https://doi.org/10.1090/amsip/017/02

AMS/IP Studies in Advanced Mathematics
Volume 17, 2000

Adaptive Control of Discrete-Time Nonlinear Systems with Structural Uncertainties

Liang-Liang XIE and Lei GUO

Contents

1. Introduction

Adaptive control is usually regarded as a control method that can deal with systems with unknown or changing dynamics. However, since ordinary feedback control also has the same objective of reducing the effects of the system uncertainties on the control performance, "a meaningful definition of adaptive control, which would make it possible to look at a controller hardware and software and decide whether or not it is adaptive, is still lacking" (see [**1**, p.1]). In this paper, we do not intend to distinguish adaptive control from nonlinear feedback control, and instead, place our focus on controller design of systems with uncertainties.

The uncertainties of a system usually stem from two sources: structure uncertainties and external disturbances (noises). In general, the later is easier to cope with than the former. A fundamental question in the area of systems and control is: What are the limitations and capabilities of feedback for controlling uncertain systems? This is a conundrum, on which only a few control areas could shed some light. Robust control and adaptive control are two such areas where structure uncertainties of a system are the main concern in the controller design. Robust control usually requires that the true system lies in a small ball centered at a known nominal model (cf. [**20**]), whereas adaptive control does not need such a prerequest (cf. [**1**] [**13**]). Although much progress has been made in these two areas over the

1991 *Mathematics Subject Classification.* 93C10, 93C40, 93C55, 93D15, 93D21.

past two decades, the understanding of the fundamental question concerning about capabilities and limitations of feedback is far from being complete.

For linear finite dimensional systems with uncertain parameters, a well-developed theory of adaptive control exists today, both for stochastic systems (cf. [**4**] [**9**] [**11**]) and for deterministic systems with small unmodelled dynamics (cf. [**13**]). The three fundamental problems, i.e., least-square-based adaptive tracking, adaptive pole-placement and linear-quadratic-Gaussian control for finite-dimensional linear stochastic systems with time-invariant unknown parameters have been extensively studied and solved by now (see, e.g. [**11**] [**9**]).

In recent years, attempts have been made towards a theory of adaptive nonlinear control. Because of the difficulties involved in estimating general parametric models, almost all of the existing works concern with models where the unknown parameters appear linearly. If the nonlinearity is only involved in the input part, or if the output part of a system is nonlinear but has a linear growth rate, then the adaptive linear methods for both design and analysis can still be applied (see, e.g., [**17**], [**18**]). We will briefly delineate this in Chapter II.

However, the situation changes dramatically when one attempts to deal with discrete-time systems with output nonlinearities having growth rates faster than linear. Not only the existing methods are no longer useful, but also the stabilizability itself will become a problem ([**10**]). In Chapter III, we will present a group of conditions, under which feedback is no longer capable of stabilizing the uncertain nonlinear systems, which is a phenomenon that does not exist in the linear case. For the counterpart nonlinear continuous-time systems, however, if one can measure the output signals continuously and change the control input accordingly at the same time, then stabilizability is not a problem. Indeed, as will be shown in Remark 3.3.3 of Chapter III, we can always globally stabilize such class of continuous-time systems by the nonlinear damping control.

Nevertheless, if only sampled data of the outputs of continuous-time systems are accessible and/or the control inputs cannot follow the design in time, then stabilizability becomes a problem again. The nonlinear damping methods are no longer powerful even for deterministic systems as will be shown in Chapter IV. In fact, in the same chapter we will further show that even the classical Lyapunov-based adaptive control design methodology for linear systems also loses its ability when only sampled data of the outputs are accessible.

In the last chapter (Chapter V), instead of considering parameterized models, we analyze nonparametric models. As a novel direction in adaptive control, some nonparametric functional estimation methods developed in mathematical statistics are analyzed for the case where the data are generated by the underlining nonlinear dynamical systems. With these methods combined with the ideas common in adaptive control, the stability and optimality of a class of nonlinear nonparametric control systems are established.

2. Adaptive Control of Systems with Linearly-Growing Output Dynamics

2.1. Nonlinear Systems with Linearly-Growing Nonlinearities. Consider the following discrete-time nonlinear stochastic systems:

$$y_{t+1} = \theta^\tau f(y_t \cdots y_{t-p+1}, u_t \cdots u_{t-q+1}) + w_{t+1}. \tag{1}$$

where y_t , u_t and w_t are the system outputs , inputs and noise respectively; $\theta \in R^d$ is an unknown parameter vector; f is a known d-dimensional nonlinear function vector defined on R^{p+q}.

Denote

$$\varphi_t = (y_t \cdots y_{t-p+1}, u_t \cdots u_{t-q+1})^\tau; \tag{2}$$
$$f_t = f(y_t \cdots y_{t-p+1}, u_t \cdots u_{t-q+1}). \tag{3}$$

Our objective is to design a feedback control law based on the input-output measurements to make the system outputs track a deterministic sequence $\{y_t^*\}$. In order to analyze this control problem, we assume that

(A2.1.1) The noises $\{w_t, \mathcal{F}_t\}$ is a martingale difference sequence ($\{\mathcal{F}_t\}$ is a sequence of non-decreasing sub-σ-algebras), and there exists $\beta > 2$ such that

$$\sup_t E[|w_{t+1}|^\beta | \mathcal{F}_t] < \infty \quad \text{a.s.}; \tag{4}$$

(A2.1.2) (nonlinear minimum phase condition) There exists a constant $\lambda \in (0,1)$ such that

$$u_{t-1}^2 = O\left(\sum_{i=0}^{t} \lambda^{t-i}(y_i^2 + w_i^2)\right);$$

(A2.1.3) $\{y_t^*\}$ is bounded;

(A2.1.4) There exist some constants K_1, $K_2 > 0$ such that

$$\|f(x)\| \leq K_1 + K_2\|x\|, \qquad \forall x \in R^{p+q}. \tag{5}$$

Remark 2.1.1. *The Assumptions (A2.1.1)-(A2.1.3) correspond to the standard assumptions in the linear case (cf. e.g.* [**8**]*), while Assumption (A2.1.4) is a condition on the nonlinearity.*

If the parameter θ were known, then for our control objective, we could choose u_t to be the solution of the equation

$$y_{t+1}^* = \theta^\tau f(y_t \cdots y_{t-p+1}, u_t \cdots u_{t-q+1}).$$

But in the present case, θ is unknown, so we have to estimate it online. Here we adopt the recursive weighted least squares (WLS) algorithm ([**2**] [**3**] [**9**]).

Let θ_t be the estimated values of θ at the t-th step, recursively defined by the following WLS algorithm:

$$\theta_{t+1} = \theta_t + a_t P_t f_t (y_{t+1} - f_t^\tau \theta_t); \tag{6}$$
$$P_{t+1} = P_t - a_t P_t f_t f_t^\tau P_t; \tag{7}$$
$$a_t = (\alpha_t^{-1} + f_t^\tau P_t f_t)^{-1}, \tag{8}$$

where the initial values θ_0 and $P_0 > 0$ can be chosen arbitrarily; $\{\alpha_t\}$ is the weighting sequence defined by

$$\alpha_t = \frac{1}{h(r_t)}, \quad r_t = 1 + \sum_{i=0}^{t} \|f_i\|^2 \tag{9}$$

with $h(x) = \log^{1+\delta} x$ ($\delta > 0$), or more generally, see [**9**].

According to the certainty equivalence (CE) principle, we choose the adaptive control u_t to be the solution of the equation

$$y^*_{t+1} = \theta_t^\tau f(y_t \cdots y_{t-p+1}, u_t \cdots u_{t-q+1}). \tag{10}$$

(Here, for a general result, we assume that the solution exists almost surely.)

It is expected that under the above adaptive control law the tracking error be asymptotically minimal in the averaging sense, or, equivalently the accumulated tracking errors defined by

$$R_t \triangleq \sum_{i=1}^{t} (y_i - y_i^* - w_i)^2 \tag{11}$$

satisfies

$$R_t = o(t) \qquad \text{a.s..} \tag{12}$$

Let $\{d_t\}$ be a positive non-decreasing deterministic sequence $\{d_t\}$ with $d_{t+1} = O(d_t)$ and

$$w_t^2 = O(d_t) \qquad \text{a.s..} \tag{13}$$

It can be shown ([**11**]) that under Assumption (A2.1.1), d_t can be chosen as

$$d_t = t^\delta \qquad \forall \delta \in (\frac{2}{\beta}, 1), \tag{14}$$

where β is given by (4).

The following theorem gives the asymptotic minimality and the convergence rate of the closed-loop tracking error ([**17**]).

Theorem 2.1.1. *Let the system (1) satisfy the Assumptions (A2.1.1)-(A2.1.4). If the adaptive control law is defined by (6)-(10), then the closed-loop system is stable and optimal, and has the convergence rate*

$$R_t = o(h(t)) + O(\varepsilon_t) \quad a.s., \tag{15}$$

where,

$$\varepsilon_t = \max_{1 \le i \le t} \{\delta_i d_i\},$$

$h(\cdot), R_t, d_t$ *are defined by (9), (11) and (13) respectively, and*

$$\delta_t \triangleq tr(P_t - P_{t+1}). \tag{16}$$

To prove Theorem 2.1.1, we introduce some notations:

$$\beta_t \triangleq \frac{(f_t^\tau \widetilde{\theta}_t)^2}{\alpha_t^{-1} + f_t^\tau P_t f_t}, \quad \widetilde{\theta}_t \triangleq \theta - \theta_t, \tag{17}$$

$$\bar{r}_t \triangleq 1 + \sum_{i=1}^{t} \|\varphi_i\|^2. \tag{18}$$

We first present several lemmas.

Lemma 2.1.1. ([**9**]) *If Assumption (A2.1.1) holds, then the estimate* θ_t *defined by (6)-(8) satisfies*

$$\sum_{i=1}^{\infty} \frac{(f_i^\tau \widetilde{\theta}_i)^2}{\alpha_i^{-1} + f_i^\tau P_i f_i} < \infty, \qquad \text{a.s.,}$$

where $\widetilde{\theta}_i$ *is defined in (17).*

Lemma 2.1.2. *Under the conditions of Theorem 2.1.1, there exists some positive stochastic sequence* $\{L_t\}$ *such that*

$$y_t^2 \le L_t \quad \text{a.s.,} \quad \forall t$$

and $\{L_t\}$ *satisfies*

$$L_{t+1} \le (\lambda + c\beta_t\delta_t)L_t + \xi_t,$$

where the constant $\lambda \in (0,1), c > 0$ *and*

$$\xi_t = O(d_t) + o(\alpha_t^{-1}).$$

Proof. By (1) and (10), we have

$$y_{t+1} = f_t^\tau \widetilde{\theta}_t^\tau + y_{t+1}^* + w_{t+1}. \tag{19}$$

Thus, by Lemma 2.1.1, Assumptions (A2.1.2)-(A2.1.4), (13) and (16) -(17), we know

$$\begin{aligned} y_{t+1}^2 &\le 2(f_t^\tau \widetilde{\theta}_t)^2 + O(d_t) \\ &\le 2\beta_t\{\alpha_t^{-1} + f_t^\tau P_{t+1} f_t + f_t^\tau (P_t - P_{t+1}) f_t\} + O(d_t) \\ &\le 2\beta_t\{2\alpha_t^{-1} + \delta_t \|f_t\|^2\} + O(d_t) \\ &= O(\beta_t\delta_t \sum_{i=0}^{t+1} \lambda^{t+1-i} y_i^2) + O(d_t) + o(\alpha_t^{-1}), \end{aligned} \tag{20}$$

Where we have used the fact that $\alpha_t f_t^\tau P_{t+1} f_t \le 1$.

By Lemma 2.1.1, we have $\beta_i \to 0$. Then by (20), we know that there exists some constant $c > 0$ such that

$$y_{t+1}^2 \le c\beta_t\delta_t \sum_{i=0}^{t} \lambda^{t-i} y_i^2 + O(d_t) + o(\alpha_t^{-1}) = c\beta_t\delta_t L_t + O(d_t) + o(\alpha_t^{-1}) \tag{21}$$

where, $L_t \triangleq \sum_{i=0}^{t} \lambda^{t-i} y_i^2$, and obviously, $y_t^2 \le L_t$. By (21) and the definition of L_t, we have

$$L_{t+1} = \lambda L_t + y_{t+1}^2 \le (\lambda + c\beta_t\delta_t)L_t + O(d_t) + o(\alpha_t^{-1}). \tag{22}$$

Hence Lemma 2.1.2 is proved. □

Lemma 2.1.3. Under the conditions of Lemma 2.1.2, we have

$$\|\varphi_t\|^2 = O(d_t) + o(\alpha_t^{-1}) \quad \text{a.s.,} \quad \forall \varepsilon > 0,$$

where α_t and d_t are defined in (8) and (13) respectively.

Proof. Since $\lambda \in (0,1)$, and $\beta_t \to 0$, $\delta_t \to 0$, by (22), we have

$$L_{t+1} = O(d_t) + o(\alpha_t^{-1}).$$

Hence,

$$y_{t+1}^2 = O(d_t) + o(\alpha_t^{-1}).$$

Then by the minimum phase condition (A2.1.2), we have

$$u_t^2 = O(d_t) + o(\alpha_t^{-1}).$$

Combining the last two equations and noting the definition of φ_t in (2), we see that Lemma 2.1.3 holds. □

Proof of Theorem 2.1.1. First, noticing that $\varepsilon_t = O(d_t)$, we can get the optimality by (14) if (15) holds; Further, by the optimality and Assumptions (A2.1.1)-(A2.1.3), we immediately have $\sum_{i=1}^{t}(y_i^2 + u_i^2) = O(t)$, a.s. and the closed-loop system will be stable.

Hence we need only to prove (15).

By (16)-(17), Lemmas 2.1.1 and 2.1.3 and Assumption (A2.1.4), we deduce from (19) that

$$\begin{aligned} R_{t+1} &= \sum_{i=0}^{t}(y_{i+1} - y_{i+1}^* - w_{i+1})^2 = \sum_{i=0}^{t}(f_i^\tau \widetilde{\theta}_i)^2 \\ &= \sum_{i=0}^{t}\beta_i(\alpha_i^{-1} + f_i^\tau P_i f_i) = o(\alpha_t^{-1}) + \sum_{i=0}^{t}\beta_i(\delta_i\|f_i\|^2 + f_i^\tau P_{i+1} f_i) \\ &= o(\alpha_t^{-1}) + \sum_{i=0}^{t}\beta_i\delta_i(K_1^2 + K_2^2\|\varphi_i\|^2) = o(\alpha_t^{-1}) + O(\max_{1\le i\le t}\delta_i\|\varphi_i\|^2) \\ &= o(\alpha_t^{-1}) + O(\max_{1\le i\le t}\delta_i d_i) = o(h(r_t)) + O(\max_{1\le i\le t}\delta_i d_i) \end{aligned}$$

Hence, it remains to prove that $r_t = O(t)$.

By the equation above and Assumptions (A2.1.1) and (A2.1.3), we have

$$\sum_{i=1}^{t+1} y_i^2 = O(t) + o(h(r_t)).$$

By this and Assumption (A2.1.2), we have

$$\sum_{i=1}^{t} u_i^2 = O(t) + o(h(r_t)).$$

Therefore, by the last two equations above and Assumption (A2.1.4), we have

$$r_t = 1 + \sum_{i=0}^{t}\|f_i\|^2 = O(t) + o(h(r_t)).$$

Finally, by the definition of $h(\cdot)$, we have $r_t = O(t)$. □

Consider the following discrete-time nonlinear stochastic system:

$$y_{t+1} = \alpha^\tau f(\varphi_t) + \beta^\tau g(\varphi_t)u_t + w_{t+1} \tag{23}$$

$$\varphi_t = (y_t \cdots y_{t-p+1}, u_{t-1} \cdots u_{t-q})^\tau \tag{24}$$

where y_t , u_t and w_t are the output, input and random noise sequences, respectively; $\alpha \in R^m$ and $\beta \in R^l$ are unknown parameter vectors; $f(\cdot)$ and $g(\cdot)$ are nonlinear vector functions defined on R^{p+q}.

In terms of the connections with the input u_t in (23), $f(\cdot)$ may be called additive nonlinearity, while $g(\cdot)$ multiplicative nonlinearity, and the system (23) may be (formally) regarded as an affine nonlinear input/output model.

To analyze the control problem, we need the following conditions on the system (23):

(A2.2.1) There exist constants K_1 and K_2 such that

$$\|f(x)\| \leq K_1 + K_2\|x\|, \qquad \forall x \in R^{p+q}.$$

(A2.2.2) $p \geq q$, and there exists a decomposition $p = p_1 + p_2$ with $p_2 \geq \max(q, p_1)$, $p_1 \geq 0$ such that the function

$$g(x) = g(x_1, x_2, x_3), \quad x_1 \in R^{p_1}, x_2 \in R^{p_2}, x_3 \in R^q$$

is uniformly bounded for bounded x_2, and uniformly tends to ∞ as $\|x_2\| \to \infty$, where the uniformity is w.r.t. $(x_1, x_3) \in R^{p_1+q}$.

(A2.2.3) There exists a nonzero multivariate polynomial function $P(\gamma)$, $\gamma \in R^l$, such that the set

$$\mathcal{B} \stackrel{\triangle}{=} \{\gamma : P(\gamma) \neq 0\}. \tag{25}$$

contains the true system parameter β defined in (23), and for any $\gamma \in \mathcal{B}$ there exist constants $L(\gamma) > 0$ and $M(\gamma) > 0$, such that for all $\|x_2\| \geq L(\gamma)$,

$$\|g(x_1, x_2, x_3)\| \leq M(\gamma)(|\gamma^\tau g(x_1, x_2, x_3)|), \qquad \forall (x_1, x_3) \in R^{p_1+q}.$$

(A2.2.4) $\{w_t, \mathcal{F}_t\}$ is a martingale difference sequence, where $\{\mathcal{F}_t\}$ is a non-decreasing sequence of sub-σ-algebras. Assume also that

$$\sup_t E[|w_{t+1}|^2|\mathcal{F}_t] < \infty, \qquad \text{a.s.}$$

and

$$\sum_{t=1}^n w_t^2 = O(n). \tag{26}$$

Obviously, (A2.2.1) and (A2.2.4) are standard conditions. (A2.2.2) mainly says that $\|g(x_1, x_2, x_3)\|$ grows to ∞ as $\|x_2\| \to \infty$, while (A2.2.3) requires that its growth rate is unchanged when it is multiplied by any γ defined in (25). Note that if there exist two constants $0 < \alpha_1 < \alpha_2$ such that $\alpha_1 \leq \beta^\tau g(\varphi_t) \leq \alpha_2$ $\forall t \geq 0$, then it will be reduced to the case considered in Section 2.1.

We now give two examples to illustrate (A2.2.2) and (A2.2.3).

Example 2.2.1 Consider the following system with the multiplicative nonlinearity being a polynomial of y_t :

$$y_{t+1} = a_1 f_1(y_t, u_{t-1}) + \cdots + a_m f_m(y_t, u_{t-1}) + (b_0 + b_1 y_t + \cdots + b_l y_t^l)u_t + w_{t+1},$$

where $\alpha = (a_1, \cdots, a_m)^\tau$ and $\beta = (b_0, \cdots, b_l)^\tau$ are unknown parameters; $|f_i(x)| \leq M(\|x\|+1), x \in R^2, 1 \leq i \leq m$, and $b_l \neq 0, l \geq 1$. Set $p_1 = 0$, $p_2 = 1$ and $q = 1$, and define $g(x_1, x_2, x_3) = (1, x_2, \cdots, x_2^l)^\tau$, and $P(\gamma) = \gamma_l$ for $\gamma = (\gamma_0, \cdots, \gamma_l)^\tau \in R^{l+1}$, then it is easy to see that (A2.2.2) and (A2.2.3) hold.

It is worth noting that the power l in the above example can be arbitrarily large. We next present an example where the multiplicative part also contains the input sequence u_t.

Example 2.2.2 Consider the system :

$$\begin{aligned} y_{t+1} = \ & a_1 f_1(y_t, y_{t-1}, u_{t-1}) + \cdots + a_m f_m(y_t, y_{t-1}, u_{t-1}) \\ & +[b_0 + b_1 B_1(y_t) + b_2 |y_{t-1}|^\delta + b_3 B_2(u_{t-1})] u_t + w_{t+1}, \end{aligned}$$

where, $\alpha = (a_1, \cdots, a_m)^\tau$, $\beta = (b_0, \cdots, b_3)^\tau$ are unknown parameters; $\delta > 0$; $B_1(\cdot)$ and $B_2(\cdot)$ are two bounded functions; $|f_i(x)| \leq M(\|x\| + 1), x \in R^3, 1 \leq i \leq m$, and $b_2 \neq 0$. To verify (A2.2.2) and (A2.2.3), we just set $p_1 = p_2 = q = 1$, $g(x_1, x_2, x_3) = [1, B_1(x_1), |x_2|^\delta, B_2(x_3)]^\tau$, and $P(\gamma) = \gamma_2$ for $\gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3)^\tau$.

Now, we consider the following weighted one-step-ahead control performance:

$$J(u_t) = E\left\{y_{t+1}^2 + \lambda u_t^2 | \mathcal{F}_t\right\}, \qquad \lambda > 0. \tag{27}$$

Here, to guarantee the finiteness of the control energy, we do not choose the pure minimum variance cost $J_1(u_t) = E\left\{y_{t+1}^2 | \mathcal{F}_t\right\}$, since even for simple bilinear systems the usual minimum phase condition may not be satisfied (cf. [**5**]).

The optimal nonadaptive control law that minimizes (27) is given by

$$u_t = -\frac{[\alpha^\tau f(\varphi_t)][\beta^\tau g(\varphi_t)]}{[\beta^\tau g(\varphi_t)]^2 + \lambda}. \tag{28}$$

For estimating the unknown parameters in this control law, we adopt the random regularization method introduced in [**9**] and the weighted least squares (WLS) algorithm proposed in [**3**] and further studied in [**2**] and [**9**].

Set

$$\theta = [\alpha^\tau, \beta^\tau]^\tau, \tag{29}$$

$$\phi_t = [f^\tau(\varphi_t), g^\tau(\varphi_t) u_t]^\tau. \tag{30}$$

Let θ_t be the estimated values of θ, which are recursively defined by the following WLS algorithm:

$$\theta_{t+1} = \theta_t + a_t P_t \phi_t (y_{t+1} - \phi_t^\tau \theta_t) \tag{31}$$

$$P_{t+1} = P_t - a_t P_t \phi_t \phi_t^\tau P_t \tag{32}$$

$$a_t = (\lambda_t^{-1} + \phi_t^\tau P_t \phi_t)^{-1} \tag{33}$$

where, the initial values θ_0 and $P_0 > 0$ can be chosen arbitrarily; $\{\lambda_t\}$ is the weighting sequence defined by

$$\lambda_t = \frac{1}{h(r_t)}, \quad r_t = \|P_0^{-1}\| + \sum_{i=0}^{t} \|\phi_i\|^2 \tag{34}$$

with $h(x) = \log^{1+\delta} x$ $(\delta > 0)$, or see [**9**] for more general choices.

Since the estimate for β given by the above WLS may not belong to the set $\mathcal{B}$ defined by (25), we now resort to the random regularization method introduced in [**9**] to secure this.

Let $\{\zeta_t\}$ be an independent sequence of $(m+l)$–dimensional random vectors that are uniformly distributed on the unit ball $\{x \in R^{m+l} : \|x\| \leq 1\}$ and independent of $\{w_t\}$. Define $T_t(x) = |P(\theta_t + P_t^{\frac{1}{2}} x)|$, $\quad x \triangleq (x_1 \cdots x_{m+l}) \in R^{m+l}$, where $P(x_1 \cdots x_{m+l}) \triangleq P(x_{m+1} \cdots x_{m+l})$ is the polynomial function defined in (25). Take a number $\sigma \in (0, \sqrt{2}-1)$, and define a sequence $\{\eta_t\}$ recursively as follows:

$$\eta_t = \begin{cases} \zeta_t, & \text{if } T_t(\zeta_t) \geq (1+\sigma)T_t(\eta_{t-1}); \\ \eta_{t-1}, & \text{otherwise}, \end{cases} \tag{35}$$

with initial value $\eta_0 = \zeta_0$. Let

$$\widehat{\theta}_t = \theta_t + P_t^{\frac{1}{2}}\eta_t, \tag{36}$$

by which we replace θ in (28), and get the following certainty-equivalence control:

$$u_t = -\frac{[\widehat{\alpha}_t^\tau f(\varphi_t)][\widehat{\beta}_t^\tau g(\varphi_t)]}{[\widehat{\beta}_t^\tau g(\varphi_t)]^2 + \lambda}, \qquad \widehat{\theta}_t = [\widehat{\alpha}_t^\tau, \widehat{\beta}_t^\tau]^\tau. \tag{37}$$

Now we arrive at a stability result for the affine nonlinear system (23).

Theorem 2.2.1. (see [**18**]) *For the system (23), let the conditions (A2.2.1)–(A2.2.4) be satisfied, and let the adaptive control law be defined by (30)-(37), then the closed–loop system is globally stable, i.e., for any initial condition,*

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n}(y_t^2 + u_t^2) < \infty \qquad \text{a.s..}$$

3. Limitations of Feedback for Controlling Uncertain Nonlinear Systems

In the last chapter, we have treated the case where the additive output nonlinearities have a linear growth rate. A natural question is : What can be said about the other cases?

We start with a typical system

$$y_{t+1} = \theta \cdot f(y_t) + u_t + w_{t+1}, \quad \theta \in \mathbb{R}^1, \tag{38}$$

where we assume that

(A3.1.1) θ is a Gaussian unknown parameter with distribution $N(\bar{\theta}, 1)$;

(A3.1.2) $\{w_t\}$ is an i.i.d. noise sequence independent of θ and has Gaussian distribution $N(0, \sigma_w^2)$;

(A3.1.3) There exist positive constants $b > 0$, $M > 0$ such that

$$f(x) \sim Mx^b \quad \text{as } x \to \infty.$$

Then Guo ([**10**]) found that $b = 4$ is a critical case for stabilizability.

Theorem 3.1.1. ([**10**]) *Let the above three conditions (A3.1)-(A3.3) be satisfied for the system 38. Then*

(i) If $b \geq 4$, then the system 38 is not a.s. stabilizable, i.e., for any feedback control $u_t \in \sigma\{y_0, \cdots, y_t\}$, there always exists a set D with $P(D) > 0$ such that

$$\frac{1}{T}\sum_{t=1}^{T}(y_t^2 + u_t^2)^2 \to \infty \quad \text{on } D$$

at a rate faster than exponential.

(ii) If $b < 4$, then the system 38 can be stabilized and asymptotically optimized by the LS-based adaptive control

$$u_t = -\theta_t f(y_t)$$

where θ_t is the LS estimate for θ at time t. Furthermore, the rate of tracking is the best possible, i.e.,

$$\sum_{t=1}^{T}(y_t - y_t^* - w_t)^2 = O(\log T). \quad \text{a.s.}$$

This theorem reveals the fundamental limitations of feedback control for systems with additive output nonlinearities containing only one unknown parameter. The main purpose of this chapter is to consider more general cases.

3.2. Conditional Cramer-Rao Inequalities for Dynamical Systems. In the proof of Theorem 2.1 in [**19**], the Gaussian assumption on θ played an important role since the optimality of the Kalman Filter was used. In order to treat non-Gaussian cases, we first extend the classical Cramer-Rao inequality to the conditional case for dynamical systems.

Theorem 3.2.1. (Conditional Cramer-Rao Inequality) *Let $\theta \in \mathbb{D}_n \triangleq (\underline{\theta}_1, \bar{\theta}_1) \times \cdots \times (\underline{\theta}_n, \bar{\theta}_n) \subset \mathbb{R}^n$ be a random parameter vector with probability density function (p.d.f.) $p(\theta)$ (where $\underline{\theta}_i$ and $\bar{\theta}_i$ may be either infinite or finite), and x be a given random vector. Denote $E_x y \triangleq E\{y|x\}$ for any random vector*

y. Let $g(x,\theta) = [g_1(x,\theta), g_2(x,\theta), \cdots, g_m(x,\theta)]^T$ be any measurable function vector having partial derivatives of first order with respect to (w.r.t.) θ and with $E_x g(x,\theta)$, $E_x \dfrac{\partial g(x,\theta)}{\partial \theta}$ existing. Assume that $p(x,\theta)$, the joint p.d.f. of x and θ, has partial derivatives of second order w.r.t. θ, and that for any fixed x and $\theta_1, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_n$,

(i) $\lim_{\theta_i \to \underline{\theta}_i \text{ or } \bar{\theta}_i} p(x,\theta) = 0 \quad 1 \le i \le n;$

(ii) $\lim_{\theta_i \to \underline{\theta}_i \text{ or } \bar{\theta}_i} g_j(x,\theta) p(x,\theta) = 0 \quad 1 \le i \le n,\ 1 \le j \le m;$

(iii) $\lim_{\theta_i \to \underline{\theta}_i \text{ or } \bar{\theta}_i} \dfrac{\partial p(x,\theta)}{\partial \theta_j} = 0 \quad 1 \le i, j \le n.$

Then,

$$\begin{aligned} &E_x\{[g(x,\theta) - E_x g(x,\theta)][g(x,\theta) - E_x g(x,\theta)]^T\} \\ &\ge E_x \frac{\partial g(x,\theta)}{\partial \theta} \left\{ -E_x \left[\frac{\partial^2 \log p(x,\theta)}{\partial \theta^2} \right] \right\}^{-1} E_x^T \frac{\partial g(x,\theta)}{\partial \theta}, \end{aligned}$$

and particularly for $g(x,\theta) = \theta$,

$$E_x[(\theta - E_x\theta)(\theta - E_x\theta)^T] \ge \left\{ -E_x \left[\frac{\partial^2 \log p(x,\theta)}{\partial \theta^2} \right] \right\}^{-1},$$

where we assume that $p(\theta,x) > 0$, $\forall \theta \in I\!D_n$, $\forall x$.

Proof. We first prove the conditional version of an inequality in [**12**, p.73]. Let ξ, η be two r-dimensional random vectors and let all the entries of these vectors possess finite variance. If $E_x(\eta - E_x\eta)(\eta - E_x\eta)^T > 0$, then

(39)
$$\begin{aligned} E_x(\xi - E_x\xi)(\xi - E_x\xi)^T \ &\ge [E_x(\xi - E_x\xi)(\eta - E_x\eta)^T] \\ &\times [E_x(\eta - E_x\eta)(\eta - E_x\eta)^T]^{-1} [E_x(\eta - E_x\eta)(\xi - E_x\xi)^T]. \end{aligned}$$

In fact, without loss of generality, we can suppose $E_x\eta = 0, E_x\xi = 0$. Then we only need to prove

$$E_x\xi\xi^T \ge (E_x\xi\eta^T)(E_x\eta\eta^T)^{-1}(E_x\eta\xi^T). \tag{40}$$

For any $z \in \sigma\{x\}$, $E_x(\xi - z\eta)(\xi - z\eta)^T \ge 0$, i.e.,

$$E_x\xi\xi^T - (E_x\xi\eta^T)z^T - z(E_x\eta\xi^T) + z(E_x\eta\eta^T)z^T \ge 0.$$

Let $z = (E_x\xi\eta^T)(E_x\eta\eta^T)^{-1}$, then we get (40).

Now set $\xi = g(x,\theta)$, $\eta = \dfrac{\partial p(x,\theta)}{\partial \theta} / p(\theta|x) \triangleq [\eta_1, \cdots, \eta_n]^T$.

First, for $1 \le i \le n$,

(41)
$$\begin{aligned} E_x\eta_i &= E_x \left\{ \frac{\partial p(x,\theta)}{\partial \theta_i} / p(\theta|x) \right\} \\ &= \int_{I\!D_n} \frac{\partial p(x,\theta)}{\partial \theta_i} d\theta \\ &= \int_{\underline{\theta}_n}^{\bar{\theta}_n} \cdots \int_{\underline{\theta}_{i+1}}^{\bar{\theta}_{i+1}} \int_{\underline{\theta}_{i-1}}^{\bar{\theta}_{i-1}} \cdots \int_{\underline{\theta}_1}^{\bar{\theta}_1} \int_{\underline{\theta}_i}^{\bar{\theta}_i} \frac{\partial p(x,\theta)}{\partial \theta_i} d\theta_i d\theta_1 \cdots d\theta_{i-1} d\theta_{i+1} \cdots d\theta_n. \end{aligned}$$

Inspired by the method used in [**16**, p.72], by Assumption (i), we have

$$\int_{\underline{\theta}_i}^{\bar{\theta}_i} \frac{\partial}{\partial \theta_i} p(x,\theta) d\theta_i = 0 \quad \text{for any fixed } x, \theta_1, \cdots, \theta_{i-1}, \theta_{i+1}, \cdots, \theta_n.$$

Hence, by (41),

$$E_x \eta_i = 0, \quad 1 \le i \le n. \tag{42}$$

Next, we have

$$\begin{aligned} & E_x[(\xi - E_x\xi)\eta^T] \\ = \; & E_x \left\{ [g(x,\theta) - E_x g(x,\theta)] \frac{\partial^T p(x,\theta)}{\partial \theta} / p(\theta|x) \right\} \\ = \; & \int_{\mathbb{D}_n} [g(x,\theta) - E_x g(x,\theta)] \frac{\partial^T p(x,\theta)}{\partial \theta} d\theta. \end{aligned} \tag{43}$$

By Assumptions (i) (ii), for $1 \le i \le n, 1 \le j \le m$,

$$\int_{\underline{\theta}_i}^{\bar{\theta}_i} \frac{\partial [g_j(x,\theta) p(x,\theta)]}{\partial \theta_i} d\theta_i = 0, \quad \int_{\underline{\theta}_i}^{\bar{\theta}_i} \frac{\partial \{[E_x g_j(x,\theta)] p(x,\theta)\}}{\partial \theta_i} d\theta_i = 0.$$

Hence,

$$\begin{aligned} 0 &= \int_{\mathbb{D}_n} \frac{\partial \{[g(x,\theta) - E_x g(x,\theta)] p(x,\theta)\}}{\partial \theta} d\theta \\ &= \int_{\mathbb{D}_n} \left\{ \frac{\partial g(x,\theta)}{\partial \theta} p(x,\theta) + [g(x,\theta) - E_x g(x,\theta)] \frac{\partial^T p(x,\theta)}{\partial \theta} \right\} d\theta. \end{aligned}$$

That is,

$$-p(x) E_x \frac{\partial g(x,\theta)}{\partial \theta} = \int_{\mathbb{D}_n} [g(x,\theta) - E_x g(x,\theta)] \frac{\partial^T p(x,\theta)}{\partial \theta} d\theta. \tag{44}$$

By (43) and (44),

$$E_x(\xi - E_x\xi)\eta^T = -p(x) E_x \frac{\partial g(x,\theta)}{\partial \theta}. \tag{45}$$

By the definition of η, we have

$$\begin{aligned} E_x \eta\eta^T &= E_x \frac{\left[\frac{\partial}{\partial \theta} p(x,\theta)\right] \left[\frac{\partial}{\partial \theta} p(x,\theta)\right]^T}{p^2(\theta|x)} \\ &= E_x \frac{\left[\frac{\partial}{\partial \theta} \log p(x,\theta)\right] \left[\frac{\partial}{\partial \theta} \log p(x,\theta)\right]^T p^2(x,\theta)}{p^2(\theta|x)} \\ &= p^2(x) E_x \left[\frac{\partial \log p(x,\theta)}{\partial \theta} \cdot \frac{\partial^T \log p(x,\theta)}{\partial \theta} \right] \end{aligned} \tag{46}$$

For the same reason as getting 42, by Assumption (iii) we have

$$\begin{aligned} & \int_{\mathbb{D}_n} \frac{\partial^2 p(x,\theta)}{\partial \theta_i \partial \theta_j} d\theta \\ = \; & \int_{\underline{\theta}_n}^{\bar{\theta}_n} \cdots \int_{\underline{\theta}_{j+1}}^{\bar{\theta}_{j+1}} \int_{\underline{\theta}_{j-1}}^{\bar{\theta}_{j-1}} \cdots \int_{\underline{\theta}_1}^{\bar{\theta}_1} \int_{\underline{\theta}_j}^{\bar{\theta}_j} \frac{\partial}{\partial \theta_j} \left[\frac{\partial p(x,\theta)}{\partial \theta_i} \right] d\theta_j d\theta_1 \\ & \cdots d\theta_{j-1} d\theta_{j+1} \cdots d\theta_n \\ = \; & 0 \qquad\qquad 1 \le i,j \le n. \end{aligned} \tag{47}$$

On the other hand, since $p(x,\theta)$ has partial derivatives of second order with respect to θ,

$$
\begin{aligned}
&\int_{\mathbb{D}_n} \frac{\partial^2 p(x,\theta)}{\partial\theta_i \partial\theta_j} d\theta \\
= &\int_{\mathbb{D}_n} \frac{\partial}{\partial\theta_j}\left[\frac{\partial \log p(x,\theta)}{\partial\theta_i}\cdot p(x,\theta)\right] d\theta \\
= &\int_{\mathbb{D}_n} \left[\frac{\partial^2 \log p(x,\theta)}{\partial\theta_i\partial\theta_j}\cdot p(x,\theta) + \frac{\partial \log p(x,\theta)}{\partial\theta_i}\cdot\frac{\partial \log p(x,\theta)}{\partial\theta_j}\cdot p(x,\theta)\right] d\theta \\
= &\; p(x)E_x\left[\frac{\partial^2 \log p(x,\theta)}{\partial\theta_i\partial\theta_j} + \frac{\partial \log p(x,\theta)}{\partial\theta_i}\cdot\frac{\partial \log p(x,\theta)}{\partial\theta_j}\right].
\end{aligned} \tag{48}
$$

Combining (47) and (48), we have

$$E_x\left[\frac{\partial \log p(x,\theta)}{\partial\theta_i}\cdot\frac{\partial \log p(x,\theta)}{\partial\theta_j}\right] = -E_x\left[\frac{\partial^2 \log p(x,\theta)}{\partial\theta_i\partial\theta_j}\right] \tag{49}$$

Then, by (46) and (49),

$$E_x\eta\eta^T = p^2(x)\left\{-E_x\left[\frac{\partial^2 \log p(x,\theta)}{\partial\theta^2}\right]\right\}. \tag{50}$$

At last, by (39), (42), (45) and (50), we have

$$
\begin{aligned}
&E_x\{[g(x,\theta)-E_xg(x,\theta)][g(x,\theta)-E_xg(x,\theta)]^T\} \\
&\geq E_x\frac{\partial g(x,\theta)}{\partial\theta}\left\{-E_x\left[\frac{\partial^2 \log p(x,\theta)}{\partial\theta^2}\right]\right\}^{-1} E_x^T\frac{\partial g(x,\theta)}{\partial\theta},
\end{aligned}
$$

which is just the conclusion of this theorem. □

Remark 3.2.1. *By the elementary properties of Lebesgue integration, we can see that* $\frac{\partial^2 p(x,\theta)}{\partial\theta^2}$ *is only needed to exist almost everywhere for Theorem 3.2.1 to hold. That is, there can be a set with zero Lebesgue measure on which* $\frac{\partial^2 p(x,\theta)}{\partial\theta^2}$ *does not exist.*

With only a small modification on the proof of Theorem 3.2.1, we can easily get the Conditional Cramer-Rao Inequality for general bounded parameters as follows.

Theorem 3.2.2 *Let* $\theta \in \Theta \subset \mathbb{R}^n$ *be a random parameter vector, where* Θ *is closed, bounded and convex. Let* x *be a given random vector and denote* $E_x y \triangleq E\{y|x\}$ *for any random vector* y. *Let* $\partial\Theta$ *and* Θ^0 *denote respectively the boundary and the interior of* Θ. *Assume that* $p(x,\theta)$ *(the joint p.d.f. of* x *and* θ*) is continuous with respect to (w.r.t.)* θ *on* Θ, *has continuous partial derivatives of first order w.r.t.* θ *on* Θ^0 *and has bounded partial derivatives of second order w.r.t.* θ *on* $\Theta^0 - G$, *where* $G \subset \Theta^0$ *with* $V(G) = 0$ *(*$V(\cdot)$ *denotes the Lebesgue measure on* $\mathbb{R}^n$*). Furthermore, assume that for any fixed* x,

(i) $p(x,\theta) = 0, \forall\theta \in \partial\Theta$;
(ii) $p(x,\theta) > 0, \forall\theta \in \Theta^0$;
(iii) $\lim_{\theta\to\partial\Theta} \frac{\partial p(x,\theta)}{\partial\theta} = 0$.

Then for any measurable vector function $g(x,\theta) = [g_1(x,\theta), g_2(x,\theta), \cdots, g_m(x,\theta)]^T$ *having partial derivatives of first order w.r.t.* θ *and with* $E_xg(x,\theta)$ *and* $E_x\frac{\partial g(x,\theta)}{\partial\theta}$

existing, we have

$$E_x\{[g(x,\theta)-E_xg(x,\theta)][g(x,\theta)-E_xg(x,\theta)]^T\}$$
$$\geq E_x\frac{\partial g(x,\theta)}{\partial\theta}\left\{-E_x\left[\frac{\partial^2\log p(x,\theta)}{\partial\theta^2}\right]\right\}^{-1}E_x^T\frac{\partial g(x,\theta)}{\partial\theta},$$

and particularly for $g(x,\theta)=\theta$,

$$E_x[(\theta-E_x\theta)(\theta-E_x\theta)^T]\geq\left\{-E_x\left[\frac{\partial^2\log p(x,\theta)}{\partial\theta^2}\right]\right\}^{-1}.$$

Now we apply the Conditional Cramer-Rao Inequality to general dynamical systems.

Theorem 3.2.3. *Let* θ *be defined as in Theorem 3.2.1 (or Theorem 3.2.2) with p.d.f.* $p(\theta)$, *and let* $\{w_k\}$ *be an i.i.d. random sequence with p.d.f.* $q(w)$ *and independent of* θ. *Let* $Ew_k=0$ *and* $Ew_k^2=\sigma_w^2$. *For the* $t+1$ *recursive equations: (*y_0 *is deterministic)*

$$y_{k+1}=f(\theta,\varphi_k)+u_k+w_{k+1},\quad k=0,1,\cdots,t,$$

where $\varphi_k\triangleq(y_k,\cdots,y_{k-p},u_{k-1},\cdots,u_{k-q})$, $u_k\in\mathcal{F}_k^y\triangleq\sigma\{y_1,\cdots,y_k\}$, *let* $x\triangleq\{y_1,\cdots,y_t\}$ *and denote* $g(x,\theta)\triangleq f(\theta,\varphi_t)$. *If the assumptions for* $p(x,\theta)$ *and* $g(x,\theta)$ *in Theorem 3.2.1 (or Theorem 3.2.2) are satisfied, then*

$$\begin{aligned}&E_x[f(\theta,\varphi_t)-E_xf(\theta,\varphi_t)]^2\\&\geq E_x^T\frac{\partial f(\theta,\varphi_t)}{\partial\theta}\\&\times\left\{-E_x\left[\sum_{k=1}^t\frac{\partial^2\log q(y_k-f_{k-1}-u_{k-1})}{\partial\theta^2}+\frac{\partial^2\log p(\theta)}{\partial\theta^2}\right]\right\}^{-1}E_x\frac{\partial f(\theta,\varphi_t)}{\partial\theta},\end{aligned}\tag{51}$$

where $f_{k-1}\triangleq f(\theta,\varphi_{k-1})$. *Furthermore, we have*

$$\begin{aligned}&E_xy_{t+1}^2\\&\geq E_x^T\frac{\partial f(\theta,\varphi_t)}{\partial\theta}\\&\times\left\{-E_x\left[\sum_{k=1}^t\frac{\partial^2\log q(y_k-f_{k-1}-u_{k-1})}{\partial\theta^2}+\frac{\partial^2\log p(\theta)}{\partial\theta^2}\right]\right\}^{-1}E_x\frac{\partial f(\theta,\varphi_t)}{\partial\theta}+\sigma_w^2.\end{aligned}\tag{52}$$

Proof. Directly applying Theorem 3.2.1 (or Theorem 3.2.2), we have

$$E_x[f(\theta,\varphi_t)-E_xf(\theta,\varphi_t)]^2$$
$$\geq E_x^T\frac{\partial f(\theta,\varphi_t)}{\partial\theta}\left\{-E_x\left[\frac{\partial^2[\log p(x|\theta)+\log p(\theta)]}{\partial\theta^2}\right]\right\}^{-1}E_x\frac{\partial f(\theta,\varphi_t)}{\partial\theta},$$

where,

$$\begin{aligned}&p(x|\theta)\\=\;&p(y_1,y_2,\cdots,y_t|\theta)\\=\;&p(y_1|\theta,y_0)\cdot p(y_2|\theta,y_0,y_1)\cdots p(y_t|\theta,y_0,\cdots,y_{t-1})\\=\;&q(y_1-f_0-u_0)\cdot q(y_2-f_1-u_1)\cdot\cdots\cdot q(y_t-f_{t-1}-u_{t-1}).\end{aligned}\tag{53}$$

Thus, after some simple manipulations, we arrive at (51).

Furthermore, since $u_t \in \sigma\{x\}$, w_{t+1} is independent of x and

$$y_{t+1} = f(\theta,\varphi_t) - E_x f(\theta,\varphi_t) + E_x f(\theta,\varphi_t) + u_t + w_{t+1},$$

we have

$$\begin{aligned} E_x y_{t+1}^2 &= E_x[f(\theta,\varphi_t) - E_x f(\theta,\varphi_t)]^2 + [E_x f(\theta,\varphi_t) + u_t]^2 + \sigma_w^2 \\ &\geq E_x[f(\theta,\varphi_t) - E_x f(\theta,\varphi_t)]^2 + \sigma_w^2. \end{aligned} \tag{54}$$

Hence (52) follows from (54) and (51). □

Theorem 3.2.4. *If in the conditions of Theorem 3.2.3, let $p(\theta)$, instead of $p(x,\theta)$, satisfy the assumptions for $p(x,\theta)$ in Theorem 3.2.1 (or Theorem 3.2.2) and $q(w) > 0, \forall w \in I\!R^1$, then (51) and (52) still hold.*

Proof. By (53) and $q(w) > 0, \forall w \in I\!R^1$, we know that for any x, $p(x|\theta) > 0$. Hence $p(x,\theta) = p(\theta)p(x|\theta)$ also satisfies the assumptions that $p(\theta)$ satisfies. Thus the conditions of Theorem 3.2.3 are satisfies. Hence (51) and (52) hold. □

Theorem 3.2.5. *Under the conditions of Theorem 3.2.4, if the parameters enter the function linearly, i.e., $f(\theta,\varphi_k) = \theta^T g(\varphi_k) + h(\varphi_k) \triangleq \theta^T\phi_k + h(\varphi_k)$, and $w_i \sim N(0,\sigma_w^2)$, i.e., $q(w) = \frac{1}{\sqrt{2\pi}\sigma_w}\exp\left(-\frac{w^2}{2\sigma_w^2}\right)$, then*

$$E_x[(\theta - E_x\theta)(\theta - E_x\theta)^T] \geq \left\{\frac{1}{\sigma_w^2}\sum_{k=1}^{t}\phi_{k-1}\phi_{k-1}^T - E_x\frac{\partial^2 \log p(\theta)}{\partial\theta^2}\right\}^{-1}, \tag{55}$$

and

$$E_x y_{t+1}^2 \geq \phi_t^T \left\{\frac{1}{\sigma_w^2}\sum_{k=1}^{t}\phi_{k-1}\phi_{k-1}^T - E_x\frac{\partial^2 \log p(\theta)}{\partial\theta^2}\right\}^{-1}\phi_t + \sigma_w^2. \tag{56}$$

Proof. Since $q(y_k - f_{k-1} - u_{k-1}) = \frac{1}{\sigma_w\sqrt{2\pi}}\exp\{-\frac{1}{2\sigma_w^2}(y_k - f_{k-1} - u_{k-1})^2\}$, $k = 1, 2, \cdots, t$, we have

(57)

$$\begin{aligned} \frac{\partial^2 \log q(y_k - f_{k-1} - u_{k-1})}{\partial\theta^2} &= \frac{\partial^2}{\partial\theta^2}\{-\frac{1}{2\sigma_w^2}[y_k - \theta^T g(\varphi_{k-1}) - h(\varphi_{k-1}) - u_{k-1}]^2\} \\ &= -\frac{1}{\sigma_w^2}\phi_{k-1}\phi_{k-1}^T. \end{aligned}$$

Then by (53) and Theorem 3.2.1 (or Theorem 3.2.3) with $g(x,\theta) = \theta$, we get (55).

To prove (56), we can either substitute (55) and $f(\theta,\varphi_t) = \theta^T\phi_t + h(\varphi_t)$ into (54), or directly apply (52) with $f(\theta,\varphi_t) = \theta^T\phi_t + h(\varphi_t)$ and (57). □

Remark 3.2.2 *If, in addition to the conditions of Theorem 3.2.5,*

$$-\frac{\partial^2 \log p(\theta)}{\partial\theta^2} \leq M\,I, \quad \forall\theta \in I\!D_n \text{ or } \Theta,$$

for some $M > 0$, then

$$E_x[(\theta - E_x\theta)(\theta - E_x\theta)^T] \geq \left\{\frac{1}{\sigma_w^2}\sum_{k=1}^{t}\phi_{k-1}\phi_{k-1}^T + M\,I\right\}^{-1}.$$

In this remark, the conditions imposed on θ are fairly general for unbounded distributions, which includes, for example, the familiar Gaussian and Cauchy distributions.

For bounded parameters, we can get the following conclusion.

Theorem 3.2.6. *Consider the following dynamical system:*

$$y_{k+1} = \theta^T f(\varphi_k) + u_k + w_{k+1}, \quad k = 0, 1, 2, \cdots,$$

where $\varphi_k \triangleq (y_k, \cdots, y_{k-p}, u_{k-1}, \cdots, u_{k-q})$, $u_k \in \mathcal{F}_k^y \triangleq \sigma\{y_1, \cdots, y_k\}$, y_0 *is deterministic and* $y_i = 0$, $\forall i < 0$; $\{w_k\}$ *is an i.i.d. random sequence with distribution* $N(0, \sigma_w^2)$ *and independent of the unknown parameter vector* θ.
Let $\theta \in \Theta \subset I\!R^n$, *where* Θ *is closed, bounded and convex. Assume that* $p(\theta)$ *is continuous on* Θ, *has continuous partial derivatives of first order on* Θ^0 *and bounded partial derivatives of second order on* $\Theta^0 - G$, *where* $G \subset \Theta^0$ *with* $V(G) = 0$ *(*$V(\cdot)$ *denotes the Lebesgue measure on* $I\!R^n$*) and that*

(i) $p(\theta) = 0, \forall \theta \in \partial\Theta$;

(ii) $p(\theta) > 0, \forall \theta \in \Theta^0$;

(iii) $\lim\limits_{\theta \to \partial\Theta} \dfrac{\partial p(\theta)}{\partial \theta} = 0$;

(iv) $\left| \dfrac{\partial p(\theta)}{\partial \theta_i} \cdot \dfrac{\partial p(\theta)}{\partial \theta_j} \cdot \dfrac{1}{p(\theta)} \right| \le M, \quad \forall \theta \in \Theta^0, \quad 1 \le i, j \le n$;

(v) *There exist* $\varepsilon_0 > 0$, $N > 0$ *and* $M_1 > 0$ *such that for any* $\theta \in \Theta - A_t$, *where* $A_t \triangleq \{\theta : d(\theta, \partial\Theta) < \varepsilon_t \triangleq \varepsilon_0/t^3, \theta \in \Theta\}$ *(*$d(\theta, \partial\Theta)$ *denotes the distance between* θ *and* $\partial\Theta$*),*

$$p(\theta) > \frac{1}{M_1 t^N} \quad \text{and} \quad \Theta - A_0 \ne \phi \ (\phi \text{ denotes the null set}).$$

Further assume that $\|f(\varphi_k)\| \ne 0$, *a.s.*, $k = 0, 1, 2, \cdots$.
Then there exists some $D_0 \subset \Omega$ *with* $Prob(D_0) > 0$ *such that for* $t = 1, 2, \cdots$,

$$E[(\theta - \widehat{\theta}_t)(\theta - \widehat{\theta}_t)^T | \mathcal{F}_t^y] \ge \left\{ \frac{1}{\sigma_w^2} \sum_{k=0}^{t-1} \phi_k \phi_k^T + K t^N I \right\}^{-1} \quad \text{on } D_0,$$

where, $\widehat{\theta}_t \triangleq E\{\theta | \mathcal{F}_t^y\}, t = 1, 2, \cdots$; $K > 0$ *is some constant;* I *denotes the identity matrix and* $\phi_k \triangleq f(\varphi_k)$. *Furthermore, there exists some* $D_1 \subset D_0$ *with* $Prob(D_1) > 0$ *such that for* $t = 0, 1, 2, \cdots$,

$$E[y_{t+1}^2 | \mathcal{F}_t^y] \le (K_1 t^4 + 4) y_{t+1}^2 + (K_1 t^4 + 4)(t+1)^2 + \sigma_w^2 \quad \text{on } D_1,$$

where $K_1 > 0$ *is some constant.*

Proof. Let us denote $x = \{y_1, \cdots, y_t\}$, then $E_x(\cdot) = E[\cdot | \mathcal{F}_t^y]$.
First, since the conditions in Theorem 3.2.5 are satisfied, we have

$$E_x[(\theta - \widehat{\theta}_t)(\theta - \widehat{\theta}_t)^T] \ge \left\{ \frac{1}{\sigma_w^2} \sum_{k=1}^{t} \phi_{k-1} \phi_{k-1}^T - E_x \frac{\partial^2 \log p(\theta)}{\partial \theta^2} \right\}^{-1}.$$

We can choose some $\theta_0 \in \Theta - A_0$ and $r > 0$, such that $B(\theta_0, 2r) \cap (\Theta - A_0) = B(\theta_0, 2r)$ ($B(\theta_0, 2r)$ denotes the ball centered at θ_0 with radius $2r$). Define $\Theta_0 \triangleq B(\theta_0, r)$ and $D_0 \triangleq \bigcap_{k=1}^{\infty} \{\omega : |w_k| \le k\} \cap \{\omega : \theta \in \Theta_0\}$. Since $\sum_{k=1}^{\infty} \text{Prob}(\{\omega : |w_k| > k\}) < \infty$, $\{w_k\}$ and Θ_0 are independent, we have $\text{Prob}(D_0) > 0$.
Now, we only need to prove that

$$-E_x \frac{\partial^2 \log p(\theta)}{\partial \theta^2} \le K t^N I \quad \text{on } D_0,$$

where $K > 0$ is some constant.

Since $\dfrac{\partial^2 p(\theta)}{\partial\theta^2}$ and $\dfrac{1}{p(\theta)}\left(\dfrac{\partial p(\theta)}{\partial\theta}\right)\left(\dfrac{\partial p(\theta)}{\partial\theta}\right)^T$ are bounded, by some simple manipulations we have

$$\frac{\partial^2 \log p(\theta)}{\partial\theta^2} = -\frac{1}{p^2(\theta)}\left(\frac{\partial p(\theta)}{\partial\theta}\right)\left(\frac{\partial p(\theta)}{\partial\theta}\right)^T + \frac{1}{p(\theta)}\frac{\partial^2 p(\theta)}{\partial\theta^2} \geq -\frac{M_2}{p(\theta)}I,$$

where $M_2 > 0$ is some constant. Then

$$-E_x\frac{\partial^2 \log p(\theta)}{\partial\theta^2} \leq M_2 I\cdot E_x\frac{1}{p(\theta)}. \tag{58}$$

Since $p(\theta) > \dfrac{1}{M_1 t^N}$ for $\theta\in\Theta - A_t$, we have

$$\begin{aligned} E_x\frac{1}{p(\theta)} &= \int_\Theta \frac{1}{p(\theta)}p(\theta|x)d\theta = \frac{1}{p(x)}\int_\Theta\frac{1}{p(\theta)}p(x,\theta)d\theta\\ &= \frac{\int_\Theta \frac{1}{p(\theta)}p(x,\theta)d\theta}{\int_\Theta p(x,\theta)d\theta} \leq \frac{\int_{\Theta-A_t} p(x|\theta)d\theta + \int_{A_t} p(x|\theta)d\theta}{\int_{\Theta-A_t} p(\theta)p(x|\theta)d\theta}\\ &\leq M_1 t^N \frac{\int_{\Theta-A_t} p(x|\theta)d\theta + \int_{A_t} p(x|\theta)d\theta}{\int_{\Theta-A_t} p(x|\theta)d\theta}, \end{aligned} \tag{59}$$

where, $p(x|\theta) = \prod_{k=1}^{t}\dfrac{1}{\sqrt{2\pi}\sigma_w}\exp\{-\dfrac{1}{2\sigma_w^2}(y_k - \theta^T\phi_{k-1} - u_{k-1})^2\}$.

For any $\omega_0\in D_0$, we have $\theta(\omega_0)\in\Theta_0$, $|w_k(\omega_0)|\leq k$, $k = 1,2,\cdots$, and

$$p(x|\theta) = \prod_{k=1}^{t}\frac{1}{\sqrt{2\pi}\sigma_w}\exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0) + (\theta(\omega_0) - \theta)^T\phi_{k-1}]^2\}.$$

We take

$$\alpha - \theta(\omega_0) = (\theta - \theta(\omega_0))\frac{r}{r + c\varepsilon_t},$$

where $c > 1$ such that $\alpha\in\Theta - A_t$ for any $\theta\in A_t$. (By the convexity of Θ, we know such $c > 1$ always exists.) Then

$$\begin{aligned} \int_{A_t} p(x|\theta)d\theta = &\int_{A_t}\prod_{k=1}^{t}\frac{1}{\sqrt{2\pi}\sigma_w}\exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0) + (\theta(\omega_0) - \theta)^T\phi_{k-1}]^2\}d\theta\\ \leq &\int_{\Theta-A_t}\prod_{k=1}^{t}\frac{1}{\sqrt{2\pi}\sigma_w}\exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0) + (\theta(\omega_0) - \alpha)^T\phi_{k-1}\\ &+\frac{c\varepsilon_t}{r}(\theta(\omega_0) - \alpha)^T\phi_{k-1}]^2\}d\alpha\cdot\left(\frac{r + c\varepsilon_t}{r}\right)^n. \end{aligned}$$

Next we consider two cases separately.

Case 1. If $[w_k(\omega_0) + (\theta(\omega_0) - \alpha)^T\phi_{k-1}]\times[\frac{c\varepsilon_t}{r}(\theta(\omega_0) - \alpha)^T\phi_{k-1}]\geq 0$, then

$$\begin{aligned} &\exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0) + (\theta(\omega_0) - \alpha)^T\phi_{k-1} + \frac{c\varepsilon_t}{r}(\theta(\omega_0) - \alpha)^T\phi_{k-1}]^2\}\\ \leq\ &\exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0) + (\theta(\omega_0) - \alpha)^T\phi_{k-1}]^2\}. \end{aligned}$$

Case 2. If $[w_k(\omega_0)+(\theta(\omega_0)-\alpha)^T\phi_{k-1}]\times[\frac{c\varepsilon_t}{r}(\theta(\omega_0)-\alpha)^T\phi_{k-1}]<0$, then

$$w_k(\omega_0)\times(\theta(\omega_0)-\alpha)^T\phi_{k-1}<0 \quad \text{and} \quad |(\theta(\omega_0)-\alpha)^T\phi_{k-1}|\le|w_k(\omega_0)|\le k.$$

Thus,

$$\begin{aligned}
& \exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0)+(\theta(\omega_0)-\alpha)^T\phi_{k-1}+\frac{c\varepsilon_t}{r}(\theta(\omega_0)-\alpha)^T\phi_{k-1}]^2\} \\
\le\ & \exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0)+(\theta(\omega_0)-\alpha)^T\phi_{k-1}]^2 \\
& +\frac{1}{\sigma_w^2}[w_k(\omega_0)+(\theta(\omega_0)-\alpha)^T\phi_{k-1}]\cdot\frac{c\varepsilon_t}{r}(\theta(\omega_0)-\alpha)^T\phi_{k-1}\} \\
\le\ & \exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0)+(\theta(\omega_0)-\alpha)^T\phi_{k-1}]^2+\frac{c\varepsilon_t}{\sigma_w^2 r}k^2\} \\
=\ & \exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0)+(\theta(\omega_0)-\alpha)^T\phi_{k-1}]^2\}\cdot\exp\left\{\frac{c\varepsilon_0}{\sigma_w^2 r}\frac{k^2}{t^3}\right\}.
\end{aligned}$$

Combining Case 1 and Case 2, we have

(60)

$$\begin{aligned}
\int_{A_t}p(x|\theta)d\theta \le & \left(\frac{r+c\varepsilon_0}{r}\right)^n\cdot\exp\left\{\frac{c\varepsilon_0}{\sigma_w^2 r}\sum_{k=1}^t\frac{k^2}{t^3}\right\} \\
& \times\int_{\Theta-A_t}\prod_{k=1}^t\frac{1}{\sqrt{2\pi}\sigma_w}\exp\{-\frac{1}{2\sigma_w^2}[w_k(\omega_0)+(\theta(\omega_0)-\alpha)^T\phi_{k-1}]^2\}d\alpha \\
\le & \ M_3\int_{\Theta-A_t}p(x|\theta)d\theta,
\end{aligned}$$

where $M_3>0$ is some constant.

Hence, by (58), (59) and (60), we have

$$-E_x\frac{\partial^2\log p(\theta)}{\partial\theta^2}\le Kt^N I, \quad K=M_2M_1(1+M_3)>0.$$

To prove the second conclusion of this theorem, we define

$$\Delta_t\triangleq\{\theta\in\Theta:|\theta^T\phi_t+u_t|\le\frac{\varepsilon}{t^2}\|\phi_t\|\},\quad 0<\varepsilon<\frac{\phi(D_0)}{2\cdot S\cdot P\sum_{t=0}^{\infty}\frac{1}{t^2}},\quad t=0,1,2,\cdots,$$

where, $P\triangleq\sup_{\theta\in\Theta}p(\theta)$, and $S\triangleq\sup_{L\in\mathcal{L}}V_{n-1}(L\bigcap\Theta)$ with $\mathcal{L}$ denoting the set of all (n-1)-dimensional hyperplane and $V_{n-1}(\cdot)$ denoting the Lebesgue measure on $I\!R^{n-1}$. Since Θ is bounded, we have $S<\infty$.

Recursively define $\Theta_{t+1}\triangleq\Theta_t-\Delta_t$, $t=0,1,\cdots$. Let $\Theta_\infty\triangleq\lim_{t\to\infty}\Theta_t$, $D_1\triangleq\bigcap_{k=1}^{\infty}\{\omega:|w_k|\le k\}\cap\{\omega:\theta\in\Theta_\infty\}$.

Now we prove that $\text{Prob}(D_1)>0$.

For any $\theta\in\Theta$, let $\theta=\theta_1+\theta_2$, where $\theta_1\parallel\phi_t$, $\theta_2\perp\phi_t$. since $\phi_t\ne0$ a.s. by our assumption, we have for some $a\in R^1$,

$$\theta_1=\frac{a\phi_t}{\|\phi_t\|}, \quad \text{and} \quad \theta_2^T\phi_t=0.$$

Hence, $\theta^T\phi_t + u_t = (a + \frac{u_t}{\|\phi_t\|})\|\phi_t\|$ and

$$|\theta^T\phi_t + u_t| \le \frac{\varepsilon}{t^2}\|\phi_t\| \Longleftrightarrow \left|a + \frac{u_t}{\|\phi_t\|}\right| \le \frac{\varepsilon}{t^2}.$$

Hence, $V(\Delta_t) \le S\cdot\frac{2\varepsilon}{t^2}$ and $\text{Prob}(\{\omega : \theta \in \Delta_t\}) \le P\cdot S\cdot\frac{2\varepsilon}{t^2}$. Then by the definition of ε, we have

$$\text{Prob}(\{\omega : \theta \in \bigcup_{t=0}^{\infty}\Delta_t\}) \le \sum_{t=0}^{\infty}\text{Prob}(\{\omega : \theta \in \Delta_t\}) \le P\cdot S\cdot\sum_{t=0}^{\infty}\frac{1}{t^2}\cdot 2\varepsilon < \text{Prob}(D_0).$$

Hence, we have

$$\begin{aligned}
&\text{Prob}(D_1)\\
=\ &\text{Prob}\left(\bigcap_{k=1}^{\infty}\{\omega : |w_k| \le k\}\bigcap\{\omega : \theta \in \Theta_0 - \bigcup_{t=0}^{\infty}\Delta_t\}\right)\\
=\ &\text{Prob}\left(\bigcap_{k=1}^{\infty}\{\omega : |w_k| \le k\}\bigcap\{\omega : \theta \in \Theta_0\} - \bigcap_{k=1}^{\infty}\{\omega : |w_k| \le k\}\right.\\
&\left.\bigcap\{\omega : \theta \in \bigcup_{t=0}^{\infty}\Delta_t\}\ \right)\\
\ge\ &\text{Prob}\left(\bigcap_{k=1}^{\infty}\{\omega : |w_k| \le k\}\bigcap\{\omega : \theta \in \Theta_0\}\right)\\
&-\text{Prob}\left(\bigcap_{k=1}^{\infty}\{\omega : |w_k| \le k\}\bigcap\{\omega : \theta \in \bigcup_{t=0}^{\infty}\Delta_t\}\right)\\
\ge\ &\text{Prob}(D_0) - \text{Prob}\left(\{\omega : \theta \in \bigcup_{t=0}^{\infty}\Delta_t\}\right)\\
>\ &0.
\end{aligned}$$

For any $\omega_1 \in D_1$, we have $|\theta(\omega_1)^T\phi_t + u_t| > \frac{\varepsilon}{t^2}\|\phi_t\|$ and $|w_t(\omega_1)| \le t$. Then by the boundedness of Θ, we have on D_1

$$\begin{aligned}
E_x y_{t+1}^2 &= E_x(\theta^T\phi_t + u_t)^2 + \sigma_w^2\\
&\le 2E_x(\theta(\omega_1)^T\phi_t + u_t)^2 + 2E_x(\theta(\omega_1)^T\phi_t - \theta^T\phi_t)^2 + \sigma_w^2\\
&\le 2(\theta(\omega_1)^T\phi_t + u_t)^2 + M_4\|\phi_t\|^2 + \sigma_w^2\\
&\le \left(2 + \frac{M_4t^4}{\varepsilon^2}\right)(\theta(\omega_1)^T\phi_t + u_t)^2 + \sigma_w^2\\
&\le \left(2 + \frac{M_4t^4}{\varepsilon^2}\right)[y_{t+1} - w_{t+1}(\omega_1)]^2 + \sigma_w^2\\
&\le \left(4 + \frac{2M_4t^4}{\varepsilon^2}\right)y_{t+1}^2 + \left(4 + \frac{2M_4t^4}{\varepsilon^2}\right)(t+1)^2 + \sigma_w^2\\
&\le (K_1t^4 + 4)y_{t+1}^2 + (K_1t^4 + 4)(t+1)^2 + \sigma_w^2,
\end{aligned}$$

where, $M_4 > 0$ is some constant and $K_1 = \frac{2M_4}{\varepsilon^2} > 0$ is a constant. □

Remark 3.2.3. *A typical class of distributions that satisfy the conditions of Theorem 3.2.6 are*

$$p(\theta) = \begin{cases} c\,(2^{1-2k}R^{2k} - \|\theta\|^{2k}) & \text{if } 0 \le \|\theta\| \le R/2;\\ c\,(R - \|\theta\|)^{2k} & \text{if } R/2 < \|\theta\| \le R;\\ 0 & \text{otherwise,}\end{cases}$$

where $R > 0, k \geq 1, c$ *is some constant to make* $\int_{\|\theta\| \leq R} p(\theta) d\theta = 1.$

3.3. On the Non-Stabilizability of Multi-Parameter Case. Consider the following discrete-time nonlinear regression model

$$y_{t+1} = \theta^T f(\varphi_t) + u_t + w_{t+1}, \quad t \geq 0, \tag{61}$$

$$\varphi_t = [y_t, y_{t-1}, \cdots, y_{t-p+1}]^T, \quad p \geq 1, \tag{62}$$

where y_t and u_t are the system output and input signals respectively, $\theta \triangleq [\theta_1, \theta_2, \cdots, \theta_n]^T$ is an unknown parameter vector, $f \triangleq [f_1, f_2, \cdots, f_n]^T$ is a known nonlinear vector function and w_t is the noise signal.

Assume that

(A3.3.1) There exist $M > 0$, $0 < \alpha < \beta < \infty$ and $b_1 > b_2 > \cdots > b_n > 0$, such that for $|x_1| > M$, $|x_j| \leq |x_1|, 2 \leq j \leq p$,

$$\alpha |x_1|^{b_i} \leq |f_i(x_1, x_2, \cdots, x_p)| \leq \beta |x_1|^{b_i}, \quad \forall 1 \leq i \leq n; \tag{63}$$

(A3.3.2) $\{w_t\}$ is a Gaussian white noise sequence with distribution $N(0, \sigma_w^2)$;

(A3.3.3) The unknown parameter vector $\theta = [\theta_1, \cdots, \theta_n]^T$ is independent of $\{w_t\}$ and satisfies the conditions in Theorem 3.2.6.

Our objective is to study the global stabilizability of (61) under the above conditions. First, we give a precise definition of stabilizability.

Definition 3.3.1. *Let* $\sigma\{y_i, 0 \leq i \leq t\}$ *be the* σ*-field generated by the observations* $\{y_i,\ 0 \leq i \leq t\}$. *The system (61) is said to be a.s. globally stabilizable, if there exists a feedback control*

$$u_t \in \mathcal{F}_t^y \triangleq \sigma\{y_i, 0 \leq i \leq t\}, \quad t = 0, 1, \cdots \tag{64}$$

such that for any initial condition $y_0 \in R^1$, $\limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} y_t^2 < \infty$, *a.s.*

Remark 3.3.1. *We remark that the global stabilization of (61) is a trivial task in either the case where* θ *is known or the case where the noise is free (i.e.* $w_t \equiv 0$*). To be precise, if* θ *were known, we can put* $u_t \equiv -\theta^T f(\varphi_t)$, *which gives* $y_{t+1} \equiv w_{t+1}$, *and the system is stabilized since*

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} y_t^2 = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} w_t^2 < \infty.$$

In the case where θ *is unknown but the noise is free* ($w_t \equiv 0$), *we can obtain the true value of the parameter* θ *by solving* n *independent linear equations. For example, if in the first* $(n+1)$ *steps, we choose* $\{u_t, 0 \leq t \leq n\}$ *to be independently identically distributed random variables with probability density function* $p(u)$, *then it is not* *difficult to prove the nonsingularity of the following matrix*[1] *(see Section 3.4)*

$$A \triangleq \begin{bmatrix} f_1(\varphi_1) & f_2(\varphi_1) & \cdots & f_n(\varphi_1) \\ f_1(\varphi_2) & f_2(\varphi_2) & \cdots & f_n(\varphi_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(\varphi_n) & f_2(\varphi_n) & \cdots & f_n(\varphi_n) \end{bmatrix} \tag{65}$$

Hence the true value of the parameter vector θ can easily be obtained by solving the linear equation:

$$A \cdot \theta = [y_2 - u_1, y_3 - u_2, \cdots, y_{n+1} - u_n]^T.$$

Then again we can take the control as $u_t = -\theta^T f(\varphi_t)$ for $t > n$, which globally stabilizes the noise-free system. For more general parametric-strict-feedback models with no noises, similar approaches can also be applied to design a globally stabilizing adaptive controller regardless of the growth rate of the nonlinearities.

Unfortunately, the main drawback of the above approach is that the resulting adaptive controller is not robust with respect to noises. In fact, the presence of noises will even change the stabilizability of discrete-time nonlinear systems dramatically if the growth rate of the nonlinearities is faster than linear, as will be shown by the following theorem together with its corollaries. Their proofs will be given in Section 3.4.

Theorem 3.3.1. *Under Assumptions (A3.3.1)–(A3.3.3), the system (61) is* **not** *a.s. globally stabilizable by feedback whenever the following inequality*

$$P(z) < 0, \quad z \in (1, b_1) \tag{66}$$

has a solution, where $P(z)$ is a polynomial defined by

$$P(z) = z^{n+1} - b_1 z^n + (b_1 - b_2) z^{n-1} + \cdots + (b_{n-1} - b_n) z + b_n. \tag{67}$$

To understand the implications of Theorem 3.3.1, we now give some detailed discussions on the inequality (66).

Corollary 3.3.1. ([**19**]) *Let b_i $(1 \le i \le n)$ satisfy*

$$b_1 > 1 \quad \text{and} \quad 0 < b_i - b_{i+1} \le \frac{\sqrt{b_1}}{2}(\sqrt{b_1} - 1)^2, \quad 1 \le i \le n-1,$$

then (66) has a solution whenever $n \ge 2\log\left(\frac{\sqrt{b_1}+1}{\sqrt{b_1}-1}\right) \Big/ \log b_1$. Consequently, whenever $b_1 > 1$ and the number of unknown parameters n is suitably large, there always exist $0 < b_n < b_{n-1} < \cdots < b_1$ such that (61) is not a.s. globally stabilizable.

Remark 3.3.2. *By Corollary 3.3.1 we know that the usual linear growth condition imposed on the nonlinear function $f(\cdot)$ of the general control model*

$$y_{t+1} = \theta^T f(\varphi_t) + u_t + w_{t+1}, \quad \theta \in R^n \tag{68}$$

cannot be essentially relaxed in general for global adaptive stabilization, unless additional conditions on the number n and the structure of $f(\cdot)$ are imposed.

[1]Here, we assume that for any fixed $x_2, \cdots, x_p$ and any $a_i \in I\!R^1, 1 \le i \le n$ with $\sum_{i=1}^n |a_i| > 0$, $g(x_1) \triangleq \sum_{i=1}^{n} a_i f_i(x_1, x_2, \cdots, x_p)$ has at most countable zeroes on $x_1 \in (-\infty, +\infty)$. By the theory of complex functions, we know this holds for any analytic function $f_i(x), 1 \le i \le n$ with $\sum_{i=1}^n a_i f_i(x) \not\equiv 0$.

Remark 3.3.3. *Let us consider the following counter-part continuous-time model*

$$dy_t = [\theta^T f(y_t) + u_t]dt + dw_t, \quad t \geq 0, \tag{69}$$

where $\theta \in R^n$ is an unknown parameter vector, and $f(x) : R^1 \to R^n$ is a continuous function satisfying the local Lipschitz condition, and $\{w_t\}$ is a standard Brownian motion. Assume that $\|f(x)\| \leq L_1 + L_2|x|^k$ for some integer $k > 0$ and constants $L_1, L_2 > 0$. Then it can be shown (see Section 3.4) that the following feedback control of nonlinear damping type:

$$u_t = -cy_t - y_t^{2k+1}, \quad c > 0$$

can stabilize the systems regardless of the growth rate of the nonlinearities (measured by k).

Corollary 3.3.2 *Let $b_1 > 2$, then for $n > 1 + 2\log\left(\frac{2}{b_1 - 2}\right) \Big/ \log\left(\frac{b_1}{2}\right)$, (66) has a solution for any $\{b_i\}$ satisfying $1 \leq b_n < b_{n-1} < \cdots < b_2 < b_1$. On the other hand, if $b_1 \leq 2$, then for any n, there always exist $1 \leq b_n < b_{n-1} < \cdots < b_2 < b_1$ such that (66) has no solution.*

Corollary 3.3.3 *For any $n \geq 1$ and any $b_1 > b_2 > \cdots > b_n > 0$,*

(i) A necessary condition for (66) to have a solution is $\sum_{i=1}^{n} b_i > 4$;

(ii) A sufficient condition for (66) to have a solution is either $b_1 > 4$, or

$$\sum_{i=1}^{n} b_i > (n+1)(1 + \frac{1}{n})^n.$$

The above three corollaries give us a picture concerning about situations where the nonlinear model (61) is not a.s. globally stabilizable by feedback.

3.4. Proof of the Main Results in Section 3.3

We first present the proof of Theorem 3.3.1, which is prefaced with a lemma.

Lemma 3.4.1 *Assume that for some $\delta > 0$ and $t \geq 1$, $|y_i| \geq |y_{i-1}|^{1+\delta}$, $i = 1, 2, \cdots, t$, and that the initial condition $|y_0| \geq 1$ is sufficiently large, then the determinants of the matrices*

$$P_{t+1}^{-1} \triangleq \sigma_w^2 K(t+1)^N I + \sum_{i=0}^{t} \phi_i \phi_i^T \quad \text{and} \quad Q_{t+1}^{-1} \triangleq \sigma_w^2 K t^N I + \sum_{i=0}^{t} \phi_i \phi_i^T \tag{70}$$

with $\phi_t = [f_1(\varphi_t), f_2(\varphi_t), \cdots, f_n(\varphi_t)]^T$ satisfy

$$|P_{t+1}^{-1}| \leq [(\sigma_w^2 K n^N)^n + \frac{1}{2}] f_1^2(\varphi_t) f_2^2(\varphi_{t-1}) \cdots f_n^2(\varphi_{t-n+1}); \tag{71}$$

$$|Q_{t+1}^{-1}| \geq \frac{1}{2} f_1^2(\varphi_t) f_2^2(\varphi_{t-1}) \cdots f_n^2(\varphi_{t-n+1}), \tag{72}$$

where by definition $f_j(\varphi_i) \triangleq 1$ for $i < 0, 1 \leq j \leq n$, $K > 0$ and $N > 0$ are constants defined in Theorem 3.2.6. (Here, without contradicting to Theorem 3.2.6, we assume $K > 0$ large enough such that $\sigma_w^2 K > 1$.)

Proof. By (70), we have

$$|Q_{t+1}^{-1}| = |\sigma_w^2 K t^N I + \sum_{i=0}^{t} \phi_i \phi_i^T|$$

$$= \begin{vmatrix} \sigma_w^2 K t^N + \sum_{i=0}^{t} f_1^2(\varphi_i) & \sum_{i=0}^{t} f_1(\varphi_i) f_2(\varphi_i) & \cdots & \sum_{i=0}^{t} f_1(\varphi_i) f_n(\varphi_i) \\ \sum_{i=0}^{t} f_1(\varphi_i) f_2(\varphi_i) & \sigma_w^2 K t^N + \sum_{i=0}^{t} f_2^2(\varphi_i) & \cdots & \sum_{i=0}^{t} f_2(\varphi_i) f_n(\varphi_i) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=0}^{t} f_1(\varphi_i) f_n(\varphi_i) & \sum_{i=0}^{t} f_2(\varphi_i) f_n(\varphi_i) & \cdots & \sigma_w^2 K t^N + \sum_{i=0}^{t} f_n^2(\varphi_i) \end{vmatrix}$$

Now, let us denote $\alpha_m(i) \triangleq [f_1(\varphi_i) f_m(\varphi_i), f_2(\varphi_i) f_m(\varphi_i), \cdots f_n(\varphi_i) f_m(\varphi_i)]^T$, $1 \le m \le n$, $0 \le i \le t$, and let $\alpha_m(-1) \triangleq \sigma_w^2 K t^N e_m$ (e_m denotes the m-th column of the identity matrix I_n), then

$$|Q_{t+1}^{-1}| = \det\left(\sum_{i=-1}^{t} \alpha_1(i), \sum_{i=-1}^{t} \alpha_2(i), \cdots, \sum_{i=-1}^{t} \alpha_n(i) \right).$$

By the elementary properties of determinants, we have

$$(73) \qquad |Q_{t+1}^{-1}| = \sum_{i_1, i_2, \cdots, i_n = -1}^{t} \det(\alpha_1(i_1), \alpha_2(i_2), \cdots, \alpha_n(i_n)).$$

It is clear that if in the group $(i_1, i_2, \cdots, i_n)$ there are at least two integers having the same value (but different from -1), then $\det(\alpha_1(i_1), \cdots, \alpha_n(i_n)) = 0$. So in the discussions below we will exclude this kind of zero-valued determinants.

We proceed to prove (72) by considering two cases separately:

Case (I) $t < n - 1$

In this case, in order that $\det(\alpha_1(i_1), \alpha_2(i_2), \cdots, \alpha_n(i_n)) \neq 0$, the number of (-1)'s in $(i_1, i_2, \cdots, i_n)$ must at least be $(n - 1 - t)$ and the other integers must be distinct. Then each term in the expansion of the non-zero determinant

$$\det(\alpha_1(i_1), \alpha_2(i_2), \cdots, \alpha_n(i_n))$$

has the general form

(74)
$$(\sigma_w^2 K t^N)^{n-r} f_{j_1}(\varphi_{i_1}) f_{k_1}(\varphi_{i_1}) \cdot f_{j_2}(\varphi_{i_2}) f_{k_2}(\varphi_{i_2}) \cdots f_{j_r}(\varphi_{i_r}) f_{k_r}(\varphi_{i_r}), \quad r \le t+1,$$

where, $i_m \neq i_l$, $j_m \neq j_l$ and $k_m \neq k_l$ for $m \neq l$. Note that one of such terms is $(\sigma_w^2 K t^N)^{n-t-1} f_1^2(\varphi_t) f_2^2(\varphi_{t-1}) \cdots f_{t+1}^2(\varphi_0)$ (from the products of the main diagonal elements of the matrix $[\alpha_1(t), \cdots, \alpha_{t+1}(0), \alpha_{t+2}(-1), \cdots, \alpha_n(-1)]$), and it is different from other terms. Denote $a_t \triangleq f_1^2(\varphi_t) f_2^2(\varphi_{t-1}) \cdots f_{t+1}^2(\varphi_0)$. Now, we proceed to prove that the absolute value of any other term of the form (74) is not greater than

$$\frac{a_t}{|y_0|^{\min(1,\delta) \cdot \min_{1 \le i \le n-1}(b_i - b_{i+1})}} \cdot \left(\frac{\beta}{\alpha}\right)^{2n} \cdot (\sigma_w^2 K n^N)^n,$$

where $\alpha, \beta, b_i, 1 \le i \le n$ are defined in Assumption (A3.3.1).

From the proof of Lemma 3.2 in [**19**], we know that

$$\left|y_{i_1}^{b_{j_1}+b_{k_1}}y_{i_2}^{b_{j_2}+b_{k_2}}\cdots y_{i_r}^{b_{j_r}+b_{k_r}}\right| \le \frac{1}{|y_0|^{\min(1,\delta)\cdot\min_{1\le i\le n-1}(b_i-b_{i+1})}}y_t^{2b_1}y_{t-1}^{2b_2}\cdots y_0^{2b_{t+1}}$$

for $y_{i_1}^{b_{j_1}+b_{k_1}}y_{i_2}^{b_{j_2}+b_{k_2}}\cdots y_{i_r}^{b_{j_r}+b_{k_r}} \neq y_t^{2b_1}y_{t-1}^{2b_2}\cdots y_0^{2b_{t+1}}$.

Then by Assumption (A3.3.1) and $|y_i| \ge |y_{i-1}|^{1+\delta}, i = 1, 2, \cdots, t$, we have for $|y_0| > \max\{M, 1\}$,

$$\begin{aligned}
&\left|(\sigma_w^2 Kt^N)^{n-r} f_{j_1}(\varphi_{i_1})f_{k_1}(\varphi_{i_1})\cdot f_{j_2}(\varphi_{i_2})f_{k_2}(\varphi_{i_2})\cdots f_{j_r}(\varphi_{i_r})f_{k_r}(\varphi_{i_r})\right| \\
\le\;& \beta^{2n}(\sigma_w^2 Kt^N)^n \left|y_{i_1}^{b_{j_1}+b_{k_1}}y_{i_2}^{b_{j_2}+b_{k_2}}\cdots y_{i_r}^{b_{j_r}+b_{k_r}}\right| \\
\le\;& \beta^{2n}(\sigma_w^2 Kt^N)^n \frac{1}{|y_0|^{\min(1,\delta)\cdot\min_{1\le i\le n-1}(b_i-b_{i+1})}}y_t^{2b_1}y_{t-1}^{2b_2}\cdots y_0^{2b_{t+1}} \qquad (75)\\
\le\;& \left(\frac{\beta}{\alpha}\right)^{2n}(\sigma_w^2 Kn^N)^n \frac{a_t}{|y_0|^{\min(1,\delta)\cdot\min_{1\le i\le n-1}(b_i-b_{i+1})}}.
\end{aligned}$$

Now, rewrite (73) as

$$|Q_{t+1}^{-1}| = R_t + (\sigma_w^2 Kt^N)^{n-t-1}a_t \ge R_t + a_t,$$

where R_t denotes the summation of all the terms different from $(\sigma_w^2 Kt^N)^{n-t-1}a_t$. It is obvious that R_t has at most $[(t+2)^n\cdot n! - 1] \le (n^n\cdot n! - 1)$ terms. Hence, by (75) we obtain

$$|R_t| \le \frac{(n^n\cdot n! - 1)a_t}{|y_0|^{\min(1,\delta)\min_{1\le i\le n-1}(b_i-b_{i+1})}}\left(\frac{\beta}{\alpha}\right)^{2n}(\sigma_w^2 Kn^N)^n.$$

Therefore, by choosing the initial value $|y_0|$ large enough we can make $|R_t|$ less than $\frac{1}{2}a_t$. Consequently (72) follows.

Case (II) $t \ge n-1$

First of all, any nonzero determinant

$$\det(\alpha_1(i_1), \alpha_2(i_2), \cdots, \alpha_n(i_n)),\ -1 \le i_1, \cdots, i_n \le t$$

can be expanded as the summation of $n!$ terms whose general form is

$$(\sigma_w^2 Kt^N)^{n-r} f_{j_1}(\varphi_{i_1})f_{k_1}(\varphi_{i_1})\cdot f_{j_2}(\varphi_{i_2})f_{k_2}(\varphi_{i_2})\cdots f_{j_r}(\varphi_{i_r})f_{k_r}(\varphi_{i_r}), \quad r\le n,$$

where, $i_m \ne i_l$, $j_m \ne j_l$ and $k_m \ne k_l$ for $m \ne l$. Obviously one of such terms is $a_t \triangleq f_1^2(\varphi_t)f_2^2(\varphi_{t-1})\cdots f_n^2(\varphi_{t-n+1})$ (one term in $\det(\alpha_1(t), \cdots, \alpha_n(t-n+1))$). We now show that for any other terms, the following inequality holds:

$$\begin{aligned}
&\left|(\sigma_w^2 Kt^N)^{n-r} f_{j_1}(\varphi_{i_1})f_{k_1}(\varphi_{i_1})\cdot f_{j_2}(\varphi_{i_2})f_{k_2}(\varphi_{i_2})\cdots f_{j_r}(\varphi_{i_r})f_{k_r}(\varphi_{i_r})\right| \\
\le\;& \frac{a_t}{|y_{t-n+1}|^{\frac{\delta}{1+\delta}\min_{1\le i\le n}(b_i-b_{i+1})}}\cdot\left(\frac{\beta}{\alpha}\right)^{2n}\cdot(\sigma_w^2 Kt^N)^n, \qquad (76)
\end{aligned}$$

where $b_{n+1} \triangleq 0$.

From the proof of Lemma 3.2 of [**19**], we know that

$$\left|y_{i_1}^{b_{j_1}+b_{k_1}}y_{i_2}^{b_{j_2}+b_{k_2}}\cdots y_{i_r}^{b_{j_r}+b_{k_r}}\right| \le \frac{1}{|y_{t-n+1}|^{\frac{\delta}{1+\delta}\min_{1\le i\le n}(b_i-b_{i+1})}}y_t^{2b_1}\cdots y_{t-n+1}^{2b_n}.$$

for $y_{i_1}^{b_{j_1}+b_{k_1}}y_{i_2}^{b_{j_2}+b_{k_2}}\cdots y_{i_r}^{b_{j_r}+b_{k_r}} \neq y_t^{2b_1}\cdots y_{t-n+1}^{2b_n}$. Then similarly as in Case (I), by Assumption (A3.3.1) we get (76) for $|y_0| > \max\{M, 1\}$.

Thus, similar to the arguments in Case (I), we rewrite (73) as

$$|Q_{t+1}^{-1}| = R_t + a_t,$$

where R_t denotes the summation of $[(t+2)^n \cdot n! - 1]$ terms in the determinant expansions, which are different from a_t. Then by (76) we know that

$$\begin{aligned}|R_t| &\leq \frac{[(t+2)^n \cdot n! - 1] \cdot a_t}{|y_{t-n+1}|^{\frac{\delta}{1+\delta}\min_{1\leq i\leq n}(b_i - b_{i+1})}} \left(\frac{\beta}{\alpha}\right)^{2n} \cdot (\sigma_w^2 K t^N)^n \\ &\leq \frac{[(t+2)^n \cdot n! - 1] \cdot a_t}{|y_0|^{(1+\delta)^{t-n+1}\cdot\frac{\delta}{1+\delta}\cdot\min_{1\leq i\leq n}(b_i - b_{i+1})}} \left(\frac{\beta}{\alpha}\right)^{2n} \cdot (\sigma_w^2 K t^N)^n \\ &\leq \frac{1}{2} a_t,\end{aligned}$$

where the last inequality holds for sufficiently large $|y_0|$. Hence, (72) is proved.

To prove (71), noting the definition (70), we only need to replace $\sigma_w^2 K t^N$ by $\sigma_w^2 K (t+1)^N$ in the proof above. □

Proof of Theorem 3.3.1.

We only need to prove that if the inequality (66) has a solution, then for any feedback control $u_t \in \mathcal{F}_t^y$, there always exist an initial condition y_0 and a set D_1 with positive probability such that the output signal y_t of the closed-loop control system tends to infinity at a rate faster than exponential on D_1.

Since by Assumption (A3.3.3) Theorem 3.2.6 holds, we choose D_1 to be defined as in Theorem 3.2.6. Then on D_1, we have

$$E[(\theta - \widehat{\theta}_t)(\theta - \widehat{\theta}_t)^T | \mathcal{F}_t^y] \geq \left\{ \frac{1}{\sigma_w^2} \sum_{k=0}^{t-1} \phi_k \phi_k^T + K t^N I \right\}^{-1}, \quad t \geq 1 \tag{77}$$

equation and

$$E[y_{t+1}^2 | \mathcal{F}_t^y] \leq (K_1 t^4 + 4) y_{t+1}^2 + (K_1 t^4 + 4)(t+1)^2 + \sigma_w^2, \quad t \geq 0, \tag{78}$$

where $\widehat{\theta}_t \triangleq E[\theta | \mathcal{F}_t^y]$, $K > 0$ and $K_1 > 0$ are constants such that $\sigma_w^2 K > 1$.

Next, by (61) we know that

$$y_{t+1} = \phi_t^T \widetilde{\theta}_t + (\phi_t^T \widehat{\theta}_t + u_t) + w_{t+1}, \tag{79}$$

where $\widetilde{\theta}_t \triangleq \theta - \widehat{\theta}_t$ and $\phi_t = [f_1(\varphi_t), f_2(\varphi_t), \cdots, f_n(\varphi_t)]^T$. Consequently, by the fact that $E[\widetilde{\theta}_t | \mathcal{F}_t^y] = 0$ and $E[w_{t+1} | \mathcal{F}_t^y] = 0$ it follows that for any $u_t \in \mathcal{F}_t^y$,

$$\begin{aligned}E[y_{t+1}^2 | \mathcal{F}_t^y] &= \phi_t^T E[\widetilde{\theta}_t \widetilde{\theta}_t^T | \mathcal{F}_t^y] \phi_t + (\phi_t^T \widehat{\theta}_t + u_t)^2 + \sigma_w^2 \\ &\geq \phi_t^T E[\widetilde{\theta}_t \widetilde{\theta}_t^T | \mathcal{F}_t^y] \phi_t + \sigma_w^2.\end{aligned}$$

Then by (77), we have on D_1,

$$\begin{aligned}E[y_{t+1}^2 | \mathcal{F}_t^y] &\geq \phi_t^T \left\{ \frac{1}{\sigma_w^2} \sum_{k=0}^{t-1} \phi_k \phi_k^T + K t^N I \right\}^{-1} \phi_t + \sigma_w^2 \\ &= \sigma_w^2 \phi_t^T P_t \phi_t + \sigma_w^2 = \sigma_w^2 \frac{|P_t^{-1} + \phi_t \phi_t^T|}{|P_t^{-1}|} \\ &= \sigma_w^2 \frac{|Q_{t+1}^{-1}|}{|P_t^{-1}|}, \qquad t \geq 1,\end{aligned} \tag{80}$$

where P_t, Q_t are defined by (70).

Hence by (78) and (80), we have on D_1,

$$y_{t+1}^2 \geq \frac{1}{K_1 t^4+4}\left[\sigma_w^2 \frac{|Q_{t+1}^{-1}|}{|P_t^{-1}|} - (K_1 t^4+4)(t+1)^2 - \sigma_w^2\right], \quad t \geq 1. \tag{81}$$

Now, let $z_0 \in (1, b_1)$ be a solution of the inequality (66). We proceed to prove that on D_1,

$$|y_i| \geq |y_{i-1}|^{z_0}, \quad i = 1, 2, \cdots. \tag{82}$$

We adopt the induction argument.

First, we consider the case where $i = 1$. Since

$$E[(\theta-\widehat{\theta}_0)(\theta-\widehat{\theta}_0)^T|\mathcal{F}_0^y] = E[(\theta-\widehat{\theta}_0)(\theta-\widehat{\theta}_0)^T] \geq \sigma_\theta^2 I,$$

we have by (3.3) $E[y_1^2|\mathcal{F}_0^y] \geq \sigma_\theta^2\|\phi_0\|^2 + \sigma_w^2$. Then by (78),

$$y_1^2 \geq \frac{1}{4}(\sigma_\theta^2\|\phi_0\|^2 - 4) \geq \frac{\sigma_\theta^2}{4} f_1^2(\varphi_0) - 1.$$

Then by Assumption (A3.3.1), we have $|y_1| > |y_0|^{z_0}$ for large $|y_0|$ satisfying $|y_0| > M$ and $\left(\frac{\sigma_\theta^2\alpha^2}{4} y_0^{2b_1-2z_0} - y_0^{-2z_0}\right) > 1$. Hence (82) is true for $i = 1$.

Now let us assume that for some $t \geq 1$,

$$|y_i| \geq |y_{i-1}|^{z_0}, \quad i = 1, 2, \cdots, t, \text{ on } D_1,$$

then by Lemma 3.4.1, it follows that

$$|P_t^{-1}| \leq [(\sigma_w^2 K n^N)^n + \frac{1}{2}] f_1^2(\varphi_{t-1}) f_2^2(\varphi_{t-2}) \cdots f_n^2(\varphi_{t-n})$$

and

$$|Q_{t+1}^{-1}| \geq \frac{1}{2} f_1^2(\varphi_t) f_2^2(\varphi_{t-1}) \cdots f_n^2(\varphi_{t-n+1}).$$

Consequently, by (81) and Assumption (A3.3.1) we have for $|y_0| \geq \max\{M, 1\}$,

(83)

$$\begin{aligned} y_{t+1}^2 &\geq \frac{1}{K_1 t^4+4}\left[\sigma_w^2 \frac{|Q_{t+1}^{-1}|}{|P_t^{-1}|} - (K_1 t^4+4)(t+1)^2 - \sigma_w^2\right] \\ &\geq \frac{1}{K_1 t^4+4}\left\{K_2 \frac{f_1^2(\varphi_t)\cdots f_n^2(\varphi_{t-n+1})}{f_1^2(\varphi_{t-1})\cdots f_n^2(\varphi_{t-n})} - (K_1 t^4+4)(t+1)^2 - \sigma_w^2\right\} \\ &\geq \frac{1}{K_1 t^4+4}\left\{K_2 \left(\frac{\alpha}{\beta}\right)^{2n} \frac{y_t^{2b_1}\cdots y_{t-n+1}^{2b_n}}{y_{t-1}^{2b_1}\cdots y_{t-n}^{2b_n}} - (K_1 t^4+4)(t+1)^2 - \sigma_w^2\right\}, \end{aligned}$$

where $K_2 \triangleq \frac{\sigma_w^2}{2[(\sigma_w^2 K n^N)^n + \frac{1}{2}]} > 0$.

However, by the induction assumption we have

$$|y_{t-i}| \leq |y_t|^{z_0^{-i}}, \quad 1 \leq i \leq t,$$

and so by $b_{i+1} - b_i < 0$ $(1 \leq i \leq n)$,

$$|y_{t-i}|^{b_{i+1}-b_i} \geq |y_t|^{(b_{i+1}-b_i)z_0^{-i}}, \quad 1 \leq i \leq n. \tag{84}$$

Note that this inequality also holds for $i > t$, since $y_j \triangleq 1$ for $j < 0$ by definition.

Hence, on D_1,

$$\frac{y_t^{2b_1}\cdots y_{t-n+1}^{2b_n}}{y_{t-1}^{2b_1}\cdots y_{t-n}^{2b_n}} \begin{aligned}[t] &\geq y_t^{2[b_1+(b_2-b_1)z_0^{-1}+\cdots+(b_n-b_{n-1})z_0^{-(n-1)}-b_n z_0^{-n}]} \\ &\geq y_t^{-2z_0^{-n}P(z_0)}\cdot y_t^{2z_0} \\ &\geq [y_0^{-2z_0^{-n}P(z_0)}]^{z_0^t}\cdot y_t^{2z_0}. \end{aligned} \tag{85}$$

At last, since $-z_0^{-n}P(z_0) > 0$, by (83) and (85) we have $y_{t+1}^2 \geq y_t^{2z_0}$ for sufficiently large $|y_0|$.

Hence, by induction, (82) is true. Thus for all large initial conditions $|y_0|$, the output process $|y_t|$ diverges to infinity at a rate faster than exponential on D_1. This completes the proof of Theorem 3.3.1. □

Proof of Corollary 3.3.2.

We first consider the inequality

$$z^{n+1} - b_1 z^n + (b_1-1)z^{n-1} + 1 < 0, \quad z \in (1, b_1). \tag{86}$$

By taking $z = \frac{b_1}{2} \in (1, b_1)$, we have after some simple manipulations

$$\begin{aligned} & z^{n+1} - b_1 z^n + (b_1-1)z^{n-1} + 1 \\ = & -\left[b_1^2 + 4 - 4b_1 - 4(\frac{2}{b_1})^{n-1}\right]\frac{b_1^{n-1}}{2^{n+1}} \\ = & -\left[(b_1-2)^2 - 4(\frac{2}{b_1})^{n-1}\right]\frac{b_1^{n-1}}{2^{n+1}} < 0, \end{aligned}$$

where the last inequality is guaranteed by the assumption on n. Hence, $z = \frac{b_1}{2}$ is a solution of (86).

Next, let $z \in (1, b_1)$ be any solution of (86). Then by (67) we know that

$$\begin{aligned} P(z) &\leq z^{n+1} - b_1 z^n + z^{n-1}[(b_1-b_2)+\cdots+(b_{n-1}-b_n)] + b_n \\ &= z^{n+1} - b_1 z^n + z^{n-1}(b_1-1) + z^{n-1}(1-b_n) + b_n \\ &\leq z^{n+1} - b_1 z^n + z^{n-1}(b_1-1) + (1-b_n) + b_n \\ &= z^{n+1} - b_1 z^n + z^{n-1}(b_1-1) + 1 < 0 \end{aligned}$$

Therefore, (66) has a solution.

For the proof of the second part of the corollary, we use contradiction argument. For $b_1 \leq 2$, if the second assertion were not true, then there would exist some n such that (66) would have a solution for any $b_1 > b_2 > \cdots > b_n \geq 1$.

For any small $\varepsilon > 0$, let $z_\varepsilon \in (1, b_1)$ be a solution of (66) corresponding to the following choice (note that $b_1 > 1$): $b_2 = 1 + (n-2)\varepsilon$, $b_3 = 1 + (n-3)\varepsilon$, $\cdots$, $b_n = 1$. Then we would have from (66)

$$P(z_\varepsilon) = z_\varepsilon^{n+1} - b_1 z_\varepsilon^n + z_\varepsilon^{n-1}[b_1 - 1 - (n-2)\varepsilon] + \varepsilon(z_\varepsilon^{n-2} + \cdots + z_\varepsilon) + 1 < 0.$$

Let $\varepsilon \to 0$ and z_0 be a limiting point of z_ε, we get $z_0^{n+1} - b_1 z_0^n + z_0^{n-1}(b_1-1) + 1 \leq 0$, which implies that

$$z_0^{n+1} - b_1 z_0^n + z_0^{n-1}(b_1-1) < 0,$$

or $z_0^2 - b_1 z_0 + b_1 - 1 < 0$, or $(z_0-1)(z_0+1-b_1) < 0$, which is impossible since $z_0 \geq 1$ and $b_1 \leq 2$. Hence, the proof of Corollary 3.3.2 is completed. □

Proof of Corollary 3.3.3.

(i) First, we rewrite the polynomial $P(z)$ defined by (67) as

$$P(z) = z^{n+1} - b_1 z^{n-1}(z-1) - \cdots - b_n(z-1). \tag{87}$$

Next, denote $b \triangleq \sum_{i=1}^{n} b_i$, and let $z_0 \in (1, b_1)$ be a solution of $P(z) < 0$. Then (87) implies

$$\begin{aligned} 0 > P(z_0) &> z_0^{n+1} - z_0^{n-1}[b_1(z_0-1) + \cdots + b_n(z_0-1)] \\ &= z_0^{n-1}(z_0^2 - bz_0 + b), \end{aligned}$$

or $z^2 - bz + b < 0$ has a solution $z_0 \in (1, b_1)$. This necessarily implies that $b > 4$.

(ii) If $b_1 > 4$, then $z^{n+1} - b_1 z^n + b_1 z^{n-1} < 0$ has a solution $z_0 \in (1, b_1)$. This combined with (87) easily induces $P(z_0) < 0$.

Moreover, if $\sum_{i=1}^n b_i > (n+1)(1+\frac{1}{n})^n$, then we have

$$\frac{n}{n+1}\left(\frac{b}{n+1}\right)^{1/n} > 1, \quad b \triangleq \sum_{i=1}^{n} b_i. \tag{88}$$

Let us denote $z_0 = [b/(n+1)]^{1/n}$, then (88) implies that $z_0 > 1$. Moreover by the fact that $b \le nb_1$ we have $1 < z_0 \le [nb_1/(n+1)]^{1/n} b_1^{1/n} < b_1^{1/n} \le b_1$. Hence $z_0 \in (1, b_1)$. Furthermore, by (88) it is easy to check $z_0^{n+1} - bz_0 + b < 0$. Consequently, by (87) and the fact that $z_0 > 1$ we have $0 > z_0^{n+1} - bz_0 + b > P(z_0)$. Therefore, Corollary 3.3.3 holds. □

Proof of the nonsingularity of the matrix A in (65).

We write the k-th main sub-matrix of A as $A^{(k)}$, i.e., $A^{(k)}$ is defined in a similar way as A in (65) but with n replaced by k. Clearly, $A = A^{(n)}$. We adopt the induction argument to prove $\text{Prob}\{\det(A^{(k)}) = 0\} = 0$ for any $1 \le k \le n$.

First, for $k = 1$, $A^{(1)} = [f_1(\varphi_1)]$. Since $\phi(f_1(\varphi_1) = 0) = 0$ by the choice of u_0 and the countability of the zeroes of $f_1(x)$, we have $\text{Prob}\{\det(A^{(1)}) = 0\} = 0$.

Second, if $\text{Prob}\{\det(A^{(k)}) = 0\} = 0$ for some $k \ge 1$, then we use the contradiction argument to prove that $\text{Prob}\{\det(A^{(k+1)}) = 0\} = 0$. Suppose that $\text{Prob}\{\det(A^{(k+1)}) = 0\} > 0$, then noticing $\text{Prob}\{\det(A^{(k)}) = 0\} = 0$, we know that there exist random variables $a_1, a_2, \cdots, a_k$ and a set B with positive probability, on which

$$\begin{cases} f_{k+1}(\varphi_1) &= a_1 f_1(\varphi_1) + \cdots + a_k f_k(\varphi_1) \\ f_{k+1}(\varphi_2) &= a_1 f_1(\varphi_2) + \cdots + a_k f_k(\varphi_2) \\ \quad\vdots & \qquad\qquad \vdots \\ f_{k+1}(\varphi_{k+1}) &= a_1 f_1(\varphi_{k+1}) + \cdots + a_k f_k(\varphi_{k+1}). \end{cases}$$

Since $\det(A^{(k)}) \ne 0$ a.s., we can solve the first k equations for the values of $a_1, a_2, \cdots, a_k$. Hence $a_i \in \mathcal{F}_k^y = \sigma\{y_0, y_1, \cdots, y_k\}$, $i = 1, \cdots, k$. Now denote $C \triangleq \{\omega : f_{k+1}(\varphi_{k+1}) = a_1 f_1(\varphi_{k+1}) + \cdots + a_k f_k(\varphi_{k+1})$. By the independence of u_k and $\mathcal{F}_k^y$, the fact that $a_i \in \mathcal{F}_k^y$, $i = 1, \cdots, k$ and the properties of conditional expectation, we have

$$\begin{aligned} \text{Prob}(C) &= E[E[I_C | \mathcal{F}_k^y]] \\ &= E\int_{-\infty}^{\infty} I_{[g_k(\theta^T\phi_k + u) = 0]} \times p(u) \cdot du \\ &= E\int_{-\infty}^{\infty} I_{[g_k(x) = 0]} \times p(x - \theta^T\phi_k) \cdot dx, \end{aligned} \tag{89}$$

where $g_k(x) \triangleq \sum_{i=1}^{k} a_i f_i(x, y_k, \cdots, y_{k-p+2}) - f_{k+1}(x, y_k, \cdots, y_{k-p+2})$, $p(\cdot)$ is the density function of u_k, $\theta^T \phi_k = \theta_1 f_1(\varphi_k) + \cdots + \theta_n f_n(\varphi_k)$ and

$$I_C(\omega) \triangleq \begin{cases} 1, & \text{if } \omega \in C, \\ 0, & \text{otherwise.} \end{cases}$$

Now, since by our assumption, for any fixed $x_2, \cdots, x_p$,

$$g(x_1) = a_1 f_1(x_1, x_2, \cdots, x_p) + \cdots + a_k f_k(x_1, x_2, \cdots, x_p) - f_{k+1}(x_1, x_2, \cdots, x_p)$$

has at most countable real zeroes on $x_1 \in (-\infty, \infty)$, the integral in (89) must be zero, or $\text{Prob}(C) = 0$. This contradicts to the fact that $\text{Prob}(C) \geq \text{Prob}(B) > 0$ by our definitions of B and C. Hence $\text{Prob}\{\det(A^{(k+1)}) = 0\} = 0$, and the proof of the nonsingularity of A is completed. □

Proof of Remark 3.3.3.

By Lemma 2.2 of [**7**] and the fact that

$$(90) \qquad x\,[\,\theta^T f(x) - x^{2k+1}\,] \leq \|\theta\| \cdot \sup_{|x| \leq M} \|xf(x)\| \leq M^2 \sup_{|x|\leq M} \|f(x)\|, \quad \forall x \in R^1,$$

where $M \triangleq \max\{\|\theta\|, L_1 + L_2 + 1\}$, it is easy to know that the closed-loop equation

$$dy_t = [-cy_t + \theta^T f(y_t) - y_t^{2k+1}]dt + dw_t$$

has a unique strong solution on $[0, \infty)$.

Now, by the Ito formula, we have

$$dy_t^2 = [-2cy_t^2 + 2y_t \theta^T f(y_t) - 2y_t^{2k+2}]dt + 2y_t dw_t + dt.$$

So, by (90) again we obtain

$$(91) \qquad y_t^2 \leq y_0^2 - 2c \int_0^t y_s^2 ds + (2M^2 \sup_{|x|\leq M} \|f(x)\| + 1)t + 2\int_0^t y_s dw_s.$$

Since by Lemma 12.3 of [**4**], we know that for any $\varepsilon > 0$,

$$\int_0^t y_s dw_s = o\left(\left\{\int_0^t y_s^2 ds\right\}^{\frac{1}{2}+\varepsilon}\right), \quad \text{a.s.},$$

we conclude from (91) that

$$(92) \qquad \limsup_{t\to\infty} \frac{1}{t}\int_0^t y_s^2 ds \leq \frac{1}{2c}(1 + 2M^2 \sup_{|x|\leq M} \|f(x)\|), \quad \text{a.s.}.$$

Then we only need to show that the control energy is finite. For this, again using Ito's formula, we have

$$\begin{aligned} dy_t^{2k+2} = \; & [(2k+2)y_t^{2k+1}(\theta^T f(y_t) - cy_t - y_t^{2k+1}) \\ & + (k+1)(2k+1)y_t^{2k}]dt + [(2k+2)y_t^{2k+1}]dw_t. \end{aligned}$$

thus,

$$\begin{aligned} y_t^{2k+2} = \; & y_0^{2k+2} + \int_0^t [(2k+2)y_s^{2k+1}(\theta^T f(y_s) - cy_s - y_s^{2k+1}) \\ & + (k+1)(2k+1)y_s^{2k}]ds + \int_0^t (2k+2)y_s^{2k+1} dw_s \\ \leq \; & y_0^{2k+2} + \int_0^t (-\lambda_1 y_s^{4k+2} + \lambda_2)ds \end{aligned}$$

for some $\lambda_1,\ \lambda_2 > 0$. The last inequality follows from our assumption $\|f(x)\| \leq L_1 + L_2|x|^k$ and Lemma 12.3 of [**4**],

$$\int_0^t y_s^{2k+1} dw_s = o\left(\left\{\int_0^t y_s^{4k+2} ds\right\}^{\frac{1}{2}+\varepsilon}\right).$$

Hence,

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t y_s^{4k+2} ds \leq \frac{\lambda_2}{\lambda_1}.$$

This together with (92) ensure that

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t u_s^2 ds < \infty.$$

Hence the system (69) is globally stabilizable regardless of the growth rate of $\|f(x)\|$ as $|x| \to \infty$.

4. Hybrid Control Systems

Many practical systems are naturally modelled by continuous-time models via the classical physical laws. However, the input signals are usually generated by the digital computers in practical implementations. Hence, it is of considerable interests to consider hybrid dynamical systems where the process involves in continuous-time, while the control algorithm is implemented in discrete-time.

Now, consider the continuous-time control system:

$$\frac{dy_t}{dt} = \theta \cdot f(y_t) + u_t, \tag{93}$$

where θ is an unknown parameter. Assume that $f(x)$ is continuous and

$$|f(x)| \leq M|x|^b, \quad b \geq 1, \quad x \in R^1 \tag{94}$$

The typical nonlinear damping controller for system (93) is

$$u_t = -\text{sgn}(y_t) \cdot |y_t|^{b+\varepsilon}, \quad \forall \varepsilon > 0. \tag{95}$$

If the output signal y_t can be measured continuously and the control signal u_t can be updated continuously, then it is not difficult to prove the global stability of the closed-loop system (93)– (95) (cf. [**15**]).

However, in many applications, we can only get the sampled data of the output signals and implement the digital controller. Therefore, we consider the control input like:

$$u_t = -\text{sgn}(y_{nh}) \cdot |y_{nh}|^{b+\varepsilon}, \quad t \in [nh, (n+1)h), \quad n = 0, 1, 2, \cdots, \tag{96}$$

where h denotes the sampling period. Nevertheless, in this way the hybrid control system will no longer be globally stable.

Theorem 4.1. *The continuous-time system (93) -(94) with the digital control (96), is* **not** *globally stable no matter how small the sampling period h is. That is, for any $h > 0$, there always exists some y_0 with $|y_0| > 1$ large enough such that*

$$\limsup_{t \to +\infty} |y_t| = \infty.$$

In order to prove this theorem we first establish an auxiliary result.

Lemma 4.1. *Consider the one dimensional autonomous system $\dot{x} = f(x)$, where $f(x)$ depends explicitly on x only. If $f(x)$ satisfies the following condition: On any interval $(a, b) \in \mathbb{R}$, there exists a sub-interval $(a', b') \subset (a, b)$, on which the sign of $f(x)$ does not change, then any solution of the system $\dot{x} = f(x)$ is monotonously increasing or decreasing.*

proof. We adopt the contradiction argument.

Suppose for a solution $x(t)$, there exist three points $t_1 < t_2 < t_3$, such that $x(t_1) < x(t_2)$ and $x(t_3) < x(t_2)$.

Without loss of the generality, we suppose $x(t_1) \leq x(t_3)$.

By the condition of the lemma, there must exist some interval $(x_1, x_2) \subset (x(t_3), x(t_2))$, such that

$$\begin{array}{lll} & f(x) > 0 & \text{for any } x \in (x_1, x_2) \\ \text{or} & f(x) < 0 & \text{for any } x \in (x_1, x_2) \\ \text{or} & f(x) = 0 & \text{for any } x \in (x_1, x_2). \end{array} \tag{97}$$

It is obvious from the system equation that the solution $x(t)$ must be continuous. Hence, by

$$x(t_1) \leq x(t_3) \leq x_1 < x_2 \leq x(t_2),$$

we know that there exist t_1', t_2' $(t_1 \leq t_1' < t_2' \leq t_2)$, such that

$$x(t_1') = x_1, \quad x(t_2') = x_2, \quad x_1 < x(t) < x_2, \quad \forall t \in (t_1', t_2'); \tag{98}$$

and also that there exist t_2'', t_3' $(t_2 \leq t_2'' < t_3' \leq t_3)$, such that

$$x(t_2'') = x_2, \quad x(t_3') = x_1, \quad x_1 < x(t) < x_2, \quad \forall t \in (t_2'', t_3'). \tag{99}$$

By the system equation, we have

$$x(t_2') - x(t_1') = \int_{t_1'}^{t_2'} f(x(t))dt,$$

then by (98),

$$x_2 - x_1 = \int_{t_1'}^{t_2'} f(x(t))dt.$$

On the other hand, we have

$$x(t_3') - x(t_2'') = \int_{t_2''}^{t_3'} f(x(t))dt,$$

and by (99),

$$x_1 - x_2 = \int_{t_3'}^{t_2''} f(x(t))dt.$$

Thus by (97), we have that $x_2 - x_1$ and $x_1 - x_2$ are both positive or negative or zero, which contradicts our definition of x_1 and x_2. Hence our supposition is incorrect and any solution must be monotonous. □

Remark 4.1. *Any piece-wise continuous function $f(x)$ satisfies the condition in Lemma 4.1.*

Proof of Theorem 4.1.

Let us first assume that for some $n \in \{0, 1, 2, \cdots\}$

(i) $|y_{nh}| > 1$;

(ii) $|y_{nh}|^{\frac{\varepsilon}{2}} > M \cdot |\theta| + 1$;

(iii) $|y_{nh}|^{b-1+\frac{\varepsilon}{2}-\frac{\varepsilon}{3b}} \cdot h > 2 > |y_{nh}|^{-\frac{\varepsilon}{3b}} + 1$.

Then for all $|y| \leq |y_{nh}|^{1+\frac{\varepsilon}{2b}}$,

$$|u_{nh}| - |\theta \cdot f(y)| \geq |y_{nh}|^{b+\varepsilon} - |\theta| \cdot M \cdot |y_{nh}|^{b+\frac{\varepsilon}{2}} \geq |y_{nh}|^{b+\frac{\varepsilon}{2}}.$$

Therefore, for all $|y| \leq |y_{nh}|^{1+\frac{\varepsilon}{2b}}$,

$$[\theta \cdot f(y) + u_{nh}] \cdot [-\text{sgn}(y_{nh})] = |u_{nh}| - \text{sgn}(y_{nh}) \cdot \theta \cdot f(y) \geq |y_{nh}|^{b+\frac{\varepsilon}{2}}. \tag{100}$$

Now we prove that there exists some $t_0 \in [nh, (n+1)h)$, such that $|y_{t_0}| \geq |y_{nh}|^{1+\frac{\varepsilon}{4b}}$. We adopt the contradiction argument. Suppose that

$$|y_t| \leq |y_{nh}|^{1+\frac{\varepsilon}{4b}}, \quad \forall t \in [nh, (n+1)h), \tag{101}$$

then by (100),

$$\frac{dy_t}{dt} \cdot [-\text{sgn}(y_{nh})] \geq |y_{nh}|^{b+\frac{\varepsilon}{2}}, \quad \forall t \in [nh, (n+1)h).$$

Hence, $|y_{(n+1)h} - y_{nh}| \geq |y_{nh}|^{b+\frac{\varepsilon}{2}} \cdot h$, then

$$\begin{aligned} |y_{(n+1)h}| &\geq |y_{nh}|^{b+\frac{\varepsilon}{2}} \cdot h - |y_{nh}| \\ &= [|y_{nh}|^{b-1+\frac{\varepsilon}{2}-\frac{\varepsilon}{3b}} \cdot h - |y_{nh}|^{-\frac{\varepsilon}{3b}}] \cdot |y_{nh}|^{\frac{\varepsilon}{3b}+1} \\ &> |y_{nh}|^{1+\frac{\varepsilon}{3b}}, \end{aligned}$$

which contradicts (101) by the continuity of y_t.

So the supposition (101) is incorrect and there must exist some $t_0 \in [nh, (n+1)h)$, such that $|y_{t_0}| > |y_{nh}|^{1+\frac{\varepsilon}{4b}}$. Since the closed-loop system is autonomous in the time interval $t \in [nh, (n+1)h)$, y_t is monotonous on $t \in [nh, (n+1)h)$ by Lemma 4.1. So we have $|y_{(n+1)h}| \geq |y_{nh}|^{1+\frac{\varepsilon}{4b}}$.

Hence, if the system initial value y_0 satisfies Conditions (i)-(iii), then $|y_h| \geq |y_0|^{1+\frac{\varepsilon}{4b}}$ also satisfies Conditions (i)-(iii); then $|y_{2h}| \geq |y_h|^{1+\frac{\varepsilon}{4b}} \geq |y_0|^{(1+\frac{\varepsilon}{4b})^2}$ also satisfies Conditions (i)-(iii); $\cdots\cdots$.

Thus, we have

$$\lim_{n\to\infty} |y_{nh}| \geq \lim_{n\to\infty} |y_0|^{(1+\frac{\varepsilon}{4b})^n} = \infty.$$

Hence the hybrid control system is not globally stable. □

Actually, even for linear systems, such a phenomenon would happen.

Consider the simplest continuous-time linear control model:

$$\dot{y}_t = ay_t + u_t, \qquad t \geq 0, \tag{102}$$

where a is an unknown parameter.

The classical Lyapunov-based adaptive controller design (cf. [**15**]) is

$$u_t = -(c + \widehat{a}_t)y_t, \tag{103}$$

$$\dot{\widehat{a}}_t = \gamma y_t^2, \tag{104}$$

where, $c > 0$, $\widehat{a}_t$ is the estimate of a, γ is some constant appearing in the Lyaponov function candidate

$$V_t = \frac{1}{2}y_t^2 + \frac{1}{2\gamma}(\widehat{a}_t - a)^2. \tag{105}$$

We can easily prove that the closed-loop system (102)-(104) is globally asymptotically stable since

$$\begin{aligned} \dot{V}_t &= y_t(ay_t + u_t) + \frac{1}{\gamma}(\widehat{a}_t - a)\dot{\widehat{a}}_t \\ &= ay_t^2 - (c + \widehat{a}_t)y_t^2 + (\widehat{a}_t - a)y_t^2 \\ &= -cy_t^2. \end{aligned}$$

But if only sampled data of the output y_t are accessible and we still use the controller (103)-(104) in a approximating manner, i.e., (here,without loss of generality for linear systems, set the sample period $h = 1$)

$$u_t \triangleq -(c + \widehat{a}_n)y_n, \quad t \in [n, n+1), \quad n = 0, 1, 2, \cdots; \tag{106}$$

$$\widehat{a}_n \triangleq a_0 + \sum_{i=0}^{n} \gamma y_i^2, \quad a_0 \in R \text{ is chosen arbitrarily,} \tag{107}$$

then the situation changes completely as stated in the theorem below.

Theorem 4.2. *For the continuous-time linear system (102), if the sampled-data controller (106)-(107) is used, then the closed-loop system is not globally stable.*

Proof. Denote $b_n \stackrel{\triangle}{=} c + \widehat{a}_n$. Clearly, b_n is non-decreasing. If for some $n \in \{0, 1, 2, \cdots\}$,

$$b_n > 4|a|; \tag{108}$$

$$b_n > 6, \tag{109}$$

then we can prove $|y_{n+1}| \geq 2|y_n|$.

We adopt the contradiction argument. Suppose that

$$|y_{n+1}| < 2|y_n|. \tag{110}$$

Then by Lemma 4.1, we have $|y_t| < 2|y_n|, \quad \forall t \in [n, n+1)$. By this inequality and (102) (106) (108) (109), we have for $t \in [n, n+1)$,

$$\begin{aligned} |\dot{y}_t| &= |ay_t + u_t| = |ay_t - (c + \widehat{a}_n)y_n| = |ay_t - b_n y_n| \\ &\geq b_n|y_n| - |a|\,|y_t| \geq b_n|y_n| - 2|a|\,|y_n| \geq \frac{b_n}{2}|y_n| > 3|y_n|. \end{aligned} \tag{111}$$

Using Lemma 4.1 again, we have

$$|y_{n+1} - y_n| = \left| \int_n^{n+1} \dot{y}_t dt \right| = \int_n^{n+1} |\dot{y}_t| dt. \tag{112}$$

By (111) and (112), $|y_{n+1} - y_n| \geq 3|y_n|$. Hence, $|y_{n+1}| \geq 2|y_n|$. This contradicts to our supposition (110). So, if $b_n > \max\{4|a|, b\}$, then $|y_{n+1}| \geq 2|y_n|$.

Therefore, if the initial value $|y_0|$ is large enough such that (b_n is nondecreasing)

$$b_n \geq b_0 = c + \widehat{a}_0 = c + a_0 + \gamma y_0^2 > \max\{4|a|, 6\}, \quad n \geq 0,$$

then we will have that $|y_{n+1}| \geq 2|y_n|$ holds for $\forall n \geq 0$, and that

$$\limsup_{t \to \infty} |y_t| \geq \limsup_{n \to \infty} |y_n| \geq \limsup_{n \to \infty} 2^n |y_0| = \infty.$$

In fact, if only there exists some moment $n \geq 0$ such that

$$b_n = c + \widehat{a}_0 + \gamma \sum_{i=0}^{n} y_i^2 > \max\{4|a|, 6\},$$

then $\limsup_{t \to \infty} |y_t| = \infty$. □

Of course, it is not difficult to find a stabilizing sampled-data controller for the linear model (102), but for the nonlinear model (93) it is far more difficult.

5. Nonparametric Adaptive Control

In this chapter, we are going to show that if the nonlinearities have a certain linear growth rate, then optimal feedback control can be designed even for the following uncertain nonparametric model:

$$y_{t+1} = f(y_t) + u_t + \varepsilon_{t+1}, \tag{113}$$

where y_t, u_t and ε_t are the d-dimensional system output, input and white noises, and $f(\cdot)$ is an unknown nonlinear function.

Our objective is to design a feedback control u_t based on the observations $\{y_i, i \leq t\}$ at each step t, such that the system output $\{y_t\}$ tracks a known reference signal $\{y_t^*\}$ in an optimal way. If $f(\cdot)$ were known, it is obvious that such a controller would take the following form:

$$u_t = -f(y_t) + y_{t+1}^*.$$

Since in the present case, $f(\cdot)$ is unknown, we adopt the nonparametric estimation approach as used in [**6**], but without resorting to external excitations in the controller design.

Let $K(\cdot)$ be a nonnegative kernel function satisfying the following conditions:

$$K(0) > 0, \quad \int K(s)ds = 1, \quad \int K^2(s)ds < \infty, \quad \int \|s\| K(s)ds < \infty.$$

Here in our estimation process, let $K(\cdot)$ have a compact support, i.e.,

$$K(s) = 0, \quad \text{for } \|s\| > A.$$

Let $\delta_j(\cdot,\cdot)$ be a function shifted from $K(\cdot)$:

$$\delta_j(x,y) \triangleq K(j^a(x-y)), \quad \forall j > 0, \quad \delta_0 = 0, \tag{114}$$

where $a \in (0, \frac{1}{2d})$, d is the dimension of the system signals.

The nonparametric estimate of $f(y), y \in \mathbb{R}^d$ at time t is defined by

$$\widehat{f}_t(y) = \begin{cases} N_t^{-1}(y) \sum_{j=1}^{t} \delta_{j-1}(y_{j-1}, y)(y_j - u_{j-1}), & \text{if } N_t(y) > 0; \\ 0 & \text{otherwise,} \end{cases} \tag{115}$$

where,

$$N_t(y) \triangleq \sum_{j=1}^{t} \delta_{j-1}(y_{j-1}, y) \tag{116}$$

To define the adaptive feedback control, we need to introduce a sequence of truncation bounds denoted by $\{h_t\}$, which is positive, monotonically diverges to infinity, and satisfies

$$h_t = o(\sqrt{\log t}), \text{ as } t \to \infty. \tag{117}$$

Now, by the (truncated) certainty equivalence principle, the nonparametric adaptive control can be defined as

$$u_t = -\widehat{f}_t(y_t) I_{(|\widehat{f}_t(y_t)| \leq h_t)} + y_{t+1}^* \tag{118}$$

where $I_{(\cdot)}$ is the indicator function.

With this control, the closed-loop system equation is

$$y_{t+1} = f(y_t) - \widehat{f}_t(y_t) I_{(|\widehat{f}_t(y_t)| \le h_t)} + y^*_{t+1} + \varepsilon_{t+1}, \tag{119}$$

which is obviously a nonlinear dynamical system.

In order to analyze the properties of (119), we introduce the following assumptions on the system (113):

(A5.1) The nonlinear function $f(\cdot)$ is Lipschitz continuous, and there exist two constants $\alpha \in (0, 1)$ and $\beta \in (0, \infty)$ such that

$$\|f(x)\| \le \alpha \|x\| + \beta, \quad \forall x \in I\!R^d.$$

(A5.2) $\{\varepsilon_t\}$ is a Gaussian white noise sequence with mean zero and variance $\Sigma \ge \sigma^2 I_d > 0$.

(A5.3) The reference signal $\{y^*_t\}$ is bounded.

The main result of this chapter is stated as follows:

Theorem 5.1. *Consider the control system (113) where the nonlinear function $f(\cdot)$ is completely unknown. Let the assumptions (A5.1)-(A5.3) be fulfilled. Then the adaptive tracking control defined by (118) is asymptotically optimal in the sense that*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \|y_t - y^*_t - \varepsilon_t\|^2 = 0, \quad a.s..$$

We preface the proof of Theorem 5.1 with two lemmas.

Lemma 5.1 *Under the conditions of Theorem 5.1,*

$$\sup_{\|y\|^2 \le \frac{c}{4}\log t} \|\widetilde{f}_t(y)\| = o(t^{-\delta}), \quad a.s., \quad as\ t \to \infty,$$

where $\widetilde{f} = \widehat{f} - f$, $\quad c \in (0, \sigma^2(\frac{1}{2} - ad)\,)$ *and* $\delta \in (0, \min\{(\frac{1}{2} - ad - c/\sigma^2),\ (1 - ad - c/\sigma^2)a\})$.

Proof. We can divide $\widetilde{f}_t(y)$ into two parts:

$$\widetilde{f}_t(y) = \widehat{f}_t(y) - f(y) = \frac{M_t(y)}{N_t(y)} + \frac{L_t(y)}{N_t(y)}, \tag{120}$$

where, $M_t(y) \triangleq \sum_{j=1}^{t} \delta_{j-1}(y_{j-1}, y) \cdot \varepsilon_j$, and $L_t(y) \triangleq \sum_{j=1}^{t} \delta_{j-1}(y_{j-1}, y)[f(y_{j-1}) - f(y)]$.

By the closed-loop system equation (119), condition (117) and Assumptions (A5.1) and (A5.3), we have

$$\|y_{t+1}\| \le \alpha\|y_t\| + o(\sqrt{\log t}) + O(1) + \|\varepsilon_{t+1}\|.$$

Then it follows from $\alpha \in (0, 1)$ that

$$\frac{1}{t} \sum_{j=1}^{t} \|y_j\|^2 = o(\log t). \tag{121}$$

Define

$$z_t \triangleq f(y_t) + u_t, \tag{122}$$

$$\mathcal{F}_{t-1} \triangleq \sigma\{(\varepsilon_j)_{j \le t-1}, (y^*_j)_{j \le t}\},$$

then $y_t = z_{t-1} + \varepsilon_t$ and by Assumption (A5.2),

$$\begin{aligned}
&E[K(j^a(y_j - y))|\mathcal{F}_{j-1}] \\
&\geq \text{const.} \int_{\mathbb{R}^d} \exp(-\frac{\|x\|^2}{2\sigma^2}) K(j^a(z_{j-1} + x - y))dx \\
&= \text{const.} j^{-ad} \int_{\mathbb{R}^d} \exp(-\frac{1}{2\sigma^2}\|\lambda j^{-a} + y - z_{j-1}\|^2) K(\lambda) d\lambda \\
&= \text{const.} j^{-ad} \exp(-\frac{1}{2\sigma^2}\|\lambda_0 j^{-a} + y - z_{j-1}\|^2) \int_{\mathbb{R}^d} K(\lambda) d\lambda,
\end{aligned}$$

where, the last equality follows from the integral mean value theorem, and $|\lambda_0| \leq A$ since $K(\lambda) = 0$, $|\lambda| > A$.

Choose $0 < c < \sigma^2(\frac{1}{2} - ad)$. Since $\|y + A - s\|^2 \leq 4\|y\|^2 + 4\|s\|^2 + 2A^2$, we have for any $t \geq 0$,

$$\begin{aligned}
&\inf\left\{\exp(-\frac{1}{2\sigma^2}\|y + A - s\|^2); \|s\|^2 \leq \frac{c}{4}\log t \text{ and } \|y\|^2 \leq \frac{c}{4}\log t\right\} \\
&\geq \text{const.} t^{-c/\sigma^2}.
\end{aligned}$$

Then for $\|z_{j-1}\|^2 \leq \frac{c}{4}\log t$ and $\|y\|^2 \leq \frac{c}{4}\log t$,

$$\begin{aligned}
E[K(j^a(y_j - y))|\mathcal{F}_{j-1}] &\geq \text{const.} j^{-ad} t^{-c/\sigma^2} \\
&\geq \text{const.} t^{-ad-c/\sigma^2}.
\end{aligned}$$

Hence, for $\|y\|^2 \leq \frac{c}{4}\log t$,

$$\begin{aligned}
&\frac{1}{t}\sum_{j=1}^{t} E[K(j^a(y_j - y))|\mathcal{F}_{j-1}] \\
\geq\ & \text{const.} t^{-ad-c/\sigma^2} \times \frac{1}{t}\sum_{j=1}^{t} I_{(\|z_{j-1}\|^2 \leq \frac{c}{4}\log t)} \\
\geq\ & \text{const.} t^{-ad-c/\sigma^2} \times \frac{1}{t}\sum_{j=2}^{t} I_{(\|z_{j-1}\|^2 \leq \frac{c}{4}\log j)}
\end{aligned} \tag{123}$$

Now, we prove that

$$d_t \triangleq \frac{1}{t}\sum_{j=2}^{t} I_{(\|z_{j-1}\|^2 \leq \frac{c}{4}\log j)} \to 1, \quad \text{a.s..}$$

Actually,

$$\begin{aligned}
1 &= \frac{1}{t-1}\sum_{j=2}^{t}\left[I_{(\|z_{j-1}\|^2 \leq \frac{c}{4}\log j)} + I_{(\|z_{j-1}\|^2 > \frac{c}{4}\log j)}\right] \\
&\leq \frac{t}{t-1} d_t + \frac{1}{t-1}\sum_{j=2}^{t}\frac{\|z_{j-1}\|^2}{\frac{c}{4}\log j}.
\end{aligned} \tag{124}$$

Let $S_j \triangleq \sum_{i=2}^{j} \|z_{i-1}\|^2$, $j \geq 2$ and $S_1 \triangleq 0$, then $\|z_{j-1}\|^2 = S_j - S_{j-1}$. By (121) and (122), we have

$$\frac{1}{t}\sum_{j=2}^{t} \|z_{j-1}\|^2 = o(\log t), \qquad \text{i.e.,} \qquad S_t = o(t\log t).$$

Thus,

$$\begin{aligned} &\frac{1}{t}\sum_{j=2}^{t} \frac{\|z_{j-1}\|^2}{\log j} = \frac{1}{t}\sum_{j=2}^{t} \frac{S_j - S_{j-1}}{\log j} \\ &= \frac{1}{t}\left[\sum_{j=2}^{t-1}\left(\frac{S_j}{\log j} - \frac{S_j}{\log(j+1)}\right) - \frac{S_1}{\log 2} + \frac{S_t}{\log t}\right] \\ &\leq \frac{1}{t}\left[\sum_{j=2}^{t-1} \frac{S_j}{j\log^2 j} - \frac{S_1}{\log 2} + \frac{S_t}{\log t}\right] \\ &\to 0. \end{aligned} \tag{125}$$

On the other hand, apparently $d_t \leq 1$. Hence, by (124) and (125), $d_t \to 1$, a.s.. Consequently, by (123)

$$\begin{aligned} &\liminf_{t\to\infty} t^{ad+c/\sigma^2-1} \inf\left\{\sum_{j=1}^{t} E[K(j^a(y_j - y))|\mathcal{F}_{j-1}];\, \|y\|^2 \leq \frac{c}{4}\log t\right\} \\ &\geq \text{const.} \end{aligned} \tag{126}$$

By the uniform law of large numbers (Theorem 6.4.34 in [**6**]), we have

$$\begin{aligned} &\sup\left\{\left|N_t(y) - \sum_{j=1}^{t-1} E[K(j^a(y_j - y))|\mathcal{F}_{j-1}]\right|;\, \|y\|^2 \leq \frac{c}{4}\log t\right\} \\ &= o(t^\beta), \quad \text{a.s. for all } \beta > \frac{1}{2}. \end{aligned}$$

Then for $1 - c/\sigma^2 - ad > \frac{1}{2}$, (i.e. $0 < a < \frac{1}{2d}$ and $c < \sigma^2(\frac{1}{2} - ad)$)

$$\liminf_{t\to\infty} t^{ad+c/\sigma^2-1}\inf\{N_t(y);\, \|y\|^2 \leq \frac{c}{4}\log t\} \geq \text{const..} \tag{127}$$

Thus there exists some $t_1 > 0$, when $t > t_1$ and $\|y\|^2 \le \frac{c}{4}\log t$, $N_t(y) > 0$. Then noting that $f(\cdot)$ is Lipschitz continuous and $K(\cdot)$ has a compact support, we have

$$\begin{aligned}
\|L_t(y)\| &\le \text{const.} \sum_{j=1}^{t-1} K(j^a(y_j - y))\|y_j - y\| \\
&\le \text{const.} \sum_{j=1}^{t-1} j^{-a} K(j^a(y_j - y)) \\
&\le \text{const.} + \text{const.} \sum_{j=t_1+1}^{t-1} K(j^a(y_j - y))(N_j(y))^{-a} \\
&= \text{const.} + \text{const.} \sum_{j=t_1+1}^{t-1} [N_{j+1}(y) - N_j(y)] \cdot (N_j(y))^{-a} \\
&\le \text{const.} + \text{const.}(N_t(y))^{1-a}/(1-a)
\end{aligned}$$

Again, applying the uniform law of large numbers (Theorem 6.4.34 in [**6**]), we have

$$\sup_{\|y\|^2 \le \frac{c}{4}\log t} |M_t(y)| = o(t^\beta), \quad \text{a.s. for all } \beta > \frac{1}{2}.$$

Hence, by (120) and (127),

$$\sup\{\|\widetilde{f}_t(y)\|; \|y\|^2 \le \frac{c}{4}\log t\} = o(t^{-\delta}),$$

for all $\delta < \min\{(\frac{1}{2} - ad - c/\sigma^2),\ (1 - ad - c/\sigma^2)a\}$. □

Lemma 5.2. *Under the conditions of Theorem 5.1, for any $m \ge 1$, we have*

$$\sum_{j=1}^{t} \|y_{j+1}\|^m = O(t), \quad \text{a.s.,} \quad \text{as } t \to \infty.$$

Proof. By the closed-loop system equation (119), we have

$$\begin{aligned}
y_{t+1} &= f(y_t) - \widehat{f}_t(y_t) I_{\{\|\widehat{f}_t(y_t)\| \le h_t\}} + y^*_{t+1} + \varepsilon_{t+1} \\
&= [f(y_t) - \widehat{f}_t(y_t)] I_{\{\|\widehat{f}_t(y_t)\| \le h_t\}} + f(y_t) I_{\{\|\widehat{f}_t(y_t)\| > h_t\}} + y^*_{t+1} + \varepsilon_{t+1}
\end{aligned}$$

For any integer $m \ge 1$,

$$\begin{aligned}
\|y_{t+1}\|^m \le\ & \lambda_1 \|[f(y_t) - \widehat{f}_t(y_t)] I_{\{\|\widehat{f}_t(y_t)\| \le h_t\}} + f(y_t) I_{\{\|\widehat{f}_t(y_t)\| > h_t\}}\|^m \\
& + \lambda_2 \|y^*_{t+1} + \varepsilon_{t+1}\|^m,
\end{aligned}$$

where, $\lambda_1 > 1$ is suitably chosen to make that $\lambda_1 \alpha^m = \alpha_1 < 1$.

Thus, noting that $\{\|\widehat{f}_t(y_t)\| \le h_t\}$ and $\{\|\widehat{f}_t(y_t)\| > h_t\}$ do not intersect, we have

$$\begin{aligned}
\|y_{t+1}\|^m \le\ & \lambda_1 \|f(y_t) - \widehat{f}_t(y_t)\|^m I_{\{\|\widehat{f}_t(y_t)\| \le h_t\}} \\
& + \lambda_1 \|f(y_t)\|^m I_{\{\|\widehat{f}_t(y_t)\| > h_t\}} \\
& + \lambda_2 \|y^*_{t+1} + \varepsilon_{t+1}\|^m \\
=\ & \lambda_1 \|f(y_t) - \widehat{f}_t(y_t)\|^m I_{\{\|\widehat{f}_t(y_t)\| \le h_t,\ \|y_t\|^2 \le \frac{c}{4}\log t\}} \\
& + \lambda_1 \|f(y_t) - \widehat{f}_t(y_t)\|^m I_{\{\|\widehat{f}_t(y_t)\| \le h_t,\ \|y_t\|^2 > \frac{c}{4}\log t\}} \\
& + \lambda_1 \|f(y_t)\|^m I_{\{\|\widehat{f}_t(y_t)\| > h_t\}} + \lambda_2 \|y^*_{t+1} + \varepsilon_{t+1}\|^m
\end{aligned}$$

Hence, by Lemma 5.1,

$$\begin{aligned}
&\sum_{j=1}^{t}\|y_{j+1}\|^m\\
\leq\ & o(t)+\lambda_1\sum_{j=1}^{t}[\lambda_3\|f(y_j)\|^m+\lambda_4\|\widehat{f}_j(y_j)\|^m]I_{\{\|\widehat{f}_j(y_j)\|\leq h_j,\,\|y_j\|^2>\frac{c}{4}\log j\}}\\
&+\lambda_1\sum_{j=1}^{t}\|f(y_j)\|^m I_{\{\|\widehat{f}_j(y_j)\|>h_j\}}+O(t),
\end{aligned}$$

where, we can choose some $\lambda_3>1$ to make $\lambda_1\lambda_3\cdot\alpha^m=\lambda_3\cdot\alpha_1\stackrel{\triangle}{=}\alpha_3<1$.

Therefore,

$$\begin{aligned}
&\sum_{j=1}^{t}\|y_{j+1}\|^m\\
\leq\ & \alpha_3\sum_{j=1}^{t}\|y_j\|^m+\lambda_1\lambda_4\sum_{j=1}^{t}\|\widehat{f}_j(y_j)\|^m I_{\{\|\widehat{f}_j(y_j)\|\leq h_j,\,\|y_j\|^2>\frac{c}{4}\log j\}}+O(t)\\
\leq\ & \alpha_3\sum_{j=1}^{t}\|y_j\|^m+\lambda_1\lambda_4\sum_{j=1}^{t}h_j^m\frac{\|y_j\|^m}{(\frac{c}{4}\log j)^{\frac{m}{2}}}+O(t).
\end{aligned}$$

Thus, using (117) again, we have

$$\sum_{j=1}^{t}\|y_{j+1}\|^m=O(t).$$

Proof of Theorem 5.1. By the closed-loop system equation (119), we have

$$\begin{aligned}
&\frac{1}{t}\sum_{j=1}^{t}\|y_{j+1}-y_{j+1}^*-\varepsilon_{j+1}\|^2\\
=\ & \frac{1}{t}\sum_{j=1}^{t}\|f(y_j)-\widehat{f}_j(y_j)\|^2 I_{\{\|\widehat{f}_j(y_j)\|\leq h_j\}}+\frac{1}{t}\sum_{j=1}^{t}\|f(y_j)\|^2 I_{\{\|\widehat{f}_j(y_j)\|>h_j\}}\\
=\ & \frac{1}{t}\sum_{j=1}^{t}\|f(y_j)-\widehat{f}_j(y_j)\|^2 I_{\{\|\widehat{f}_j(y_j)\|\leq h_j,\,\|y_j\|^2\leq\frac{c}{4}\log j\}}\\
&+\frac{1}{t}\sum_{j=1}^{t}\|f(y_j)-\widehat{f}_j(y_j)\|^2 I_{\{\|\widehat{f}_j(y_j)\|\leq h_j,\,\|y_j\|^2>\frac{c}{4}\log j\}}\\
&+\frac{1}{t}\sum_{j=1}^{t}\|f(y_j)\|^2 I_{\{\|\widehat{f}_j(y_j)\|>h_j,\,\|y_j\|^2\leq\frac{c}{4}\log j\}}\\
&+\frac{1}{t}\sum_{j=1}^{t}\|f(y_j)\|^2 I_{\{\|\widehat{f}_j(y_j)\|>h_j,\,\|y_j\|^2>\frac{c}{4}\log j\}}.
\end{aligned}$$

Here, using Lemma 5.1, we have

$$\frac{1}{t}\sum_{j=1}^{t}\|f(y_j)-\widehat{f}_j(y_j)\|^2 I_{\{\|y_j\|^2\leq\frac{c}{4}\log j\}}=o(1),$$

and,

$$\|\widehat{f}_j(y_j)\|I_{\{\|y_j\|^2\leq\frac{c}{4}\log j\}}\leq\|f(y_j)\|+o(1).$$

Then

$$
\begin{aligned}
&\frac{1}{t}\sum_{j=1}^{t}\|y_{j+1}-y_{j+1}^{*}-\varepsilon_{j+1}\|^{2}\\
\le\ & o(1)\\
&+\frac{2}{t}\sum_{j=1}^{t}\|f(y_j)\|^2 I_{\{\|y_j\|^2>\frac{c}{4}\log j\}}+\frac{2}{t}\sum_{j=1}^{t}\|\widehat{f}_j(y_j)\|^2 I_{\{\|\widehat{f}_j(y_j)\|\le h_j,\,\|y_j\|^2>\frac{c}{4}\log j\}}\\
&+\frac{1}{t}\sum_{j=1}^{t}\|f(y_j)\|^2\frac{\|f(y_j)\|+o(1)}{h_j}\\
&+\frac{1}{t}\sum_{j=1}^{t}\|f(y_j)\|^2\frac{\|y_j\|^2}{\frac{c}{4}\log j}.
\end{aligned}
$$

Now, by Lemma 5.2 and Assumption (A1), we have

$$\frac{1}{t}\sum_{j=1}^{t}\frac{\|f(y_j)\|^3}{h_j}=o(1),$$

and

$$\frac{1}{t}\sum_{j=1}^{t}\|f(y_j)\|^2\frac{\|y_j\|^2}{\frac{c}{4}\log j}=o(1).$$

Also, by (117),

$$
\begin{aligned}
&\frac{1}{t}\sum_{j=1}^{t}\|\widehat{f}_j(y_j)\|^2 I_{\{\|\widehat{f}_j(y_j)\|\le h_j,\,\|y_j\|^2>\frac{c}{4}\log j\}}\\
=\ &\frac{1}{t}\sum_{j=1}^{t}h_j^2\cdot\frac{\|y_j\|^2}{\frac{c}{4}\log j}=o(1)
\end{aligned}
$$

Therefore,

$$\frac{1}{t}\sum_{j=1}^{t}\|y_{j+1}-y_{j+1}^{*}-\varepsilon_{j+1}\|^{2}=o(1),$$

which is just the conclusion of the theorem. □

References

[1] K. J. Åström and B. Wittenmark, *Adaptive Control,* second edition. New York: Addison-Wesley, 1995.

[2] B. Bercu, "Weighted estimation and tracking for ARMAX models", *SIAM J. on Contr. and Optim.*, Vol.33, No.1, pp.89-106, 1995.

[3] B. Bercu and M. Duflo, "Moindres Carrés Pondérés et Poursuite", *Ann. Inst. Henri Poincaré,* Vol. 28, pp. 403-430, 1992.

[4] H. F. Chen and L. Guo, *Identification and Stochastic Adaptive Control.* Boston, MA: Birkhäuser, 1991.

[5] H. Cho and S. I. Marcus, "On adaptive control of stochastic bilinear system", *IEEE Trans. Automat. Contr.*, Vol.32, No.12, pp.1103-1106, 1987.

[6] M. Duflo, *Random Iterative Models.* Berlin: Springer-Verlag, 1997.

[7] L. Guo, "A note on continuous-time ELS", *Systems & Control Letters*, Vol.22, No.2, pp.111-121, 1994.

[8] L. Guo, "Convergence and logarithm laws of self-tuning regulators", *Automatica*, Vol.31, No.3, pp. 435-450, 1995.

[9] L. Guo, "Self-convergence of weighted least-squares with applications to stochastic adaptive control", *IEEE Trans. Automat. Contr.*, Vol.41, No.1, pp.79-89, 1996.

[10] L. Guo, "On critical stability of discrete-time adaptive nonlinear control", *IEEE Trans. Automat. Contr.*, Vol.42, No.11, pp.1488-1499, 1997.

[11] L. Guo and H. F. Chen, "The Åström-Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers", *IEEE Trans. Automat. Contr.*, Vol.36, No.7, pp.802-812, 1991.

[12] I. A. Ibragimov and R. Z. Has'minskii, *Statistical Estimation: Asympototic Theory*, New York: Springer-Verlag, 1981.

[13] P. A. Ioannou and J. Sun, *Robust Adaptive Control.* Engle-wood Cliffs, NJ: Prentice-Hall, 1996.

[14] I. Kanellakopoulos, "A discrete-time adaptive nonlinear system", *IEEE Trans. Automat. Contr.*, Vol.39, No.11, pp.2362-23, 1994.

[15] M. Krstić, I. Kanellakopoulos and P. V. Kokotović, *Nonlinear and Adaptive Control Design.* New York: John Wiley & Sons, 1995.

[16] H. L. Van Trees, *Detection, Estimation, and Modulation Theory: PART I*, New York: John Wiley & Sons, 1968.

[17] C. Wei and L. Guo, "Adaptive control of a class of nonlinear stochastic systems", in *Proc. 1996 Chinese Control Conference,* pp.508-512, Qingdao, September 15-20, 1996.

[18] L.-L. Xie and L. Guo, "Adaptive control of a class of discrete-time affine nonlinear systems", *Systems & Control Letters*, Vol.35, pp.201-206, 1998.

[19] L.-L. Xie and L. Guo, "Fundamental limitations of discrete-time adaptive nonlinear control", *IEEE Trans. Autom. Contr.*, November, 1999, to appear.

[20] K. Zhou, J. C. Doyle and K. Glover, *Robust and Optimal Control,* Prentice-Hall, 1996.

Institute of Systems Science Chinese Academy of SciencesBeijing 100080P. R. China