SCIENCE CHINA
Information Sciences

• LETTER •

June 2023, Vol. 66 169204:1–169204:2
https://doi.org/10.1007/s11432-021-3309-7

Asymptotic behavior of least squares estimators for nonlinear autoregressive models

Zhaobo LIU[1] & Chanying LI[2,3*]

[1]College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China;
[2]Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;
[3]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

Received 27 January 2021/Revised 11 May 2021/Accepted 30 June 2021/Published online 17 November 2022

Citation Liu Z B, Li C Y. Asymptotic behavior of least squares estimators for nonlinear autoregressive models. Sci China Inf Sci, 2023, 66(6): 169204, https://doi.org/10.1007/s11432-021-3309-7

Dear editor,

Estimations of nonlinear autoregressive (AR) models in the literature typically involve ergodic series. Based on this assumption, the asymptotic theory has been established accordingly (see [1–3]). However, this good property is not always true [4]. For example, we consider

$$y_{t+1} = \theta^\tau \phi(y_t, \dots, y_{t-n+1}) + w_{t+1}, \quad t \geqslant 0, \tag{1}$$

where θ is the $m \times 1$ unknown parameter vector and y_t and w_t are the scalar observations and random noise signals, respectively. Moreover, $\phi : \mathbb{R}^n \to \mathbb{R}^m$ is a known Lebesgue measurable vector function:

$$\phi(z_1, \dots, z_n) = \mathrm{col}\{\phi^{(1)}(z_1), \dots, \phi^{(n)}(z_n)\}, \tag{2}$$

where $\phi^{(i)} = (f_{i1}, \dots, f_{im_i})^\tau : \mathbb{R} \to \mathbb{R}^{m_i}, i = 1, \dots, n$ are certain known Lebesgue measurable vector functions and $m_i \geqslant 1$ are n integers satisfying $\sum_{i=1}^n m_i = m$. Without loss of generality, let $y_t = 0$ for $t < 0$. Clearly, most functions ϕ produce non-ergodic sequences $\{y_t\}$. Therefore, in this article, parameter θ in model (1), whose outputs are not necessarily ergodic, is identified.

Least squares (LS) estimators are known as the most efficient algorithms in parameter estimation. Its strong consistency for model (1) depends crucially on matrix $P_{t+1}^{-1} = I_m + \sum_{i=0}^t \phi_t \phi_t^\tau$, where $\phi_t = \phi(y_t, \dots, y_{t-n+1})$. Let $\lambda_{\min}(t+1)$ and $\lambda_{\max}(t+1)$ denote the minimal and maximal eigenvalues of P_{t+1}^{-1}, respectively. In the Gaussian case, Ref. [5] showed

$$\left\{\lim_{t\to+\infty} \lambda_{\min}(t+1) = +\infty\right\} = \left\{\lim_{t\to+\infty} \hat{\theta}_t = \theta\right\}, \tag{3}$$

whereas Ref. [6, Theorem 1] found that when $\{w_t\}$ is an appropriate martingale difference sequence,

$$\|\hat{\theta}_{t+1} - \theta\|^2 = O\left(\frac{\log(\lambda_{\max}(t+1))}{\lambda_{\min}(t+1)}\right), \quad \text{a.s.} \tag{4}$$

Moreover, Ref. [6] pointed out that $\log(\lambda_{\max}(t+1)) = o(\lambda_{\min}(t+1))$ is in some sense the weakest condition for the strong consistency of $\hat{\theta}_t$ in general.

The eigenvalues of P_{t+1}^{-1} depend on outputs $\{y_t\}$, which are produced automatically by the nonlinear random system (1). Thus, checking $\lim_{t\to+\infty} \lambda_{\min}(t+1) = +\infty$ or $\log(\lambda_{\max}(t+1)) = o(\lambda_{\min}(t+1))$ is nontrivial. However, Ref. [7] successfully verified $\liminf_{t\to+\infty} t^{-1}\lambda_{\min}(t+1) > 0$, a.s. for the linear AR model,

$$y_{t+1} = \theta_1 y_t + \theta_2 y_{t-1} + \cdots + \theta_n y_{t-n+1} + w_{t+1}, \ t \geqslant 0, \tag{5}$$

which is a special case of (1). Then, Ref. [7] completely solved the strong consistency of the LS estimator for this basic situation. The proof in [7] attributes to the linear structure of the AR model (5) to a certain extent. Regarding the nonlinear model (1), we aim to identify whether the LS estimator retains a similar asymptotic behavior.

This study establishes the asymptotic properties of the LS estimator for model (1). By assuming mild conditions on ϕ, the minimal eigenvalue of P_{t+1}^{-1} is estimated. We find that the LS estimates converge to the true parameter almost surely on the set where vectors $(y_t, \dots, y_{t-n+1})^\tau$ do not diverge to infinity. This means the LS estimator is highly likely to have a strong consistency when applied to model (1) in practice because most real systems are non-divergent. Finally, Appendix A provides the proofs of the main theorems.

Gaussian case. We assume the following conditions.

(A1) Noise $\{w_t\}$ is an i.i.d. random sequence with $w_1 \sim N(0, 1)$, and parameter $\theta \sim N(\theta_0, I_m)$ is independent of $\{w_t\}$.

(A2) Certain open sets $\{E_i\}_{i=1}^n$ belong to $\mathbb{R}$ satisfying the following:

(i) $f_{ij} \in C(\mathbb{R})$ and $f_{ij} \in C^{m_i}(E_i)$, $1 \leqslant j \leqslant m_i$, $1 \leqslant i \leqslant n$;

(ii) For every unit vector $x \in \mathbb{R}^m$, a point $y \in \prod_{i=1}^n E_i$ exists such that $|\phi^\tau(y)x| \neq 0$.

Remark 1. By assumption (A2)(ii), for every unit vector $x \in \mathbb{R}^m$, $\ell(\{y \in \prod_{i=1}^n E_i : |\phi^\tau(y)x| > 0\}) > 0$, where ℓ denotes the Lebesgue measure.

* Corresponding author (email: cyli@amss.ac.cn)

The LS estimate $\hat{\theta}_t$ for parameter θ can be recursively defined by

$$\begin{cases} \hat{\theta}_{t+1} = \hat{\theta}_t + P_{t+1}\phi_t(y_{t+1} - \phi_t^\tau \hat{\theta}_t), \\ P_{t+1} = P_t - (1 + \phi_t^\tau P_t \phi_t)^{-1} P_t \phi_t \phi_t^\tau P_t, \quad P_0 = I_m, \end{cases} \tag{6}$$

where $\hat{\theta}_0$ is the deterministic initial condition and ϕ_0 is the random initial vector of model (1). We provide a simple way to estimate the minimal eigenvalue of P_{t+1}^{-1}. Let $N_t(M) \triangleq \sum_{i=1}^t I_{\{\|Y_i\| \leqslant M\}}$, where $Y_t \triangleq (y_{t+n-1}, \ldots, y_t)^\tau$ and $M > 0$ is a constant. Then, in terms of $N_t(M)$, our estimate of $\lambda_{\min}(t+1)$ is readily available by the following theorem.

Theorem 1. Under assumptions (A1) and (A2), for any constant $M > 0$,

$$\liminf_{t\to+\infty} \frac{\lambda_{\min}(t+1)}{N_t(M)} > 0 \quad \text{a.s. on } \Omega(M), \tag{7}$$

where $\Omega(M) \triangleq \{\lim_{t\to+\infty} N_t(M) = +\infty\}$.

Corollary 1. Let assumptions (A1) and (A2) hold. Then,

$$\lim_{t\to+\infty} \hat{\theta}_t = \theta \quad \text{a.s. on } \left\{ \liminf_{t\to+\infty} \|Y_t\| < +\infty \right\}. \tag{8}$$

Remark 2. For a typical case where $\prod_{i=1}^n E_i = \mathbb{R}^n$, if assumption (A2)(ii) fails, then $\ell(\{y \in \mathbb{R}^n : |\phi^\tau(y)x| > 0\}) = 0$ for some unit vector $x \in \mathbb{R}^m$. Therefore, as $t \to +\infty$, $\lambda_{\min}(t+1) = O(1)$, a.s. In view of (3), $\hat{\theta}_t$ cannot converge to the true parameter θ. So, assumption (A2)(ii) is necessary for the strong consistency of the LS estimates $\{\hat{\theta}_t\}_{t\geqslant 0}$.

Constant parameter. Assume θ is a non-random parameter, and

(A1$'$) $\{w_t\}$ is an i.i.d random sequence with $Ew_1 = 0$ and $E|w_1|^\beta < +\infty$ for some $\beta > 2$. Moreover, w_1 has a density $\rho(x)$ such that for every proper interval $I \subset \mathbb{R}$, $\inf_{x\in I} \rho(x) > 0$ and $\sup_{x\in\mathbb{R}} \rho(x) < +\infty$.

In this case, the LS estimator is constructed from partial data. More specifically, for some constant $C_\phi > 0$, ϕ_t is modified as $\phi_t \triangleq I_{\{\|Y_{t-n+1}\| \leqslant C_\phi\}} \phi(y_t, \ldots, y_{t-n+1})$. Then, an analogous version of Theorem 1 is deduced as follows.

Theorem 2. Under assumptions (A1$'$) and (A2), a constant $M_\phi > 0$ exists depending only on ϕ such that for any $C_\phi > M_\phi$ and $M > 0$,

$$\liminf_{t\to+\infty} \frac{\lambda_{\min}(t+1)}{N_t(M)} > 0 \quad \text{a.s. on } \Omega(M).$$

Furthermore, if $M \geqslant C_\phi$, then

$$\|\hat{\theta}_t - \theta\|^2 = O\left(\frac{\log N_t(M)}{N_t(M)} \right) \quad \text{a.s. on } \Omega(M).$$

Remark 3. Furthermore, we conclude (8) in Theorem 2. For most practical situations, the systems frequently fulfill

$$P\left\{ \liminf_{t\to+\infty} \|Y_t\| < +\infty \right\} = 1, \tag{9}$$

and the strong consistency of the LS estimates is thus guaranteed. Observe that assumption (A1$'$) and Eq. (9) imply that $\{y_t\}_{t\geqslant 1}$ in model (1) is an aperiodic Harris recurrent Markov chain and hence admits an invariant measure. Certain integrability assumptions on the invariant measure may lead to the consistency of the LS estimates (e.g., [8]). However, judging these integrability assumptions for a nonlinear AR model is generally nontrivial.

When $n = 1$ in model (1), assumption (A2) in Theorems 1 and 2 can be relaxed as below.

(A2$'$) $f_{1i} \in C^{m_1}(E_1), i = 1, \ldots, m_1$ are linearly independent on an open set $E_1 \in \mathbb{R}$ and ϕ is bounded in every compact set.

Example 1. Consider a parametric AR model:

$$y_{t+1} = \sum_{j=1}^m \theta_j g(y_t) I_{\{y_t \in D_j\}} + y_t I_{\{y_t \in D_{m+1}\}} + w_{t+1}, \tag{10}$$

where $y_0 = 0$, $g(\cdot)$ is bounded in any compact set, $\{D_j\}_{j=1}^m$ are some compact subsets of $\mathbb{R}$ with positive Lebesgue measure, and $D_{m+1} = (\bigcup_{j=1}^m D_i)^c$. Let noises $\{w_t\}_{t\geqslant 1}$ satisfy assumption (A1$'$) and unknown parameters $\theta_1, \ldots, \theta_m \in \mathbb{R}$. As can be seen, $\{y_t\}_{t\geqslant 1}$ must fall into $\bigcup_{j=1}^m D_i$ infinitely many times and Eq. (9) holds. Hence Theorems 1 and 2 can be applied and the strong consistency of the LS estimates is established. If $g(x) = x$, model (10) becomes a non-ergodic threshold autoregressive (TAR) model.

Conclusion. This study discusses the LS estimator for a basic class of nonlinear AR models, whose outputs are not necessarily ergodic. Several asymptotic properties of the LS estimator have been established under mild conditions. These properties suggest the strong consistency of the LS estimates in non-divergent nonlinear AR models.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 11925109) and National Key R&D Program of China (Grant No. 2018YFA0703800).

Supporting information Appendix A. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Chan K S. Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. Ann Statist, 1993, 21: 520–533
2 Li D, Ling S. On the least squares estimation of multiple-regime threshold autoregressive models. J Econom, 2012, 167: 240–253
3 Zhao W-X, Chen H-F, Zheng W X. Recursive identification for nonlinear ARX systems based on stochastic approximation algorithm. IEEE Trans Automat Contr, 2010, 55: 1287–1299
4 Li C, Lam J. Stabilization of discrete-time nonlinear uncertain systems by feedback based on LS algorithm. SIAM J Control Optim, 2013, 51: 1128–1151
5 Sternby J. On consistency for the method of least squares using martingale theory. IEEE Trans Automat Contr, 1977, 22: 346–352
6 Lai T L, Wei C Z. Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. Ann Statist, 1982, 10: 154–166
7 Lai T L, Wei C Z. Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters. J Multivariate Anal, 1983, 13: 1–23
8 Li D, Tjøstheim D, Gao J. Estimation in nonlinear regression with Harris recurrent Markov chains. Ann Statist, 2016, 44: 1957–1987

SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Asymptotic Behavior of Least Squares Estimator for Nonlinear Autoregressive Models

Zhaobo Liu[1] & Chanying Li[2,3*]

[1]*College of Computer Science and Software Engineering of Shenzhen University, Shenzhen* 518060*, P. R. China;*
[2]*Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing* 100190*, P. R. China;*
[3]*School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing* 100049*, P. R. China*

Appendix A Proofs of Theorems 1–2

It is obvious that to show Theorems 1–2, it suffices to prove

Proposition 1. Under Assumptions A1' and A2, let θ be a random variable independent of $\{w_t\}_{t\geqslant 1}$.Then, there is a constant $M_\phi > 0$ depending only on ϕ such that for any $C_\phi > M_\phi$ and $M, K > 0$,

$$\liminf_{t\to+\infty} \frac{\lambda_{\min}(t+1)}{N_t(M)} > 0 \quad \text{a.s. on } \Omega(M)\cap\{\|\theta\|\leqslant K\}. \tag{A1}$$

Appendix A.1 Proof of Proposition 1

Following the idea of [2], for every $x\in\mathbb{R}^m$ with $\|x\|=1$, we construct a set $\mathcal{S}\triangleq\prod_{i=1}^n\bigcup_{j=1}^{p_i}S_i^j(q)$ with disjoint open intervals $\{S_i^j(q): j=1,\ldots,p_i\}$ such that

$$\ell\left(\{y\in\mathcal{S}: |\phi^\tau(y)x|>0\}\right)>0 \quad \text{and} \quad \overline{\mathcal{S}}\subset\prod_{i=1}^n E_i. \tag{A2}$$

Define

$$U_x(\delta)\triangleq\{y: |\phi^\tau(y)x|>\delta\}\cap\mathcal{S}, \quad \delta>0. \tag{A3}$$

Next, let $\{d_k\}_{k=1}^{2n}$ be a sequence of numbers and for $k\in[n+1,2n]$ define

$$\varsigma_k\triangleq d_k - x^\tau\phi(d_{k-1},\ldots,d_{k-n}), \quad x\in\mathbb{R}^m. \tag{A4}$$

Denote $y=(d_n,\ldots,d_1)^\tau$ and $\varsigma=(\varsigma_{2n},\ldots,\varsigma_{n+1})^\tau$. Evidently, (A4) implies that there is a function $g:\mathbb{R}^{2n+m}\to\mathbb{R}^n$ such that

$$(d_{2n},\ldots,d_{n+1})^\tau = g(\varsigma,y,x). \tag{A5}$$

We take δ in (A3) according to the following lemma.

Lemma 1. Under Assumption A2, the following two statements hold:
(i) given $y\in\mathbb{R}^n$, $x\in\mathbb{R}^m$ and a box $O=\prod_{i=1}^n I_i$ with $\{I_i\}_{i=1}^n$ being some intervals, then

$$\ell(\{\varsigma: g(\varsigma,y,x)\in O\})=\ell(O); \tag{A6}$$

(ii) for any constants $M,K>0$, there is a $\delta^*>0$ such that $\inf_{\|z\|=1,\|y\|\leqslant M,\|x\|\leqslant K}\ell(\{\varsigma: |\phi^\tau(g(\varsigma,y,x))z|>\delta^*, g(\varsigma,y,x)\in\mathcal{S}\})>0$.

Proof. (i) Note that in view of (A4), $d_k=\varsigma_k+o_{k-1}$, $k=n+1,\ldots,2n$, where $o_{k-1}\in\mathbb{R}$ is a point determined by ς_{k-1}, y and x (for $k=n+1$, ς_n does not exist and o_n depends only on y and x). So, $\{\varsigma:\varsigma+o_{k-1}\in I_k\}=I_k-o_{k-1}$ is an interval with length $|I_k|$. By the definition of the Lebesgue measure in $\mathbb{R}^n$, it is straightforward that $\ell(\{\varsigma: g(\varsigma,y,x)\in O\})=\prod_{k=1}^n|I_k|=\ell(O)$.
(ii) Suppose (ii) is false. Then for each integer $k\geqslant 1$, we can take some $(z(k),y(k),x(k))$ with $\|z(k)\|=1$ in $\overline{B(0,1)}\times\overline{B(0,M)}\times\overline{B(0,K)}\subset\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^m$ such that

$$\ell(\{\varsigma: |\phi^\tau(g(\varsigma,y(k),x(k)))z(k)|>\frac{1}{k}, g(\varsigma,y(k),x(k))\in\mathcal{S}\})<\frac{1}{k}. \tag{A7}$$

Hence there is a subsequence $\{z(k_r),y(k_r),x(k_r)\}_{r\geqslant 1}$ and an accumulation point (z^*,y^*,x^*) satisfying

$$\lim_{r\to+\infty}(x(k_r),y(k_r),z(k_r))=(x^*,y^*,z^*). \tag{A8}$$

* Corresponding author (email: cyli@amss.ac.cn)

Clearly, $\|z^*\| = 1$, $\|y^*\| \leqslant M$, $\|x^*\| \leqslant K$. If $\ell(\{\varsigma : |\phi^\tau(g(\varsigma, y^*, x^*))z^*| > 0, g(\varsigma, y^*, x^*) \in \mathcal{S}\}) = 0$, then $\phi^\tau(y)z^* \equiv 0$ for all $y \in \mathcal{S}$ due to (A4), (A5) and the continuity of ϕ. It contradicts to (A2). Consequently, for any $\{\mathcal{S}_k\}_{k\geqslant 1}$ satisfying $\mathcal{S}_k \subset \mathcal{S}_{k+1}$ and $\lim_{k\to+\infty} \mathcal{S}_k = \mathcal{S}$, we have

$$\begin{aligned} &\lim_{k\to+\infty} \ell(\{\varsigma : |\phi^\tau(g(\varsigma, y^*, x^*))z^*| > \frac{1}{k}, g(\varsigma, y^*, x^*) \in \mathcal{S}_k\}) \\ =\ &\ell(\{\varsigma : |\phi^\tau(g(\varsigma, y^*, x^*))z^*| > 0, g(\varsigma, y^*, x^*) \in \mathcal{S}\}) > 0. \end{aligned}$$

Therefore, for some $h \geqslant 1$,

$$\ell(\{\varsigma : |\phi^\tau(g(\varsigma, y^*, x^*))z^*| > \frac{1}{h}, g(\varsigma, y^*, x^*) \in \mathcal{S}_h\}) > 0. \tag{A9}$$

Note that all points $\{y(k_r), x(k_r)\}_{r\geqslant 1}$ are restricted to $\overline{B(0, M)} \times \overline{B(0, K)}$, (A4) and (A5) then indicate that there is a compact set O' such that $\{\varsigma : g(\varsigma, y(k_r), x(k_r)) \in \mathcal{S}\} \subset O'$. Further, g and ϕ are continuous due to (A4), (A5) and Assumption A2(i), hence (A8) shows $\lim_{r\to\infty} \sup_{\varsigma\in O'} \|g(\varsigma, y^*, x^*) - g(\varsigma, y(k_r), x(k_r))\| = 0$ and $\lim_{r\to\infty} \sup_{\varsigma\in O'} \|\phi^\tau(g(\varsigma, y^*, x))z^* - \phi^\tau(g(\varsigma, y(k_r), x(k_r)))z(k_r)\| = 0$.

As a consequence, for all sufficiently large r,

$$\begin{aligned} &\ell(\{\varsigma : |\phi^\tau(g(\varsigma, y^*, x^*))z^*| > \frac{1}{h}, g(\varsigma, y^*, x^*) \in \mathcal{S}_h\}) \\ <\ &\ell(\{\varsigma : |\phi^\tau(g(\varsigma, y(k_r), x(k_r)))z(k_r)| > \frac{1}{k_r}, g(\varsigma, y(k_r), x(k_r)) \in \mathcal{S}\}) < \frac{1}{k_r}, \end{aligned}$$

which contradicts to (A9) by letting $r \to +\infty$. Lemma 1 follows. □

Remark 1. In Lemma 1, Assumption A2 can be weaken to Assumption A2' when $n = 1$. Statement (i) is trivial. For (ii), note that (A2) still holds by Assumption A2'. But, (A4), (A7) and (A9) yield that for all sufficiently large r,

$$\begin{aligned} \frac{1}{k_r}\ &>\ \ell(\{\varsigma : |\phi^\tau(g(\varsigma, y(k_r), x(k_r)))z(k_r)| > \frac{1}{k_r}, g(\varsigma, y(k_r), x(k_r)) \in \mathcal{S}\}) \\ &=\ \ell(\{y : |\phi^\tau(y)z(k_r)| > \frac{1}{k_r}, y \in \mathcal{S}\}) \geqslant \ell(\{y \in \mathcal{S} : |\phi^\tau(y)z^*| > \frac{1}{k_r} + \frac{1}{h}\}), \end{aligned}$$

where $\{z(k_r), y(k_r), x(k_r)\}_{r\geqslant 1}$ is defined in the proof of Lemma 1. Letting $r \to +\infty$ in the above inequality infers

$$0\ \geqslant\ \lim_{r\to+\infty} \ell(\{y \in \mathcal{S} : |\phi^\tau(y)z^*| > \frac{1}{k_r} + \frac{1}{h}\}) = \ell(\{y \in \mathcal{S} : |\phi^\tau(y)z^*| > \frac{1}{h}\}),$$

which contradicts to (A9).

Fix two positve numbers M and K and let δ^* be constructed in Lemma 1(ii). Now, for every unit vector $x \in \mathbb{R}^m$, define $U_x \triangleq U_x(\delta^*)$.

For the next lemma, fix a closed box $O = \prod_{i=1}^n I_i \in \mathbb{R}^n$ and a positive integer r. Equally divide each I_i into r closed intervals $\{I_{i,j}\}_{j=1}^r$ so that $\mathrm{int}(I_{i,j}) \cap \mathrm{int}(I_{i,j'}) = \emptyset$ if $j \neq j'$. We thus obtain r^n small closed boxes $\prod_{i=1}^n \{I_{i,j}\}_{j=1}^r$, which are denoted by $\mathcal{T}(O, r)$. It is easy to see that for any distinct boxes $U, U' \in \mathcal{T}(O, r)$, $\mathrm{int}(U) \cap \mathrm{int}(U') = \emptyset$. Define

$$\mathcal{T}_\delta(O, r) \triangleq \left\{U \in \mathcal{T}(O, r) : \mathcal{B}(\delta) \cap \overline{\mathcal{S}} \cap U \neq \emptyset\right\}, \tag{A10}$$

where $\mathcal{B}(\delta) \triangleq \partial(\{y : \phi^\tau(y)x > \delta\})$. Let $\mathcal{K}_\delta(O, x, r) \triangleq |\mathcal{T}_\delta(O, r)|$.

Lemma 2. There is a constant $C > 0$ such that for any closed box $O = \prod_{i=1}^n I_i$, non-zero vector $x \in \mathbb{R}^m$, $\delta \in \mathbb{R}$ and integer $r \geqslant 1$,

$$\mathcal{K}_\delta(O, x, r) \leqslant Cr^{n-1}. \tag{A11}$$

Proof. Denote $A(g) \triangleq \{x : g(x) = 0\}$ for function g. For $i \in [1, n]$, let $(\phi^{(i)})' = (f'_{i1}, \dots, f'_{im_i})^\tau$ and

$$\begin{cases} K_i = \mathrm{int}(A(x_i^\tau(\phi^{(i)})')) \cap \left(\bigcup_{j=1}^{p_i} \overline{S_i^j(q)}\right) \\ L_i = (A(x_i^\tau(\phi^{(i)})') \cap \left(\bigcup_{j=1}^{p_i} \overline{S_i^j(q)}\right)) \setminus K_i \end{cases}. \tag{A12}$$

We prove (A11) by induction. For $n = 1$, let $O = I_1$ be a closed box. By [2, Lemma B.10], it is easy to check that

$$\left|\mathcal{B}(\delta) \cap \bigcup_{j=1}^{p_1} S_1^j(q)\right| \leqslant 2p_1(|L_1| + 2) < +\infty. \tag{A13}$$

Moreover, since $\mathcal{B}(\delta) \cap (\bigcup_{j=1}^{p_1} \overline{S_1^j(q)}) \subset \mathcal{B}(\delta) \cap (\bigcup_{j=1}^{p_1} S_1^j(q)) \cup \partial(\bigcup_{j=1}^{p_1} S_1^j(q))$, it follows that $\mathcal{K}_\delta(O, x, r) \leqslant 2|\mathcal{B}(\delta) \cap (\bigcup_{j=1}^{p_1} \overline{S_1^j(q)})| \leqslant 4p_1(|L_1| + 2) + 4p_1$. Hence, (A11) is true for $n = 1$ by taking $C = 4p_1(|L_1| + 2) + 4p_1$.

Now, suppose (A11) holds for $n = k$ with some $k \geqslant 1$. Let us consider the case where $n = k + 1$. Take a closed box $O = \prod_{i=1}^{k+1} I_i \in \mathbb{R}^{k+1}$, and let $\mathcal{T}(O, r)$ be the set of the r^{k+1} disjoint refined boxes. These boxes correspond to two sets

$$\mathcal{T}^1 = \prod_{i=1}^{k} \{I_{i,j}\}_{j=1}^r \quad \text{and} \quad \mathcal{T}^2 = \{I_{k+1,j}\}_{j=1}^r.$$

Write vector $x = \mathrm{col}\{x_1, \ldots, x_{k+1}\} \neq \mathbf{0}$. First, assume there is an index $l \in [1, k+1]$ such that $x_l = \mathbf{0}$. Without loss of generality, let $l = k+1$, then

$$\begin{aligned} &\mathcal{B}(\delta) \cap \prod_{i=1}^{k+1} \bigcup_{j=1}^{p_i} \overline{S_i^j(q)} \cap O \\ \subset\ & \left(\partial\left(\left\{z \in \mathbb{R}^k : \sum_{i=1}^{k} x_i \phi^{(i)}(z_i) > \delta\right\}\right) \cap \prod_{i=1}^{k} \bigcup_{j=1}^{p_i} \overline{S_i^j(q)} \cap \prod_{i=1}^{k} I_i\right) \times I_{k+1}. \end{aligned} \tag{A14}$$

where $z = (z_1, \ldots, z_k)^\tau \in \mathbb{R}^k$. By applying the induction assumption for $n = k$ and for the refined boxes in $\mathcal{T}^1$, there is a constant $C > 0$ such that $\mathcal{K}_\delta\left(\prod_{i=1}^k I_i, \mathrm{col}\{x_1, \ldots, x_k\}, r\right) \leqslant Cr^{k-1}$, which, together with (A14) and $\mathcal{T}(O, a) = \mathcal{T}^1 \times \mathcal{T}^2$, yields $\mathcal{K}_\delta(O, x, r) \leqslant Cr^k$. This is exactly (A11) for $n = k+1$.

So, let $x_i \neq \mathbf{0}$ for all $i \in [1, k+1]$. For any $B \in \mathcal{T}^1$, define set

$$Z(B) \triangleq \{z_{k+1} \in I_{k+1} : (B \times z_{k+1}) \cap \mathcal{B}(\delta) \cap \prod_{i=1}^{k+1} \bigcup_{j=1}^{p_i} \overline{S_i^j(q)} \neq \emptyset\}.$$

Observe that $Z(B)$ is a closed set, then $\partial Z(B) \subset Z(B)$. Define

$$\begin{cases} \mathcal{Z}_1(B) \triangleq \{I_{k+1,j} \in \mathcal{T}^2 : Z(B) \cap I_{k+1,j} \neq \emptyset\} \\ \mathcal{Z}_2(B) \triangleq \{I_{k+1,j} \in \mathcal{T}^2 : \partial Z(B) \cap I_{k+1,j} \neq \emptyset\} \end{cases}.$$

Since any interval in $\mathcal{Z}_1(B) \setminus \mathcal{Z}_2(B)$ must be contained in $Z(B)$,

$$|\mathcal{Z}_1(B)| - |\mathcal{Z}_2(B)| = |\mathcal{Z}_1(B) \setminus \mathcal{Z}_2(B)| \leqslant \frac{r}{|I_{k+1}|} \ell(Z(B)).$$

At the same time, $\sum_{B \in \mathcal{T}^1} \ell(Z(B)) = \sum_{B \in \mathcal{T}^1} \int_{\mathbb{R}} I_{Z(B)} dz_{k+1} = \int_{I_{k+1}} \sum_{B \in \mathcal{T}^1} I_{Z(B)} dz_{k+1}$, therefore

$$\mathcal{K}_\delta(O, x, r) = \sum_{B \in \mathcal{T}^1} |\mathcal{Z}_1(B)| \leqslant \frac{r}{|I_{k+1}|} \int_{I_{k+1}} \sum_{B \in \mathcal{T}^1} I_{Z(B)} dz_{k+1} + \sum_{B \in \mathcal{T}^1} |\mathcal{Z}_2(B)|. \tag{A15}$$

The last step is to estimate the term in (A15). Since the argument is involved, it is included in Appendix A.2. In light of Lemmas 5 and 6, when $n = k+1$, there are two constants $C_1, C_2 > 0$ depending only on ϕ such that $\mathcal{K}_\delta(O, x, r) \leqslant (C_1 + C_2)\, r^k$. The proof is thus completed. □

By applying Lemma 2, we can find a constant $C_0 > 0$ depending only on ϕ such that

$$|\{U \in \mathcal{T}(O, r) : \partial(U_x) \cap U \neq \emptyset\}| \leqslant C_0 r^{n-1}. \tag{A16}$$

Now, we estimate the frequency of $\{Y_t\}_{t \geqslant 1}$, where $Y_i \triangleq (y_{i+n-1}, \ldots, y_i)^\tau$, falling into U_x. For this, define a random process g_x by

$$g_x(i) \triangleq I_{\{Y_i \in U_x\}} - P(Y_i \in U_x | \mathcal{F}_{i-1}^y), \quad i \geqslant 1,$$

where $\mathcal{F}_{i-1}^y \triangleq \sigma\{\theta, y_0, \ldots, y_{i-1}\}$. By modifying the proof of [2, Lemma B.12] slightly, we can obtain

Lemma 3. 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 $\{g_\epsilon(i)\}_{i \geqslant 1}$ with the form

$$g_\epsilon(i) = I_{\{Y_i \in U_\epsilon\}} - P(Y_i \in U_\epsilon | \mathcal{F}_{i-1}^y) - \epsilon, \quad i \geqslant 1, \tag{A17}$$

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

$$g_\epsilon(i) \leqslant g_x(i) \quad \forall i \geqslant 1. \tag{A18}$$

Proof of Proposition 1. First, recall the definition of U_x, for any $x \in \mathbb{R}^m$ with $\|x\| = 1$, Lemma 1(ii) and Assumption A1' yield

$$\begin{aligned} & P(Y_i \in U_x | \mathcal{F}_{i-1}^y) I_{\{\|Y_{i-n}\| \leqslant M, \|\theta\| \leqslant K\}} = P(Y_i \in \{y : |\phi^\tau(y)x| > \delta^*\} \cap \mathcal{S} | \mathcal{F}_{i-1}^y) \cdot I_{\{\|Y_{i-n}\| \leqslant M, \|\theta\| \leqslant K\}} \\ \geqslant\ & \inf_{\|x\|=1, \|y\| \leqslant M, \|z\| \leqslant K} \ell(\{\varsigma : |\phi^\tau(g(\varsigma, y, z))x| > \delta^*, g(\varsigma, y, z) \in \mathcal{S}\}) \cdot \left(\inf_{s \in [-S', S']} \rho(s)\right)^n I_{\{\|Y_{i-n}\| \leqslant M, \|\theta\| \leqslant K\}} \\ \triangleq\ & C_P I_{\{\|Y_{i-n}\| \leqslant M, \|\theta\| \leqslant K\}}, \end{aligned} \tag{A19}$$

where $S' = K \sup_{\|y\| \leqslant M + R'} \|\phi(y)\| + R'$ and $R' \triangleq \max_{1 \leqslant i \leqslant n} \mathrm{dist}\left(0, \bigcup_{j=1}^{p_i} S_i^j(q)\right)$.

Next, note that for any $\epsilon > 0$ and $g_\epsilon \in \mathcal{G}_\epsilon$, $\{g_\epsilon(i) + \epsilon, \mathcal{F}_i^y\}_{i \geqslant 1}$ is a martingale difference sequence. Taking account of [1, Theorem 2.8],

$$\lim_{t \to +\infty} \frac{\sum_{i=1}^t I_{\{\|Y_{i-n}\| \leqslant M\}} (g_\epsilon(i) + \epsilon)}{N_t(M)} = 0, \quad \text{a.s. } on \quad \Omega(M),$$

where $\Omega(M)$ is defined in Theorem 1. Thanks to the finite number of U_ϵ constrained in $\mathcal{S}$, it gives

$$\lim_{t\to+\infty}\inf_{U_\epsilon\subset\mathcal{S}}\frac{1}{N_t(M)}\sum_{i=1}^{t}I_{\{\|Y_{i-n}\|\leqslant M\}}g_\epsilon(i)=-\epsilon,\quad \text{a.s. on}\quad \Omega(M).$$

As a result, Lemma 3(ii) infers that for some $g_\epsilon^x\in\mathcal{G}_\epsilon$,

$$\begin{aligned}\liminf_{t\to+\infty}\inf_{\|x\|=1}\frac{1}{N_t(M)}\sum_{i=1}^{t}I_{\{\|Y_{i-n}\|\leqslant M\}}g_x(i)&\geqslant\liminf_{t\to+\infty}\inf_{\|x\|=1}\frac{1}{N_t(M)}\sum_{i=1}^{t}I_{\{\|Y_{i-n}\|\leqslant M\}}g_\epsilon^x(i)\\&\geqslant\liminf_{t\to\infty}\inf_{U_\epsilon\subset\mathcal{S}}\frac{1}{N_t(M)}\sum_{i=1}^{t}I_{\{\|Y_{i-n}\|\leqslant M\}}g_\epsilon(i)\\&=-\epsilon,\quad\text{a.s. on}\quad\Omega(M).\end{aligned}$$

Further, by the arbitrariness of ϵ, we obtain that on $\Omega(M)$

$$\liminf_{t\to+\infty}\inf_{\|x\|=1}\frac{1}{N_t(M)}\sum_{i=1}^{t}I_{\{\|Y_{i-n}\|\leqslant M\}}g_x(i)\geqslant0\quad\text{a.s..}\tag{A20}$$

Finally, by (A19)–(A20), for sufficiently small ϵ, there is a positive random integer T such that for any unit vector $x\in\mathbb{R}^m$ and all $t>T$, $\frac{1}{N_t(M)}\sum_{i=1}^{t}I_{\{\|Y_{i-n}\|\leqslant M\}}I_{\{Y_i\in U_x\}}>\frac{1}{N_t(M)}\sum_{i=1}^{t}I_{\{\|Y_{i-n}\|\leqslant M\}}P(Y_i\in U_x|\mathcal{F}_{i-1}^y)-\frac{C_P}{2}\geqslant\frac{C_P}{2}$, a.s. on $\Omega(M)\cap\{\|\theta\|\leqslant K\}$. Hence, we select C_ϕ satisfies $C_\phi>nR'$ and $U_x\subset\overline{B(0,C_\phi)}$, for sufficiently large t,

$$\begin{aligned}\lambda_{\min}(t+1)&=\inf_{\|x\|=1}x^\tau\left(I_m+\sum_{i=0}^{t}\phi_i\phi_i^\tau\right)x\\&\geqslant\sum_{i=1}^{t-n+1}I_{\{Y_i\in U_x\}}(\phi^\tau(Y_i)x)^2\geqslant(\delta^*)^2\sum_{i=1}^{t-n+1}I_{\{Y_i\in U_x\}}\\&\geqslant\frac{(\delta^*)^2C_P}{2}(N_t(M)-n),\quad\text{a.s. on}\quad\Omega(M)\cap\{\|\theta\|\leqslant K\}.\end{aligned}$$

Proposition 1 is thus proved. □

Appendix A.2 Proof of (A15)

In this section, we follow the definitions and symbols in the proof of Lemma 2 and complete the estimation details of (A15). To this end, define $\mathbb{S}_i\triangleq\bigcup_{j=1}^{p_i}\overline{S_i^j(q)}$, $i\in[1,n]$,

$$\begin{aligned}I_{k+1}^*&\triangleq\left\{z_{k+1}:\left(\prod_{i=1}^{k}I_i\times z_{k+1}\right)\cap\mathcal{B}(\delta)\cap\left(\prod_{i=1}^{k}K_i\times z_{k+1}\right)\neq\emptyset\right\}\\&\quad\cap I_{k+1}\cap\left(\bigcup_{j=1}^{p_{k+1}}\overline{S_{k+1}^j(q)}\right),\quad k\geqslant1\\\mathcal{T}^3&\triangleq\{A\in\mathcal{T}^2:A\cap I_{k+1}^*\neq\emptyset\},\\\mathcal{T}^4&\triangleq\left\{B\in\mathcal{T}^1:\bigcup_{i=1}^{k}\{z:z_i\in L_i\}\cap B\neq\emptyset\right\},\end{aligned}$$

where $\prod_{i=1}^{k+1}I_i=O$ is the given closed box in the proof of Lemma 2.

Lemma 4. The cardinals of $I_{k+1}^*,\mathcal{T}^3$ and $\mathcal{T}^4$ are bounded by

$$|I_{k+1}^*|\leqslant(2p_{k+1}(|L_{k+1}|+2)+2)\prod_{i=1}^{k}(|L_i|+p_i),\tag{A21}$$

$$|\mathcal{T}^3|\leqslant2(2p_{k+1}(|L_{k+1}|+2)+2)\prod_{i=1}^{k}(|L_i|+p_i),$$

$$|\mathcal{T}^4|\leqslant2r^{k-1}\sum_{i=1}^{k}|L_i|,\tag{A22}$$

Proof. By the definitions of $\mathcal{T}^3$ and $\mathcal{T}^4$, $\mathcal{T}^3\leqslant2|I_{k+1}^*|$ and (A22) is trivial. So, it suffices to show (A21). For this, recall the definitions of K_i and L_i, then for each $i\in[1,n]$, there is a set $\mathcal{P}_i$ consisting of some disjoint intervals such that $|\mathcal{P}_i|\leqslant|L_i|+p_i$ and $\bigcup_{I\in\mathcal{P}_i}I=K_i$. As a result, $|\prod_{i=1}^{k}\mathcal{P}_i|\leqslant\prod_{i=1}^{k}(|L_i|+p_i)$. For each box $B\in\prod_{i=1}^{k}\mathcal{P}_i$, denote $I_{k+1}^*(B)=\{z_{k+1}:(\prod_{i=1}^{k}I_i\times z_{k+1})\cap\mathcal{B}(\delta)\cap(B\times z_{k+1})\neq\emptyset\}\cap I_{k+1}\cap\mathbb{S}_{k+1}$. Since $B\subset\prod_{i=1}^{k}K_i$, it is evident that

$$\sum_{i=1}^{k}x_i^\tau\phi^{(i)}\equiv\text{constant}\qquad\text{on }B.\tag{A23}$$

So, for any $z_{k+1} \in I_{k+1}^*(B)$, arbitrarily taking a $(z_1,\ldots,z_k)^\tau \in \mathrm{int}(B)$ infers $(z_1,\ldots,z_{k+1})^\tau \in \mathcal{B}(\delta)$. Let $\{(z_{1,j},\ldots,z_{k+1,j})^\tau\}_{j=1}^{+\infty}$ be a sequence of points in $(\mathrm{int}(B)\times E_{k+1}) \cap \{y : \phi^\tau(y)x > \delta\}$ and tend to $(z_1,\ldots,z_{k+1})^\tau$. Then, $\lim_{j\to+\infty}\|z_{k+1,j}-z_{k+1}\| = 0$ and

$$x_{k+1}^\tau \phi^{(k+1)}(z_{k+1,j}) \ > \ \delta - \sum_{i=1}^k x_i^\tau \phi^{(i)}(z_{i,j}) = \delta - \sum_{i=1}^k x_i^\tau \phi^{(i)}(z_i). \tag{A24}$$

Denote $\bar\delta = \delta - \sum_{i=1}^k x_i^\tau \phi^{(i)}(z_i)$, so (A24) implies $z_{k+1} \in \partial(\{z : x_{k+1}^\tau \phi^{(k+1)}(z) > \bar\delta\}) \cap \mathbb{S}_{k+1}$, Therefore, applying Lemma A.3(iii), $|I_{k+1}^*(B)| \leqslant |\partial(\{z : x_{k+1}^\tau\phi^{(k+1)}(z) > \bar\delta\}) \cap \mathbb{S}_{k+1}| \leqslant 2p_{k+1}(|L_{k+1}|+2)+2$, and thus $|I_{k+1}^*| \leqslant (2p_{k+1}(|L_{k+1}|+2)+2)\left|\prod_{i=1}^k \mathcal{P}_i\right| \leqslant (2p_{k+1}(|L_{k+1}|+2)+2)\prod_{i=1}^k(|L_i|+p_i)$, which completes the proof. □

Lemma 5. Let Lemma 2 hold with $n = k$. Then, there is a constant $C_1 > 0$ depending only on ϕ such that

$$\frac{r}{|I_{k+1}|}\int_{I_{k+1}} \sum_{B\in\mathcal{T}^1} I_{\mathcal{Z}(B)} dz_{k+1} \leqslant C_1 r^k. \tag{A25}$$

Proof. Denote $\phi' = \mathrm{col}\{\phi^{(1)},\ldots,\phi^{(k)}\}$, $x' = \mathrm{col}\{x_1,\ldots, x_k\}$ and $z = (z_1,\ldots,z_k)^\tau$. Given $z_{k+1} \in I_{k+1}$, define $\delta' \triangleq \delta - \phi^{(k+1)}(z_{k+1})x_{k+1}$. Then,

$$\begin{aligned}
&\{z : (z_1,\ldots,z_{k+1})^\tau \in \mathcal{B}(\delta)\} \cap \prod_{i=1}^k A^c(x_i^\tau(\phi^{(i)})') \cap \prod_{i=1}^k \mathbb{S}_i\\
=\ &\partial(\{z : (\phi')^\tau(z)x' > \delta'\}) \cap \prod_{i=1}^k A^c(x_i^\tau(\phi^{(i)})') \cap \prod_{i=1}^k \mathbb{S}_i.
\end{aligned}$$

In addition, by the definition of $\{L_i, K_i\}_{i=1}^n$ in (A12), $(\prod_{i=1}^k A^c(x_i^\tau(\phi^{(i)})'))^c = (\bigcup_{i=1}^k\{z : z_i \in L_i\}) \cup \prod_{i=1}^k K_i$, so we arrive at

$$\begin{aligned}
&\{z : (z_1,\ldots,z_{k+1})^\tau \in \mathcal{B}(\delta)\} \cap \left(\prod_{i=1}^k \bigcup_{j=1}^{p_i} \overline{S_i^j(q)}\right)\\
\subset\ &\partial(\{z : (\phi')^\tau(z)x' > \delta'\}) \cap \prod_{i=1}^k \mathbb{S}_i \cup \bigcup_{i=1}^k\{z : z_i \in L_i\} \cup \prod_{i=1}^k K_i.
\end{aligned}$$

Consequently, for any $z_{k+1} \in A \in \mathcal{T}^2 \setminus \mathcal{T}^3$ and $B \in \mathcal{T}^1 \setminus \mathcal{T}^4$,

$$\{z : (z_1,\ldots,z_{k+1})^\tau \in \mathcal{B}(\delta)\} \cap \prod_{i=1}^k \mathbb{S}_i \cap B \subset \partial(\{z : (\phi')^\tau(z)x' > \delta'\}) \cap \prod_{i=1}^k \mathbb{S}_i \cap B.$$

Now, for $\partial(\{z : (\phi')^\tau(z)x' > \delta'\}) \cap \prod_{i=1}^k \mathbb{S}_i$ and $\mathcal{T}^1$, applying Lemma 2 with $n = k$ leads to

$$\sum_{B\in\mathcal{T}^1\setminus\mathcal{T}^4} I_{Z(B)}(z_{k+1}) \leqslant Cr^{k-1}. \tag{A26}$$

Based on (A26), it is readily to compute

$$\begin{aligned}
&\int_{I_{k+1}} \sum_{B\in\mathcal{T}^1} I_{\mathcal{Z}(B)} dz_{k+1} = \sum_{A\in\mathcal{T}^2}\int_A \sum_{B\in\mathcal{T}^1} I_{\mathcal{Z}(B)} dz_{k+1}\\
\leqslant\ &\sum_{A\in\mathcal{T}^2\setminus\mathcal{T}^3}\int_A \sum_{B\in\mathcal{T}^1} I_{\mathcal{Z}(B)} dz_{k+1} + \sum_{A\in\mathcal{T}^3}\int_A r^k dz_{k+1}\\
=\ &\sum_{A\in\mathcal{T}^2\setminus\mathcal{T}^3}\int_A \sum_{B\in\mathcal{T}^1\setminus\mathcal{T}^4} I_{\mathcal{Z}(B)} dz_{k+1} + \sum_{A\in\mathcal{T}^2\setminus\mathcal{T}^3}\int_A \sum_{B\in\mathcal{T}^4} I_{\mathcal{Z}(B)} dz_{k+1} + r^k\cdot\frac{|I_{k+1}|}{r}\cdot|\mathcal{T}^3|\\
\leqslant\ &\int_{I_{k+1}} Cr^{k-1}dz_{k+1} + \sum_{B\in\mathcal{T}^4}\int_{I_{k+1}} 1dz_{k+1} + r^{k-1}|I_{k+1}||\mathcal{T}^3|\\
\leqslant\ &((C+|\mathcal{T}^3|)r^{k-1} + |\mathcal{T}^4|)|I_{k+1}|.
\end{aligned}$$

The result follows from Lemmas 4 and A.3(ii). □

Lemma 6. There is a constant $C_2 > 0$ depends only on ϕ such that $\sum_{B\in\mathcal{T}^1}|\mathcal{Z}_2(B)| \leqslant C_2 r^k$.

Proof. Let

$$\mathcal{T}^5 \triangleq \left\{\prod_{i=1}^k I_i' \in \mathcal{T}^1 : \partial\left(\bigcup_{j=1}^{p_i} S_i^j(q)\right) \cap I_i' \neq \emptyset \text{ for some } i \in [1,k]\right\}.$$

Clearly, $|\mathcal{T}^5| \leqslant 4r^{k-1}\sum_{i=1}^k p_i$. Hence,

$$\sum_{B\in\mathcal{T}^1}|\mathcal{Z}_2(B)| \leqslant \sum_{B\in\mathcal{T}^1\setminus(\mathcal{T}^5\cup\mathcal{T}^4)}|\mathcal{Z}_2(B)| + r|\mathcal{T}^4| + 4r^k\sum_{i=1}^k p_i. \tag{A27}$$

It suffices to estimate the first term in the right hand side of (A27). To this end, take a set $B=\prod_{i=1}^{k} I_i' \in \mathcal{T}^1 \setminus (\mathcal{T}^5 \cup \mathcal{T}^4)$ and let $z_{k+1} \in \partial Z(B) \cap \mathrm{int}(I_{k+1})$. Select a point $(z_1,\dots,z_k)^\tau \in B$ that

$$\begin{aligned}&\mathrm{dist}((z_1,\dots,z_{k+1})^\tau, \prod_{i=1}^{k}\partial(I_i')\times z_{k+1})\\ = & \min_{y\in\mathcal{B}(\delta)\cap\prod_{i=1}^{k+1}\mathbb{S}_i\cap(B\times z_{k+1})} \mathrm{dist}(y, \prod_{i=1}^{k}\partial(I_i')\times z_{k+1}).\end{aligned} \tag{A28}$$

Clearly, $B\in\mathcal{T}^1\setminus(\mathcal{T}^5\cup\mathcal{T}^4)$ implies that for each $i=1,\dots,k$, $\mathrm{int}(I_i')\subset\bigcup_{j=1}^{p_i} S_i^j(q)$ and $\mathrm{int}(I_i')\cap L_i=\emptyset$. We consider the following two cases:

Case 1: $(z_1,\dots,z_k)^\tau \notin \prod_{i=1}^{k}\partial(I_i')$. Then, there is an integer $i\in[1,k]$ such that $z_i\in\mathrm{int}(I_i')$. By (A28), $z_i\notin K_i\cap\mathrm{int}(I_i')$. Otherwise, there is a $\rho>0$ such that $x_i^\tau(\phi^{(i)})'\equiv 0$ on $[z_i-\rho, z_i+\rho]\subset\mathrm{int}(I_i')$. Similar to (A23)–(A24), for any $z_i'\in[z_i-\rho,z_i+\rho]$, $(z_1,\dots,z_{i-1},z_i',z_{i+1},\dots,z_{k+1})^\tau\in\mathcal{B}(\delta)\cap\prod_{i=1}^{k+1}\mathbb{S}_i\cap(B\times z_{k+1})$. Then, $\min\{\mathrm{dist}((z_1,\dots,z_{i-1},z_i-\rho,z_{i+1},\dots,z_{k+1})^\tau,\prod_{i=1}^{k}\partial(I_i')\times z_{k+1}),\mathrm{dist}((z_1,\dots,z_{i-1},z_i+\rho,z_{i+1},\dots,z_{k+1})^\tau,\prod_{i=1}^{k}\partial(I_i')\times z_{k+1})\}<\mathrm{dist}((z_1,\dots,z_{k+1})^\tau,\prod_{i=1}^{k}\partial(I_i')\times z_{k+1})$, which contradicts to (A28).

Now, since $z_i\notin K_i\cap\mathrm{int}(I_i')$ and $B\notin\mathcal{T}^4$, it yields that $x_i^\tau(\phi^{(i)})'(z_i)\neq 0$. We claim

$$z_{k+1}\in\bigcup_{j=1}^{p_{k+1}}\partial(S_{k+1}^j(q)). \tag{A29}$$

Otherwise, $z_{k+1}\in\bigcup_{j=1}^{p_{k+1}}S_{k+1}^j(q)$. By the *Implicit function theorem*, there is a sufficiently small $\eta>0$ such that for any $z_{k+1}'\in(z_{k+1}-\eta,z_{k+1}+\eta)$, a point $z_i'\in\mathrm{int}(I_i)$ exists and $(z_1,\dots,z_{i-1},z_i',z_{i+1},\dots,z_k,z_{k+1}')^\tau\in\mathcal{B}(\delta)\cap\prod_{i=1}^{k+1}\mathbb{S}_i$. This means $z_{k+1}\in\mathrm{int}(Z(B))$, which is impossible due to $z_{k+1}\in\partial Z(B)$. Hence (A29) holds.

Case 2: $(z_1,\dots,z_k)^\tau\in\prod_{i=1}^{k}\partial(I_i')$. Since $z_{k+1}\in\partial(Z(B))$, $x_{k+1}^\tau\phi^{(k+1)}$ cannot be a constant on any neighbourhood of z_k. So,

$$z_{k+1}\ \in\ \partial(\{z: x_{k+1}^\tau\phi^{(k+1)}(z)\neq\bar\delta\})\cap\left(\bigcup_{j=1}^{p_{k+1}}S_{k+1}^j(q)\right)\cup\left(\bigcup_{j=1}^{p_{k+1}}\partial(S_{k+1}^j(q))\right), \tag{A30}$$

where $\bar\delta=\delta-\sum_{i=1}^{k}x_i^\tau\phi^{(i)}(z_i)$.

Combining the above two cases, $z_{k+1}\in\partial(Z(B))\cap\mathrm{int}(I_{k+1})$ implies (A30). Taking the case $z_{k+1}\in\partial(I_{k+1})$ into consideration, we obtain

$$\partial(Z(B))\subset\partial(\{y\in\mathbb{R}: x_{k+1}^\tau\phi^{(k+1)}(y)\neq\bar\delta\})\cap\left(\bigcup_{j=1}^{p_{k+1}}S_{k+1}^j(q)\right)\cup\left(\bigcup_{j=1}^{p_{k+1}}\partial(S_{k+1}^j(q))\right)\cup\partial(I_{k+1}),$$

which, together with the fact $|\partial(\{z: x_{k+1}^\tau\phi^{(k+1)}(z)\neq\bar\delta\})\cap(\bigcup_{j=1}^{p_{k+1}}S_i^j(q))|\leqslant 4p_{k+1}(|L_{k+1}|+2)$ from (A13), leads to $|\mathcal{Z}_2(B)|\leqslant 2|\partial(Z(B))|\leqslant 8p_{k+1}(|L_{k+1}|+2)+4p_{k+1}+4$. Now, in view of (A27), we derive

$$\sum_{B\in\mathcal{T}^1}|\mathcal{Z}_2(B)|\leqslant(8p_{k+1}(|L_{k+1}|+2)+4p_{k+1}+4)r^k+|\mathcal{T}^4|r+4r^k\sum_{i=1}^{k}p_i,$$

which yields the result by Lemma 4. □

References

1 Chen, H. F., Guo, L.. Identification and Stochastic Adaptive Control. 1991, Birkhauser: Boston, MA.
2 Liu, Z., Li, C.. Stabilizability theorems on discrete-time nonlinear uncertain systems. arXiv preprint arXiv:2007.11122, 2020.