\right) > 0,$$

which together with (61) implies

$$\ell\left(\{y \in E \backslash S : |\phi^T(y)x| > 0\}\right) > 0.$$

Consequently, there is a $y^* \in E \backslash S$ such that $|\phi^T(y^*)x| > 0$. By the continuity of $\phi(\cdot)$ on E, there is $\varepsilon > 0$ such that

$$|\phi^T(y)x| > 0, \quad \forall y \in \prod_{i=1}^n (y_i^* - \varepsilon, y_i^* + \varepsilon) \subset E. \tag{63}$$

On account of (62) and (63), we deduce

$$V = (y^* - \varepsilon, y^* + \varepsilon) \ \subset \ E \backslash S = \bigcup_{s \in \mathcal{H}} Z_s^2. \tag{64}$$

Next, we show (64) is impossible. To this end, note that Z_s^2 is closed for each $s \in \mathcal{H}$, and hence (64) implies that there is an integer $k \in [1, n]$ and an array $s^* \in \mathcal{H}_k^{(1,\dots,n)}$ such that $Z_{s^*}^2$ is locally dense in V. Let k be the smallest integer for such s^*.

Now, fix the above $k \in [1, n]$ and $s^* \in \mathcal{H}_k^{(1,\dots,n)}$. Since $Z_{s^*}^2$ is locally dense in V, there is an open interval $V' \subset V$ such that $Z_{s^*}^2$ is dense in V'. Moreover, $Z_{s^*}^2$ is closed, so $V' \subset Z_{s^*}^2$ and thus $V' \cap Z_{s^*}^1 = \emptyset$. In addition, $\bar{\Gamma}_{s^*}$ is continuous in $W_{s^*} \cap E$, by (59), $W_{s^*} \cap Z_{s^*}^2 \cap E \subset A(\bar{\Gamma}_{s^*})$. Consequently,

$$\text{int}\left(W_{s^*} \cap Z_{s^*}^2 \cap E\right) \subset \text{int}(A(\bar{\Gamma}_{s^*}) \cap E) = Z_{s^*}^1 \cap E. \tag{65}$$

Moreover, V' is an open interval belongs to E and $V' \cap (Z_{s^*}^2)^c = \emptyset$, then

$$\begin{aligned} V' = V' \backslash Z_{s^*}^1 \ &\subset \ V' \backslash \text{int}\left(W_{s^*} \cap Z_{s^*}^2\right) \\ &\subset \ \overline{V_i' \backslash (W_{s^*} \cap Z_{s^*}^2)} = \overline{V' \cap W_{s^*}^c} \subset \overline{W_{s^*}^c}. \end{aligned} \tag{66}$$

Note that $\bar{\Gamma}_s$ are well defined in $\mathbb{R}$ for all $s \in \mathcal{H}_1^{(1,\dots,n)}$, which shows $W_s^c \cap E = \emptyset$. Then, (66) implies $k \geq 2$. Furthermore, by $Z_{s^*}^2 \subset \overline{A(\bar{\Gamma}_{s^*})}$, it yields

$$\begin{aligned} \text{int}\left(W_{s^*}^c \cap V'\right) \ &\subset \ \text{int}\left(A^c(\bar{\Gamma}_{s^*}) \cap V'\right) = \left(\overline{(A^c(\bar{\Gamma}_{s^*}) \cap V')^c}\right)^c \\ &= \ \left(\overline{A(\bar{\Gamma}_{s^*})} \cup (V')^c\right)^c \subset \left(Z_{s^*}^2 \cup (V')^c\right)^c \\ &= \ \emptyset. \end{aligned} \tag{67}$$

Applying Lemma 11, (66) and (67) indicates that we can find some $j < k$ and $s' \in \mathcal{H}_j^{(1,\dots,n)}$ such that $A(\bar{\Gamma}_{s'})$ is locally dense in V' and int$(A(\bar{\Gamma}_{s'}) \cap V') = \emptyset$. So, there is an open interval $V'' \subset V'$ such that $V'' \subset d(A(\bar{\Gamma}_{s'}))$ and int$(A(\bar{\Gamma}_{s'})) \cap V'' = \text{int}(A(\bar{\Gamma}_{s'}) \cap V'') = \emptyset$, and then $V'' \subset Z_{s'}^2$. That is, $Z_{s'}^2$ is locally dense in V', which derives a contradiction to the definition of k. This completes the proof of Lemma 12. ■

Next, we consider a series of open sets $\{S \cap (-j, j)\}_{j \geq 1}$. Clearly, $S \cap (-j, j) \subset S \cap (-(j+1), (j+1))$ and $\lim_{j \to +\infty} S \cap (-j, j) = S$. Then, by using Lemmas 9 and 12, there is an integer $d \geq 1$ such that for any unit $x \in \mathbb{R}^n$,

$$\ell\left(\{y \in S \cap (-d, d) : |\phi^T(y)x| > 0\}\right) > 0. \tag{68}$$

Since S is open, there exists some disjoint open intervals $\{S^j\}_{j\in\Theta}$, where $\Theta = \{1,\dots,k\}$ (k can be taken infinite), such that $S\cap(-d,d) = \bigcup_{j\in\Theta} S^j$. Write $S^j = (c^j, d^j)$ and denote

$$S^j(q) \triangleq \left(c^j + \frac{d^j - c^j}{q+2}, d^j - \frac{d^j - c^j}{q+2}\right), j\in\Theta, q\in\mathbb{N}^+. \quad (69)$$

Given (68), the following lemma is natural.

Lemma 13. *If (68) holds, then there exist some integers p and $q \geq 1$ such that for any unit $x\in\mathbb{R}^n$,*

$$\ell\left(\{y\in\mathcal{S} : |\phi^T(y)x| > 0\}\right) > 0 \quad \text{and} \quad \overline{\mathcal{S}} \subset E, \quad (70)$$

where $\mathcal{S} \triangleq \bigcup_{j=1}^p S^j(q)$.

Proof. It is obvious that $\bigcup_{j\in\Theta} S^j(q) \subset \bigcup_{j\in\Theta} S^j(q+1)$, $q\in\mathbb{N}^+$. If $|\Theta| < +\infty$, then $\lim_{q\to+\infty}\bigcup_{j\in\Theta} S^j(q) = S\cap(-d,d)$. As for the case where $\Theta = \mathbb{N}^+$, it infers $\lim_{k\to+\infty}\bigcup_{j=1}^k S^j(k) = S\cap(-d,d)$. So, in view of the above two cases, by (68) and Lemma 9, there are some integers p and $q\geq 1$ such that (70) holds. ∎

2) Selection of U_x: With the foregoing preliminaries in place, we can set out to construct U_x. First, for every $x\in\mathbb{R}^n$ with $\|x\| = 1$, define

$$U_x(\delta) \triangleq \{y : |\phi^T(y)x| > \delta\}\cap\mathcal{S}, \quad \delta > 0.$$

The remaining task is to take a proper $\delta > 0$ such that $U_x = U_x(\delta)$ meet our requirement. We choose δ according to the lemma below.

Lemma 14. *Under Assumption A3, there is a $\delta^* > 0$ that*

$$\inf_{\|x\|=1} \ell(\{y : |\phi^T(y)x| > \delta^*\}\cap\mathcal{S}) > 0. \quad (71)$$

Proof. Arguing by contradiction, we assume that (71) is false. Then, for each integer $k\geq 1$, there exists some point $z(k)\in\mathbb{R}^n$ with $\|z(k)\| = 1$ such that

$$\ell(\{y : |\phi^T(y)z(k)| > \frac{1}{k}\}\cap\mathcal{S}) < \frac{1}{k}. \quad (72)$$

This sequence of points thus has a subsequence $\{z(k_r)\}_{r\geq1}$ and an accumulation point z^* such that

$$\lim_{r\to+\infty} z(k_r) = z^*. \quad (73)$$

So, $\|z^*\| = 1$. If $\ell\left(\{y : |\phi^T(y)z^*| > 0\}\cap\mathcal{S}\right) = 0$, this clearly contradicts to Lemma 13. Therefore, by (69),

$$\begin{aligned}&\lim_{k\to+\infty} \ell(\{y : |\phi^T(y)z^*| > \frac{1}{k}\}\cap\mathcal{S}) \\ = \;& \ell\left(\{y : |\phi^T(g(y)z^*| > 0\}\cap\mathcal{S}\right) > 0. \quad (74)\end{aligned}$$

This implies that there exists an integer $h\geq 1$ such that

$$\ell(\{y : |\phi^T(y)z^*| > \frac{1}{h}\}\cap\mathcal{S}) > 0. \quad (75)$$

Now, (72) and (75) yield that for all sufficiently large r,

$$\begin{aligned}\frac{1}{k_r} \;&>\; \ell(\{y : |\phi^T(y)z(k_r)| > \frac{1}{k_r}\}\cap\mathcal{S}) \\ &=\; \ell(\{y : |\phi^T(y)z(k_r)| > \frac{1}{k_r}\}\cap\mathcal{S}) \\ &\geq\; \ell(\{y\in\mathcal{S} : |\phi^T(y)z^*| > \frac{1}{k_r} + \frac{1}{h}\}).\end{aligned}$$

Letting $r\to+\infty$ in the above inequality infers

$$\begin{aligned}0 \;&\geq\; \lim_{r\to+\infty} \ell(\{y\in\mathcal{S} : |\phi^T(y)z^*| > \frac{1}{k_r} + \frac{1}{h}\}) \\ &=\; \ell(\{y\in\mathcal{S} : |\phi^T(y)z^*| > \frac{1}{h}\}), \quad (76)\end{aligned}$$

which contradicts to (75). ∎

At the end of this section, according to Lemma 14, we select a δ^* such that (71) holds. Now, for any unit vector $x\in\mathbb{R}^n$, define $U_x \triangleq U_x(\delta^*)$.

B. The Properties of U_x

To analyze the properties of U_x, we first show

Lemma 15. *Let $x = (x_1,\dots,x_n)^T\in\mathbb{R}^n$ be a non-zero vector and $d > c$ be two numbers satisfying $[c,d]\subset E$. Also, let $\{r_l\}_{l=1}^{2^{n-1}}$ be a sequence of numbers that $d\geq r_1 > r_2 > \cdots > r_{2^{n-1}} \geq c$ and*

$$\sum_{j=1}^n f_j'(r_l)x_j = 0, \quad 1\leq l\leq 2^{n-1}, \quad (77)$$

where $f_j' = Df_j, j = 1,\dots,n$. Then, the following two statements hold:
(i) there exists an array $s\in\mathcal{H}$ such that $A(\bar{\Gamma}_s)\cap[c,d]\neq\emptyset$;
(ii) if for every $s\in\mathcal{H}$, $A(\bar{\Gamma}_s)\cap[c,d]$ is either $\emptyset$ or $[c,d]$, then

$$\sum_{j=1}^n f_j'(y)x_j = 0, \quad \forall y\in[c,d]. \quad (78)$$

Proof. (i) Suppose $\bigcup_{s\in\mathcal{H}} A(\bar{\Gamma}_s)\cap[c,d] = \emptyset$. Then, for each integer $k\in[1,n]$, there exist 2^{k-1} numbers $\{\varepsilon_{k,l}\}_{l=1}^{2^{k-1}}$ satisfying $d\geq\varepsilon_{k,1} > \cdots > \varepsilon_{k,2^{k-1}} \geq c$ and

$$\sum_{j=1}^k \bar{\Gamma}_{(j,k+1,\dots,n)}(\varepsilon_{k,l})x_j = 0, \quad 1\leq l\leq 2^{k-1}, \quad (79)$$

where $\bar{\Gamma}_{(j,n+1,\dots,n)} \triangleq f_j'$.

As a matter of fact, when $k = n$, (77) leads to (79) immediately. We now prove (79) by induction. Assume (79) holds for $k = m'$, where m' is an integer in $[2,n]$. Hence we can find 2^{n-1} numbers of $\{\varepsilon_{n,l}\}_{l=1}^{2^{n-1}}$ such that $d\geq\varepsilon_{n,1} > \cdots > \varepsilon_{n,2^{n-1}} \geq c$ and $\sum_{j=1}^n \bar{\Gamma}_{(j,n+1,\dots,n)}(\varepsilon_{n,l})x_j = 0$, $1\leq l\leq 2^{n-1}$. By $\bigcup_{s\in\mathcal{H}} A(\bar{\Gamma}_s)\cap[c,d] = \emptyset$, every $\Gamma_{(j,m',\dots,n)}$ is well-defined in $[c,d]$, then for $1\leq l\leq 2^{m'-1}$,

$$\begin{aligned}\sum_{j=1}^{m'-1}\Gamma_{(j,m',\dots,n)}(\varepsilon_{m',l})x_j \;&=\; \sum_{j=1}^{m'-1}\frac{\bar{\Gamma}_{(j,m'+1,\dots,n)}}{\bar{\Gamma}_{(m',\dots,n)}}(\varepsilon_{m',l})x_j \\ &=\; -x_{m'}.\end{aligned}$$

Taking account of the *Rolle's theorem*, there are some $\{\varepsilon_{m'-1,l}\}_{l=1}^{2^{m'-2}}$ with $\varepsilon_{m'-1,l}\in(\varepsilon_{m',2l-1},\varepsilon_{m',2l})$ such that

$$\sum_{j=1}^{m'-1}\bar{\Gamma}_{(j,m',\dots,n)}(\varepsilon_{m'-1,l})x_j = 0, \quad 1\leq l\leq 2^{m'-2}.$$

So (79) holds for $k = m'-1$ and this completes the induction.

Now, by letting $k=1$ in (79), there is a number $\varepsilon_{1,1}\in[c,d]$ such that

$$\bar{\Gamma}_{(1,\ldots,n)}(\varepsilon_{1,1})x_1=0.$$

Since $\bigcup_{s\in\mathcal{H}}A(\bar{\Gamma}_s)\cap[c,d]=\emptyset$, $\bar{\Gamma}_{(1,\ldots,n)}(\varepsilon_{1,1})\neq 0$, and hence $x_1=0$. By the symmetry of $x_1,\ldots,x_n$ in (79), we conclude that $x_j=0$ for all $j\in[1,n]$. But this is impossible due to $\|x\|\neq 0$ and thus $\bigcup_{s\in\mathcal{H}}A(\bar{\Gamma}_s)\cap[c,d]\neq\emptyset$.
(ii) Let I be an open interval containing $[c,d]$. It suffices to prove the claim that for every function sequence $f_1,\ldots,f_n\in C^n(I)$ satisfying (77), if $A(\bar{\Gamma}_s)\cap[c,d]$ is either $\emptyset$ or $[c,d]$, $\forall s\in\mathcal{H}$, then (78) holds. We show it by induction. When $n=1$, (77) reduces to $f_1'(r_1)x_1=0$. Since $x_1\neq 0$, $A(f_1')\cap[c,d]\neq\emptyset$, which means $A(f_1')\supset[c,d]$ by assumption. So, $f_1'(y)x_1\equiv 0$ for all $y\in[c,d]$. Suppose the claim mentioned above holds for all $n\in[1,h-1]$, $h\geq 2$.

We now consider the claim for $n=h$. In this case, the non-zero vector $x=(x_1,\ldots,x_h)^T$. First, assume that there is an integer $j'\in[1,h]$ such that $|x_{j'}|<\|x\|$ and

$$A(f_{j'}')\cap[c,d]=\emptyset. \tag{80}$$

Without loss of generality, let $j'=h$. Define the following $h-1$ functions:

$$F_j\triangleq\Gamma_{(j,h)},\quad 1\leq j\leq h-1.$$

Owing to (80), $F_j\in C^{h-1}(I)$, $1\leq j\leq h-1$ with $h\geq 2$ are well-defined. Moreover, (77) yields

$$\sum_{j=1}^{h-1}F_j(r_l)x_j=-x_h,\quad 1\leq l\leq 2^{h-1}. \tag{81}$$

Therefore, by applying the *Rolle's theorem*, there exist 2^{h-2} numbers $\varepsilon_l\in[r_{2l},r_{2l-1}]$, $l\in[1,2^{h-2}]$ such that

$$\sum_{j=1}^{h-1}DF_j(\varepsilon_l)x_j=0,\quad 1\leq l\leq 2^{h-2}. \tag{82}$$

Here, $(x_1,\ldots,x_{h-1})^T$ is nonempty by $|x_h|<\|x\|$.

Since for every $(i_1,\ldots,i_l)\in\bigcup_{k=1}^{h-1}\mathcal{H}_k^{(1,\ldots,h-1)}$, one has $(i_1,\ldots,i_l,h)\in\bigcup_{k=1}^{h}\mathcal{H}_k^{(1,\ldots,h)}$, by (53),

$$\Lambda_l(F_{i_1},\ldots,F_{i_l})=\Lambda_{l+1}(f_{i_1},\ldots,f_{i_l},f_h)=\Gamma_{(i_1,\ldots,i_l,h)}.$$

Because $A(\bar{\Gamma}_s)\cap[c,d]$ is either $\emptyset$ or $[c,d],\forall s\in\bigcup_{k=1}^{h}\mathcal{H}_k^{(1,\ldots,h)}$, the above inequality yields

$$A\left(D\Lambda_l(F_{i_1},\ldots,F_{i_l})\right)\cap[c,d]=\emptyset\text{ or }[c,d].$$

Consequently, by the induction hypothesis with $n=h-1$ and $f_{i_j}=F_j, j\in[1,h-1]$ satisfying (82), we conclude

$$\sum_{j=1}^{h-1}DF_j(y)x_j=0,\quad\forall y\in[c,d]. \tag{83}$$

In view of (81) and (83), we deduce that $\sum_{j=1}^{h-1}F_j(y)x_j=-x_h$ for any $y\in[c,d]$, and hence

$$\sum_{j=1}^{h}f_j'(y)x_j\equiv 0\quad\forall y\in[c,d]. \tag{84}$$

Now, it remains to consider the case that for each integer $j\in[1,h]$, either $|x_j|=\|x\|$ or $[c,d]\subset A(f_j')$. If $|x_j|<\|x\|$ for all $j\in[1,h]$, then $f_j'\equiv 0$ in $[c,d]$ for all $j\in[1,h]$, which leads to (84). So, assume there is an integer $j'\in[1,h]$ that $|x_{j'}|=\|x\|$. Without loss of generality, let $j'=h$, then $x_j=0$ for all $j\in[1,h-1]$. Substituting this into (77), one has $f_h'(r_l)=0$, $1\leq l\leq 2^{h-1}$. The induction hypothesis thus yields $[c,d]\subset A(f_h')$, and hence $\sum_{j=1}^{h}f_j'(y)x_j=f_h'(y)x_h=0$, $\forall y\in[c,d]$. Therefore, the claim is true for $n=h$ and we complete the induction. ∎

We now return to analyze $Z_s^1\cap(\bigcup_{j=1}^{p}S^j(q))$. Observe that for each array $s\in\mathcal{H}$, if $Z_s^1\cap(\bigcup_{j=1}^{p}S^j(q))\neq\emptyset$, it is a countable union of disjoint open intervals. Denote the set of these intervals by $\mathcal{G}_s\triangleq\{I_s^j\}_{j\geq 1}$, where

$$I_s^j=(a_s^j,b_s^j),\quad j=1,2,\ldots. \tag{85}$$

Lemma 16. *Define* $\mathcal{G}\triangleq\bigcup_{s\in\mathcal{H}}\mathcal{G}_s$ *and* $H\triangleq\left(\bigcup_{s\in\mathcal{H}}Z_s^3\right)\cap\left(\bigcup_{j=1}^{p}S^j(q)\right)$. *Then,*

$$|\mathcal{G}|<+\infty\quad\textit{and}\quad|H|<+\infty. \tag{86}$$

Proof. Suppose $|\mathcal{G}|=+\infty$, then there is an array $s^*\in\mathcal{H}$ such that $|\mathcal{G}_{s^*}|=+\infty$. Let $y\in\bigcup_{j=1}^{p}\overline{S^j(q)}$ be an accumulation point of $\{b_{s^*}^j\}_{j\geq 1}$. By the continuity of $\bar{\Gamma}_{s^*}$ in set $\bigcup_{j=1}^{p}\overline{S^j(q)}\subset E$, $y\in\bigcup_{j=1}^{p}\overline{S^j(q)}\cap A(\bar{\Gamma}_{s^*})$. Moreover, it is evident that $y\notin Z_{s^*}^1\cup Z_{s^*}^3$, so $y\in Z_{s^*}^2\cap(\bigcup_{j=1}^{p}\overline{S^j(q)})\subset E\cap\left(\bigcup_{s\in\mathcal{H}}Z_s^2\right)$. However,

$$E\cap\left(\bigcup_{s\in\mathcal{H}}Z_s^2\right)=E_i\backslash S\subset E\backslash\bigcup_{j=1}^{p}S^j, \tag{87}$$

and

$$\left(E\backslash\bigcup_{j=1}^{p}S^j\right)\cap\left(\bigcup_{j=1}^{p}\overline{S^j(q)}\right)=\emptyset. \tag{88}$$

The contradiction is derived immediately by comparing (87), (88) and the fact $y\in E\cap\left(\bigcup_{s\in\mathcal{H}}Z_s^2\right)$. Thus, $|\mathcal{G}|<+\infty$.

As to $|H|<+\infty$, the proof is quite similar to that given for $|\mathcal{G}|<+\infty$ and is omitted. ∎

The following lemma is based on the above two lemmas.

Lemma 17. *Let* $x\in\mathbb{R}^n$ *be a non-zero vector. Denote* $(\phi)'=(f_1',\ldots,f_n')$ *and*

$$\begin{cases} K=\mathit{int}(A(x^T(\phi)'))\cap\left(\bigcup_{j=1}^{p}\overline{S^j(q)}\right)\\ L=(A(x^T(\phi)')\cap\left(\bigcup_{j=1}^{p}\overline{S^j(q)})\right)\setminus K\end{cases},$$

then $|L|\leq 2^np(3|\mathcal{G}|+|H|+2)$.

Proof. Let $\mathcal{Q}_j\triangleq(A(x^T(\phi)')\cap\overline{S^j(q)})\setminus K$, $j=1,\ldots,p$. We first show that the cardinality of each $\mathcal{Q}_j$ is finite. Otherwise, for some $j\in[1,p]$, there is a monotone sequence $\{r_l\}_{l\geq 1}$ in $\mathcal{Q}_j$ such that $r_l\neq r_{l'}$ for each $l\neq l'$ and $\lim_{l\to+\infty}r_l=y^*$ for some $y^*\in\overline{S^j(q)}$. Without loss of generality, let $r_l<r_{l'}$ if $l>l'$. Divide this sequence into infinite groups:

$$\left\{r_{2^nk+1},r_{2^nk+2},\ldots,r_{2^n(k+1)}\right\},\quad k=0,1,\ldots$$

and for each $k \geq 0$, define

$$D_k \triangleq [r_{2^n(k+1)}, r_{2^n k+1}]. \tag{89}$$

So, given $k \geq 0$, $D_k \subset \overline{S^j(q)} \subset E$ and $r_{2^n k+1} > \ldots > r_{2^n(k+1)}$ satisfy

$$x(\phi)'(r_l) = 0, \quad 2^n k + 1 \leq l \leq 2^n(k+1). \tag{90}$$

Note that the definition of $\mathcal{Q}_j$ yields $x(\phi)' \not\equiv 0$ on D_k, applying Lemma 15 with $[c,d] = D_k$ indicates that there is an array $s_k \in \mathcal{H}$ fulfilling $A(\bar{\Gamma}_{s_k}) \cap D_k \neq \emptyset$ and $D_k \not\subset A(\bar{\Gamma}_{s_k})$. Hence, at least one of following three cases occurs:
Case 1: $Z^3_{s_k} \cap D_k \neq \emptyset$.
Case 2: There is an interval $I^j_{s_k} \in \mathcal{G}_{s_k}$ such that $I^j_{s_k} \subset D_k$.
Case 3: There is an interval $I^j_{s_k} \in \mathcal{G}_{s_k}$ satisfying $I^j_{s_k} \cap D_k \neq \emptyset$ and $D_k \not\subset I^j_{s_k}$.

Let $O_l, l = 1,2,3$ present the times of Case l that occurs for some $k \geq 0$. Since $D_k \cap D_{k'} = \emptyset$ for $k \neq k'$,

$$O_1 \leq |H| \quad \text{and} \quad O_2 \leq |\mathcal{G}|. \tag{91}$$

Furthermore, for each interval $I^j_s \in G_s$, there exist at most two distinct D_k such that $I^j_s \cap D_k \neq \emptyset$ and $D_k \not\subset I^j_s$. Hence,

$$O_3 \leq 2|\mathcal{G}|. \tag{92}$$

Combining (91) and (92) yields

$$O_1 + O_2 + O_3 \leq 3|\mathcal{G}| + |H| < +\infty. \tag{93}$$

However, $O_1 + O_2 + O_3 = +\infty$ because D_k is infinite. So, the cardinality of each $\mathcal{Q}_j$ is finite.

Now, let $j \in [1,p]$ be an index that $|\mathcal{Q}_j| > 2^n(3|\mathcal{G}| + |H| + 2)$. Then, write the points of $\mathcal{Q}_j$ from left to right as $v_0, v_1, \ldots, v_{|\mathcal{Q}_j|-1}$. Define

$$h \triangleq \left\lfloor \frac{|\mathcal{Q}_j| - 2}{2^n} \right\rfloor > 3|\mathcal{G}| + |H| \tag{94}$$

and $D'_k \triangleq [v_{2^n(k+1)}, v_{2^n k+1}]$, $k = 0,1,\ldots,h-1$. By an analogous proof as (90)–(93), we arrive at $h \leq 3|\mathcal{G}| + |H|$, which arises a contradiction to (94). So $|\mathcal{Q}_j| \leq 2^n(3|\mathcal{G}| + |H| + 2)$ and hence $|L| \leq 2^n p(3|\mathcal{G}| + |H| + 2)$. ■

For any $x \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$, it is clear that set $\{y : x^T\phi(y) > \delta\} \cap (\bigcup_{j=1}^p S^j(q))$ is open. If this set is not empty, then it is a countable union of disjoint open intervals. Denote the set of these intervals by $\mathcal{U}(\delta)$.

Lemma 18. *For any non-zero $x \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$,*

$$|\mathcal{U}(\delta)| \leq p(|L| + 2). \tag{95}$$

Proof. Let $\mathcal{K}_j = \{I \in \mathcal{U}(\delta) : I \cap S^j(q) \neq \emptyset\}, j \in [1,p]$, then

$$|\mathcal{U}(\delta)| \leq \sum_{j=1}^{p} |\mathcal{K}_j|. \tag{96}$$

Fix an index $j \in [1,p]$ and $I \in \mathcal{K}_j$. By the continuity of ϕ in $S^j(q)$, each endpoint of I either belongs to the zero set $A(x^T\phi(y) - \delta)$ or is an endpoint of $S^j(q)$. If $\partial(I) \cap \partial(S^j(q)) = \emptyset$, then $\partial(I) \in A(x^T\phi(y) - \delta)$. By the *Rolle's theorem*, it follows that $\{y : x^T(\phi)'(y) = 0\} \cap I \neq \emptyset$, which together with $I \subset \{y : x^T\phi(y) > \delta\}$ leads to $L \cap I \neq \emptyset$. Note that there are at most two intervals $I \in \mathcal{K}_j$ satisfying $\partial(I) \cap \partial(S^j(q)) \neq \emptyset$ and any two intervals in $\mathcal{K}_j$ are disjoint, so

$$|\mathcal{K}_j| \leq |L| + 2. \tag{97}$$

Finally, (95) is an immediate result of (96) and (97). ■

Given a closed interval $I_1 \subset \mathbb{R}$ and a positive integer r, equally divide I_1 into r closed intervals $\{I_{1,j}\}_{j=1}^r$ that $\text{int}(I_{1,j}) \cap \text{int}(I_{1,j'}) = \emptyset$ if $j \neq j'$. So, there are r small closed intervals $\{I_{1,j}\}_{j=1}^r$. Let $\mathcal{T}(O,r)$ be the set of the r small intervals. Clearly, for any distinct intervals $U, U' \in \mathcal{T}(O,r)$, $\text{int}(U) \cap \text{int}(U') = \emptyset$. Define

$$\mathcal{T}_\delta(O,r) \triangleq \left\{U \in \mathcal{T}(O,r) : \mathcal{B}(\delta) \cap \overline{\mathcal{S}} \cap U \neq \emptyset\right\}, \tag{98}$$

where $\mathcal{B}(\delta) \triangleq \partial(\{y : \phi^T(y)x > \delta\})$ and $\mathcal{S}$ is defined in Lemma 13. Let $\mathcal{K}_\delta(O,x,r) \triangleq |\mathcal{T}_\delta(O,r)|$. The following lemma is critical to our result.

Lemma 19. *There is a number $C > 0$ such that for any closed interval I_1, non-zero vector $x \in \mathbb{R}^n$, $\delta \in \mathbb{R}$ and integer $r \geq 1$,*

$$\mathcal{K}_\delta(I_1, x, r) \leq C. \tag{99}$$

Proof. By Lemma 18 , it is easy to check that

$$\left|\mathcal{B}(\delta) \cap \left(\bigcup_{j=1}^p S^j(q)\right)\right| \leq 2p(|L| + 2). \tag{100}$$

Moreover, $\mathcal{B}(\delta) \cap \left(\bigcup_{j=1}^p \overline{S^j(q)}\right) \subset \mathcal{B}(\delta) \cap \left(\bigcup_{j=1}^p S^j(q)\right) \cup \partial\left(\bigcup_{j=1}^p S^j(q)\right)$, so it follows that

$$\mathcal{K}_\delta(I_1, x, r) \leq 2\left|\mathcal{B}(\delta) \cap \left(\bigcup_{j=1}^p \overline{S^j(q)}\right)\right| \leq 4p(|L| + 2) + 4p.$$

Hence, (99) is true by taking $C = 4p(|L| + 2) + 4p$. ■

Now, recall the definition of U_x in the end of Section A, $\partial(U_x) \subset ((\partial(\{y : \phi^T(y)x > \delta^*\}) \cup \partial(\{y : -\phi^T(y)x > \delta^*\})) \cap \mathcal{S}) \cup \partial(\mathcal{S})$. Given a closed interval I_1 and an integer $r \geq 1$, observe that $|\{U \in \mathcal{T}(I_1, r) : \partial(\mathcal{S}) \cap U \neq \emptyset\}| \leq 4p$. In addition, by applying Lemma 19 it follows that there is a constant $C_0 > 0$ depending only on ϕ such that

$$|\{U \in \mathcal{T}(I_1, r) : \partial(U_x) \cap U \neq \emptyset\}| \leq C_0. \tag{101}$$

C. The Estimation of Minimal Eigenvalue

We state a lemma modified from [10]. Now, for the set U_x we have constructed, define a random process g_x by

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

Lemma 20. *For any $\epsilon > 0$, there is a class $\mathcal{G}_\epsilon$ such that*
(i) each element of $\mathcal{G}_\epsilon$, denoted by g_ϵ, is a random series $\{g_\epsilon(i)\}_{i \geq 1}$ with the form

$$g_\epsilon(i) = I_{\{y_i \in U_\epsilon\}} - P(y_i \in U_\epsilon | \mathcal{F}^y_{i-1}) - \epsilon, \quad i \geq 1, \tag{102}$$

where U_ϵ is a set in $\mathbb{R}$;
(ii) $\mathcal{G}_\epsilon$ contains a lower process g_ϵ to each g_x in the sense

$$g_\epsilon(i) \leq g_x(i), \quad \forall i \geq 1. \tag{103}$$

Proof. (i) Let I_1 be a closed interval contains $\mathcal{S}$. Let r be an integer such that

$$r > 2\epsilon^{-1}\overline{\rho}C_0 \cdot \ell(I_1), \tag{104}$$

where C_0 is defined in (101) and $\overline{\rho} \triangleq \sup_{x\in\mathbb{R}} \rho(x)$, where $\rho(x)$ is the density function of $N(0,\sigma_w^2)$. Let U_ϵ be a union of some intervals taken from $\mathcal{T}(O,r)$. Hence, for a fixed U_ϵ, we can define a random process g_ϵ by (102). Denote $\mathcal{G}_\epsilon$ as the class of all such g_ϵ.

(ii) Note that U_x is bounded for every $x \in \mathbb{R}^n$ with $\|x\| = 1$. Then, there is a set U_ϵ such that $U_\epsilon \subset U_x$ and $\Delta U_{\epsilon,x} \triangleq U_x - U_\epsilon$ falls into a union of finite intervals $J_1,\ldots,J_l \in \{U \in \mathcal{T}(I_1,r) : \partial(U_x) \cap U \neq \emptyset\}$. By (101),

$$\sum_{k=1}^{l} \ell(J_k) \;=\; l \cdot \frac{\ell(I_1)}{r} \le C_0 \cdot \ell(I_1) \cdot \frac{1}{r} < \frac{1}{2\overline{\rho}}\epsilon. \tag{105}$$

We now calculate $P(y_t \in \Delta U_{\epsilon,x}|\mathcal{F}_{t-1}^y)$. By (105), $\sigma_{t-1} \ge \sigma_w$, and y_t possesses a conditional gaussian distribution $N(0,\sigma_{t-1}^2)$ with $\mathcal{F}_{t-1}^y$, it is easy to see

$$\begin{aligned} & P(y_t \in \Delta U_{\epsilon,x}|\mathcal{F}_{t-1}^y) \\ \le\; & P(y_t \in \bigcup_{k=1}^{l} J_k|\mathcal{F}_{t-1}^y) \le \ell\left(\bigcup_{k=1}^{l} J_k\right) \cdot \overline{\rho} < \frac{\epsilon}{2}. \end{aligned}$$

So, for any $i \ge 1$,

$$\begin{aligned} g_x(i) \;&=\; I_{\{y_i \in U_x\}} - P(y_i \in U_x|\mathcal{F}_{i-1}^y) \\ &=\; I_{\{y_i \in U_x\}} - P(y_i \in U_\epsilon|\mathcal{F}_{i-1}^y) - P(y_i \in \Delta U_{\epsilon,x}|\mathcal{F}_{i-1}^y) \\ &\ge\; I_{\{y_i \in U_\epsilon\}} - P(y_i \in U_\epsilon|\mathcal{F}_{i-1}^y) - \epsilon = g_\epsilon(i), \end{aligned}$$

which is exactly (103). ■

Proof of Proposition 1. First, recall the definition of U_x, for any $x \in \mathbb{R}^n$ with $\|x\| = 1$, Lemma 14 yields

$$\inf_{\|x\|=1} \ell(\{y : |\phi^T(y)x| > \delta\} \cap \mathcal{S}) > 0,$$

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

$$\begin{aligned} & P(y_i \in U_x|\mathcal{F}_{i-1}^y) \\ =\; & \frac{1}{\sqrt{2\pi}} \int_{s\sigma_{i-1} \in U_x} e^{-\frac{s^2}{2}}\, ds \\ =\; & \frac{1}{\sqrt{2\pi}} \int_{s\sigma_{i-1} \in U_x, |s| \le \sigma_w^{-1}R} e^{-\frac{s^2}{2}}\, ds \\ \ge\; & \frac{1}{\sqrt{2\pi}} \cdot \frac{\ell(U_x)}{\sigma_{i-1}} \cdot e^{-\frac{(\sigma_w^{-1}R)^2}{2}} \\ \ge\; & \frac{1}{\sigma_{i-1}} \cdot \frac{e^{-\frac{(\sigma_w^{-1}R)^2}{2}}}{\sqrt{2\pi}} \cdot \inf_{\|x\|=1} \ell(\{y : |\phi^T(y)x| > \delta\} \cap \mathcal{S}) \\ \triangleq\; & \frac{k_1}{\sigma_{i-1}}, \end{aligned} \tag{106}$$

where $R = \mathrm{dist}(\mathcal{S})$.

For any $\epsilon > 0$, recalling the definition of g_ϵ in Lemma 20, the strong law of large numbers for martingale differences shows that all $g_\epsilon \in \mathcal{G}_\epsilon$ fulfill

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

Since there are only finite U_ϵ satisfying $U_\epsilon \subset \mathcal{S}$, we conlude

$$\lim_{t\to\infty} \inf_{U_\epsilon \subset \mathcal{S}} \frac{1}{t} \sum_{i=1}^{t} g_\epsilon(i) = -\epsilon, \quad \text{a.s.,}$$

which, together with Lemma 20, yields

$$\begin{aligned} & \liminf_{t\to+\infty} \inf_{\|x\|=1} \frac{1}{t} \sum_{i=1}^{t} g_x(i) \\ \ge\; & \liminf_{t\to+\infty} \inf_{\|x\|=1} \frac{1}{t} \sum_{i=1}^{t} g_\epsilon^x(i) \\ \ge\; & \liminf_{t\to\infty} \inf_{U_\epsilon \subset \mathcal{S}} \frac{1}{t} \sum_{i=1}^{t} g_\epsilon(i) = -\epsilon, \quad \text{a.s..} \end{aligned}$$

Furthermore, by the arbitrariness of ϵ, we obtain

$$\liminf_{t\to+\infty} \inf_{\|x\|=1} \frac{1}{t} \sum_{i=1}^{t} g_x(i) \ge 0 \quad \text{a.s..} \tag{107}$$

Combining (106) and (107), for any given $\varepsilon > 0$, there exists a random integer $T > 0$ such that for all $t > T$,

$$\begin{aligned} \frac{1}{t} \sum_{i=1}^{t} I_{\{y_i \in U_x\}} \;&>\; \frac{1}{t} \sum_{i=1}^{t} P(y_i \in U_x|\mathcal{F}_{i-1}^y) - \frac{\sigma_w^2}{\delta^2} \cdot \varepsilon \\ &\ge\; \frac{1}{t} \sum_{i=1}^{t} \frac{k_1}{\sigma_{i-1}} - \frac{\sigma_w^2}{\delta^2} \cdot \varepsilon. \end{aligned}$$

Then

$$\begin{aligned} \lambda_{\min}(t+1) \;&=\; \inf_{\|x\|=1} x^T \left(I_n + \frac{1}{\sigma_w^2} \sum_{i=0}^{t} \phi_i \phi_i^T\right) x \\ &\ge\; \frac{1}{\sigma_w^2} \sum_{i=1}^{t} (\phi^T(y_i)x)^2 \\ &\ge\; \frac{\delta^2}{\sigma_w^2} \cdot \left(\sum_{i=1}^{t} \frac{k_1}{\sigma_{i-1}} - \frac{\sigma_w^2}{\delta^2} \cdot \varepsilon t\right) \\ &=\; \frac{\delta^2}{\sigma_w^2} \cdot k_1 \sum_{i=1}^{t} \frac{1}{\sigma_{i-1}} - \varepsilon t. \end{aligned}$$

Proposition 1 follows by letting $M = \frac{\delta^2}{\sigma_w^2} \cdot k_1$. ■

References

[1] K. J. Åström and B. Wittenmark. *Adaptive Control*. Addison-Wesley: Reading, MA, 1995.

[2] H. Chen and L. Guo. *Identification and Stochastic Adaptive Control*. Birkhauser: Boston, MA, 1991.

[3] G. C. Goodwin. *Adaptive Filtering, Prediction and Control*. Prentice-Hall, Englewood Cliffs, NJ, 1984.

[4] L. Guo. Convergence and logarithm laws of self-tuning regulators. *Automatica*, 31(3):435–450, 1995.

[5] L. Guo. On critical stability of discrete-time adaptive nonlinear control. *IEEE Transactions on Automatic Control*, 42(11):1488–1499, 1997.

[6] P. V. Kokotović and M. Arcak. Constructive nonlinear control: Progress in the 90s. *in Proc. 14th IFAC World Congr.*, pages 49–77, 1999.

[7] M. Krstić, I. Kanellakopoulos, and P. V. Kokotović. *Nonlinear and Adaptive Control Design*. Wiley, New York, 1995.

[8] T. L. Lai and C. Z. Wei. Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. *Annals of Statistics*, 10(1):154–166, 1982.

[9] C. Li and L. Guo. A new critical theorem for adaptive nonlinear stabilization. *Automatica*, 46(6):999–1007, 2010.

[10] C. Li and J. Lam. Stabilization of discrete-time nonlinear uncertain systems by feedback based on LS algorithm. *SIAM Journal on Control and Optimization*, 51(2):1128–1151, 2013.

[11] C. Li, L.-L. Xie, and L. Guo. A polynomial criterion for adaptive stabilizability of discrete-time nonlinear systems. *Communications in Information and Systems*, 6(4):273–298, 2006.

[12] Z. Liu and C. Li. Is it possible to stabilize discrete-time parameterized uncertain systems growing exponentially fast? *SIAM Journal on Control and Optimization*, 57(3):1965–1984, 2019.

[13] Z. Liu and C. Li. Stabilizability theorem of discrete-time nonlinear systems with scalar parameters. *Control Theory and Applications*, 36(11):1929–1935, 2019.

[14] H. B. Ma. Further results on limitations to the capability of feedback. *International Journal of Control*, 81(1):21–42, 2008.

[15] V. F. Sokolov. Adaptive stabilization of parameter-affine minimum-phase plants under lipschitz uncertainty. *Automatica*, 73:64–70, 2016.

[16] J. Sternby. On consistency for the method of least squares using martingale theory. *IEEE Transactions on Automatic Control*, 22(3):346–352, 1977.

[17] G. Tao and P. V. Kokotovic. *Adaptive Control of Systems with Actuator and Sensor Nonlinearities*. John Wiley and Sons Inc., New Jersey, 1996.

[18] L.-L. Xie and L. Guo. Fundamental limitations of discrete-time adaptive nonlinear control. *IEEE Transactions on Automatic Control*, 44(9):1777–1782, 1999.

[19] 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*, pages 49–90, 2000.

[20] L.-L. Xie and L. Guo. How much uncertainty can be dealt with by feedback? *IEEE Transactions on Automatic Control*, 45(12):2203–2217, 2000.

[21] S. Xu and C. Li. On instability of LS-based self-tuning systems with bounded disturbances. *Systems and Control Letters*, 129:51–55, 2019.

[22] Y. X. Zhang and L. Guo. A limit to the capability of feedback. *IEEE Transactions on Automatic Control*, 47(4):687–692, 2002.

PLACE PHOTO HERE

Zhaobo Liu received his B.S. degree in mathematics and applied mathematics from Peking University in 2015, and the Ph.D. degree in control theory from Academy of Mathematics and Systems Science, Chinese Academy of Sciences, in 2020. He is now a Postdoctoral Fellow in College of Computer Science and Software Engineering of Shenzhen University, Shenzhen, China. His current research interests include adaptive nonlinear control, system identification and evolutionary computation. He was the recipient of IEEE CSS Beijing Chapter Young Author Prize Nomination in 2021.

PLACE PHOTO HERE

Chanying Li received the B.S. degree in Mathematics from Sichuan University in 2002, and the M.S. and Ph.D. degrees in control theory from Academy of Mathematics and Systems Science, Chinese Academy of Sciences, in 2005 and 2008, respectively. After receiving the Ph.D. degree, she held postdoctoral positions at the Wayne State University and the University of Hong Kong. In November 2011, she joined the Faculty of the Institute of Systems Science at Chinese Academy of Sciences, where she is currently a professor. Her research interests include the maximum capability of feedback, adaptive nonlinear control and system identification.

Prof. Li has received the National Science Fund for Distinguished Young Scholars and the National Science Fund for Excellent Young Scholars, respectively. She was also the recipient of the Young Scholar Prize by China Society for Industrial and Applied Mathematics in 2018 and Guan Zhao-Zhi Award of the 33th Chinese Control Conference in 2014. She is currently an Associate Editor of *Automatica* and *Mathematical Control and Related Fields*.