SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

January 2013, Vol. 56 012201:1–012201:9
doi: 10.1007/s11432-012-4622-7

A dynamical inequality for the output of uncertain nonlinear systems

LI ChanYing[1] & GUO Lei[2]*

[1]*National Center for Mathematics and Interdisciplinary Sciences, Chinese Academy of Sciences, Beijing* 100090*, China;*
[2]*Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing* 100090*, China*

Received June 19, 2011; accepted November 24, 2011; published online August 2, 2012

Abstract An important task in control theory is to study the limitations of feedback principle in dealing with uncertainties. Although some progresses have been achieved in this area, they are all focused on some special classes of linearly parameterized nonlinear uncertain systems. In this paper, we will present a dynamic inequality for the output process of a quite general class of nonlinear dynamical control systems with nonlinearly parameterized uncertain parameters. This inequality will be established using a stochastic imbedding approach based on a Cramér-Rao inequality for dynamical systems, and will be shown to play a crucial role in investigating the fundamental limitations of the feedback mechanism.

Keywords dynamical systems, uncertain parameters, feedback control, Cramer-Rao bound

Citation Li C Y, Guo L. A dynamical inequality for the output of uncertain nonlinear systems. Sci China Inf Sci, 2013, 56: 012201(9), doi: 10.1007/s11432-012-4622-7

1 Introduction

Feedback, as a central concept in control theory, can be used to deal with various uncertainties including initial state uncertainty, parameter uncertainty, structure uncertainty, and unknown disturbances, etc. As has been well demonstrated in a vast literature in many areas including adaptive control and robust control [1–7]. A natural and important question from both theoretical perspective and practical applications arises: Does the feedback principle have any limitations in dealing with parameter or structure uncertainties?

To answer this fundamental question, we first observe that most of the existing literature on control of uncertain dynamical systems assumes that a continuous-time controller (or a controller with sampling rate fast enough) is implemented, and/or assumes that the uncertain systems are linear (or the nonlinear dynamics are dominated by a linear growth rate). These assumptions are obviously not true in many applications, though they are helpful for the theoretical investigation in the existing literature.

To understand what will happen when the above-mentioned assumptions are violated, a series of studies initiated by [8] on the maximum capability and fundamental limitations of the feedback principle have been carried out in recent years [9–15]. It has been found and rigorously established that the

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

feedback mechanism (or principle) does indeed have fundamental limitations in dealing with various uncertain nonlinear dynamical systems. For example, such limitations exist in the control of discrete-time parametric nonlinear stochastic systems having a nonlinear growth rate [13], in the control of discrete-time nonparametric nonlinear systems even if the nonlinearity has a linear growth rate [14], and also in the sampled-data control of continuous-time nonlinear systems with a prescribed sampling rate, even if the nonlinearity is bounded by linear growth rate [15].

All the above mentioned progresses, however, are focused on some special classes of nonlinear uncertain dynamical systems. To lay a foundation for the investigation of the fundamental limitations of the feedback mechanism for more general uncertain systems, we will in this paper establish a general dynamic inequality for a quite general class of uncertain nonlinear systems. Two key ideas and methods behind this study are as follows: i) The stochastic embedding approach. The uncertain systems are assumed to have parametric uncertainties, which means that the unknown parameters may take any values in a certain domain, which in turn means that the feedback law should be able to deal with all the systems corresponding to all the possible parameter values in this domain. Hence, if, for any given feedback law, there is always at least one system that cannot be dealt with by this law, then a fundamental limitation of the feedback principle will be found. Thus, we can imbed a random variable into the domain of the uncertain parameters, as well as a stochastic sequence into the domain of the disturbances. By doing so, a key observation is that without requiring our analysis to hold on all the sample paths, we carry out the analysis on a set with non-zero probability [9]. ii) After introducing the stochastic imbedding approach, the next key step is to establish an extended Cramér-Rao lower bound [13,16] to the uncertainties in the dynamical systems to be controlled. Such a universal lower bound to the estimation error of the system uncertainties will then be transformed to the lower bound to the output processes of the dynamical systems to be controlled. These ideas will be expounded on in detail in the paper.

The remainder of this paper is organized as follows. In Section 2, we will present the main theorem of the paper. The proof is given in Section 3. The final section will give some concluding remarks.

2 Main result

Consider the following model:

$$y_{t+1} = f(\boldsymbol{\theta}, \boldsymbol{\phi}_t) + w_{t+1}, \tag{1}$$

where $\boldsymbol{\theta} \in \mathbb{R}^p, p \geqslant 1$ is an unknown parameter vector, $\boldsymbol{\phi}_t = (y_t, u_t; \ldots; y_{t-d}, u_{t-d})$ with $d \geqslant 0$, u_t and w_t are the system regression vector, feedback law and noise signal respectively, and where $f(\cdot,\cdot) : \mathbb{R}^{2(d+1)+p} \to \mathbb{R}$ is a nonlinear measurable function. Denote the partial derivative of $f(\boldsymbol{\theta}, \boldsymbol{\phi})$ with respect to θ by $f'(\boldsymbol{\theta}, \boldsymbol{\phi}) = [f_1'(\boldsymbol{\theta}, \boldsymbol{\phi}), f_2'(\boldsymbol{\theta}, \boldsymbol{\phi}), \ldots, f_p'(\boldsymbol{\theta}, \boldsymbol{\phi})]^\tau$, which is assumed to exist and be continuous.

We assume that the parameter vector and the noise sequence satisfy the following conditions:

A1) The unknown parameter vector $\boldsymbol{\theta}$ belongs to a certain ball in $\mathbb{R}^p$: $\Theta_0 = \{\boldsymbol{\theta} : \|\boldsymbol{\theta}\| \leqslant R\}$, where $R > 0$ may not be known *a priori*.

A2) The noise sequence is an arbitrarily bounded sequence with an unknown upper bound $w > 0$, i.e.,

$$\sup_{t \geqslant 1} |w_t| \leqslant w. \tag{2}$$

We are interested in how the unknown parameter $\boldsymbol{\theta} \in \Theta_0$ may fundamentally influence the output process of (1) . For this, we need the following regularity assumption on the nonlinear function $f(\cdot,\cdot)$.

A3) The input sequence $\{u_t\}$ is any feedback sequence, i.e., u_t is a Borel measurable function of the observations $\{y_0, y_1, \ldots, y_t\}$.

A4) For any $\varepsilon > 0$, there exists a non-increasing and nonnegative function $h(\varepsilon)$ such that for any $\boldsymbol{\phi} \in \mathbb{R}^{2d+2}$ with $\|\boldsymbol{\phi}\| > h(\varepsilon)$, the set $\Delta_{\varepsilon,\boldsymbol{\phi}}$ defined by

$$\Delta_{\varepsilon,\boldsymbol{\phi}} \triangleq \{\theta \in \Theta_0 : |f(\boldsymbol{\theta}, \boldsymbol{\phi})| < \varepsilon \max_{\boldsymbol{\theta}} \|f'(\boldsymbol{\theta}, \boldsymbol{\phi})\|\}, \tag{3}$$

satisfies $L(\Delta_{\varepsilon,\phi}) \leqslant M\varepsilon$, where $L(\cdot)$ denotes the Lebesgue measure on $\mathbb{R}^p$ and $M > 0$ is some constant. Moreover, assume that for $\|\boldsymbol{\phi}\| \geqslant h(\delta)$, $f_i'(\boldsymbol{\theta}, \boldsymbol{\phi}) \neq 0$ for $\boldsymbol{\theta} \in \Theta_0, i = 1, 2, \ldots, p$, where δ is any real number in $\left(0, \frac{\pi R^2}{16M}\right]$.

Theorem 2.1. Let Assumptions A3), A4) be satisfied. Then, there exist a parameter $\boldsymbol{\theta} \in \Theta_0$ and a noise sequence $\{w_i\}$ satisfying A2) such that the corresponding outputs of system (1) satisfy

$$y_{t+1}^2 \geqslant \frac{1}{C_1(t+1)^4}\left(\frac{d_t}{r_{t-1}} - 1\right) - C_2, \tag{4}$$

provided that the regressors satisfy $\|\boldsymbol{\phi}_i\| > h((i+1)^{-2}\delta)$ for all $i \leqslant t$, where $d_t \triangleq \sum_{j=1}^p \min_{\theta \in \Theta_0} |f_j'(\boldsymbol{\theta}, \boldsymbol{\phi}_t)|^2$, and

$$r_{t-1} \triangleq 1 + t^4 \max_{\boldsymbol{\theta} \in \Theta_0}\left[\sum_{i=0}^{t-1} \|f'(\boldsymbol{\theta}, \boldsymbol{\phi}_i)\|^2\right] \quad \text{with} \quad r_{-1} \triangleq 1. \tag{5}$$

and $C_1, C_2 > 0$ are some constants.

If the uncertain parameter vector $\boldsymbol{\theta}$ enters function $f(\boldsymbol{\theta}, \boldsymbol{\phi})$ linearly, the first condition of Assumption A4) will be satisfied automatically ([10, Lemma 3.2]). To see this clearly, let us give a simple example where $f(\boldsymbol{\theta}, \boldsymbol{\phi}) = \boldsymbol{\theta}\phi$ with both $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ being scalars. Let $h(\cdot) \equiv 0$ in Assumption A4). Then for any $\phi \in \mathbb{R}$,

$$\Delta_{\varepsilon,\boldsymbol{\phi}} = \{\boldsymbol{\theta} \in \Theta_0 : |\boldsymbol{\theta}| < \varepsilon\} \subset \{|\boldsymbol{\theta}| \leqslant \min\{\varepsilon, R\}\},$$

which immediately implies $L(\Delta_{\varepsilon,\phi}) \leqslant M\varepsilon$ for some $M > 0$.

Corollary 2.1. Let Assumption A3) be satisfied and that $f(\boldsymbol{\theta}, \boldsymbol{\phi}) = \boldsymbol{\theta}^\tau g(\boldsymbol{\phi})$ for some known function $g(\cdot) \in \mathbb{R}^p$ with $g(\boldsymbol{\phi}) \neq 0$ if $\|\boldsymbol{\phi}\| > h(\varepsilon)$. Then, there exist a parameter $\boldsymbol{\theta} \in \Theta_0$ and a noise sequence $\{w_i\}$ satisfying A2) such that the corresponding outputs of system (1) satisfy

$$y_{t+1}^2 \quad \geqslant \quad \frac{1}{C_1(t+1)^4}\left(\frac{\|g(\boldsymbol{\phi}_t)\|^2}{1 + t^4 \sum_{i=0}^{t-1} \|g(\boldsymbol{\phi}_i)\|^2} - 1\right) - C_2, \tag{6}$$

provided that the regressors satisfy $\|\boldsymbol{\phi}_i\| > h((i+1)^{-2}\delta)$ for all $i \leqslant t$, where $C_1, C_2 > 0$ are some constants.

Remark 2.1. Theorem 2.1 and Corollary 2.1 give a dynamical lower bound to the system outputs. The key points lie in the fact that such inequalities no longer involve the uncertain parameters and are true for any feedback sequence $\{u_t\}$. This may enable one to analyze the possible divergence rate for any given feedback sequence $\{u_t\}$.

Remark 2.2. We remark that, under some mild conditions, the output at $t+1$ with (4) or (6) will further result in $\|\boldsymbol{\phi}_{t+1}\| > h((t+2)^{-2}\delta)$, and the dynamical lower bound to the outputs can be derived by repeatedly applying these inequalities [10].

Remark 2.3. An immediate application of Theorem 2.1 and Corollary 2.1 is to establish the limitations of the feedback capability. They can be applied directly to a large number of parameterized models, which give a much simpler analysis than the existing method [9,11,12]. For example, consider the following system:

$$y_{t+1} = \theta^\tau y_t^b + u_t + w_{t+1}, \tag{7}$$

where y_t, u_t and w_t are the system output, input and noise respectively, θ is an unknown parameter, and the exponent $b > 0$ is a real number. By [11], we know that system (7) is not globally stabilizable by feedback if $b \geqslant 4$. Such an impossibility theorem was derived in [11] by a rather involved method, which estimate the uncertain domain of θ at each step. Now, due to Corollary 2.1, the derivation becomes much easier. In fact, for any feedback control law, by time $t \geqslant 0$, we have

$$y_{t+1}^2 \geqslant \frac{1}{C_1(t+1)^4}\left(\frac{y_t^{2b}}{1 + t^4 \sum_{i=0}^{t-1} y_i^{2b}} - 1\right) - C_2,$$

if $y_i \neq 0, i \leqslant t$. When $b \geqslant 4$, it can be proved by the above inequality that the outputs tend to infinity if the initial value $|y_0|$ is large enough [10,11]. For a more general case, where the model is a nonlinearly parameterized uncertain system

$$y_{t+1} = f(\theta, y_t) + u_t + w_{t+1}, \tag{8}$$

the extension of the analysis method as used in [11] can hardly be applied due to the complexity of the problem. But, the current Theorem 2.1 turns out to be very useful in establishing the impossibility theorem on feedback (see [10] for details).

3 Proof of Theorem 2.1

Theorem 2.1 is a deterministic result in nature, but it will be established by using a stochastic imbedding approach as that used in the linearly parameterized case [9]. The reason for using this approach is that a direct analysis in the deterministic way is much more complicated even for linearly parameterized models [12], not to mention the current case where the unknown parameters enter into the systems nonlinearly.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\boldsymbol{\theta} \in \mathbb{R}^p$ be a random vector and $\{w_t\}_{t=1}^{\infty}$ be a stochastic process on this probability space. The stochastic imbedding idea is to construct a special class of $\boldsymbol{\theta}$ and $\{w_t\}_{t=1}^{\infty}$ on this probability space, such that their sample pathes are consistent with those in our Assumptions A1) and A2), and they are easily applicable to some Cramér-Rao-like inequalities for dynamical systems. This can be done by choosing a suitable class of probability density functions (p.d.f.s) as follows.

Let $\boldsymbol{\theta} \in \Theta_0$ have the following spherical p.d.f. $p(\boldsymbol{\theta})$ defined by [13, Remark 3.2.3]

$$p(\boldsymbol{\theta}) = \begin{cases} c(2^{-1}R^2 - \|\boldsymbol{\theta}\|^2), & \text{if } 0 \leqslant \|\boldsymbol{\theta}\| \leqslant R/2, \\ c(R - \|\boldsymbol{\theta}\|)^2, & \text{if } R/2 \leqslant \|\boldsymbol{\theta}\| \leqslant R, \end{cases} \tag{9}$$

where c is some constant chosen to make $\int_{\|\boldsymbol{\theta}\| \leqslant R} p(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} = 1$ and $\|\cdot\|$ denotes the Euclidean norm.

Note that by a direct calculation based on the above definition of $p(\boldsymbol{\theta})$, it can be concluded that the derivative of $p(\boldsymbol{\theta})$ exists in $\|\boldsymbol{\theta}\| < R$ but not on the boundary $\|\boldsymbol{\theta}\| = R$. Similarly, the second derivative of $p(\boldsymbol{\theta})$ exists for all $\|\boldsymbol{\theta}\| < R$ except $\|\boldsymbol{\theta}\| = \frac{R}{2}$. Fortunately, these will not affect our results and subsequent analyses, since the Lebesgue measure of $\{\boldsymbol{\theta} : \|\boldsymbol{\theta}\| = R \text{ or } R/2\}$ is zero.

Next, let $\{w_t\}$ be an independent sequence independent of $\boldsymbol{\theta}$ with w_t having a Gaussian p.d.f. $q_t(z)$ defined by $N(0, \frac{1}{t^2})$:

$$q_t(z) = \frac{t}{\sqrt{2\pi}} \exp\left(-\frac{z^2 t^2}{2} \right). \tag{10}$$

Obviously, $\{w_t\}$ is bounded almost surely, since

$$\lim_{t \to \infty} w_t = 0, \quad \text{a.s.} \tag{11}$$

Now, we will show that in the above stochastic framework, for any feedback control $u_t \in \mathcal{F}_y^t \triangleq \sigma\{y_i, 0 \leqslant i \leqslant t\}$, there always exist an initial condition y_0 and a set $D \subset \Omega$ with positive probability such that for any $\omega \in D$, $\boldsymbol{\theta}(\omega) \in \Theta_0$ and $|w_t(\omega)| \leqslant w$ hold[1], and that the corresponding output sequence $\{y_t\}$ of the closed-loop control system has a dynamical lower bound. This will naturally gives the corresponding results in the deterministic framework of this paper.

To get the above results, we need a key lemma on a conditional Cramér-Rao inequality for dynamical systems, which may be regarded as an extension of those in [13] and [16]. To this end, we first define some notations which will be used in the sequel. $f_t \triangleq f(\boldsymbol{\theta}, \boldsymbol{\phi}_t)$, where $\boldsymbol{\phi}_t$ is defined by (1). $E_{\boldsymbol{x}}\boldsymbol{y} \triangleq E\{\boldsymbol{y}|\boldsymbol{x}\}$, where $\boldsymbol{x}$, $\boldsymbol{y}$ are random vectors. $\hat{f}(\boldsymbol{\theta}, \boldsymbol{\phi}_t) \triangleq E[f(\boldsymbol{\theta}, \boldsymbol{\phi}_t)|\mathcal{F}_t^y]$; $P_{t+1}(\boldsymbol{\theta}) \triangleq KI + M_1(t+1)^4 \sum_{i=0}^{t} E[f'(\boldsymbol{\theta}, \boldsymbol{\phi}_i) f'^{\tau}(\boldsymbol{\theta}, \boldsymbol{\phi}_i)|\mathcal{F}_t^y]$ with $P_0(\boldsymbol{\theta}) \triangleq KI$; where $\mathcal{F}_t^y \triangleq \sigma\{y_1, \ldots, y_t\}$, $K, M_1 > 0$ are two nonnegative random variables and $f'(\boldsymbol{\theta}, \cdot)$ is the derivative of $f(\boldsymbol{\theta}, \cdot)$ with respect to $\boldsymbol{\theta}$.

1) It is worth noting that by (11), w_t may not satisfy the upper bound in A2) at the first finite time steps, but this will not essentially influence the analysis in what follows

We first give a Cramér-Rao inequality-like result to get the best prediction of the uncertain function $f(\boldsymbol{\theta}, \boldsymbol{\phi}_t)$ given $\{y_1, \dots, y_t\}$.

Proposition 3.1. Let $\boldsymbol{\theta}$ be a random parameter vector with p.d.f. $p(\boldsymbol{\theta})$ defined in (9), and be independent of $\{w_k\}$ which is an independent random sequence with p.d.f. $q_t(z)$ defined in (10). Let the nonlinear function $f(\cdot, \cdot)$ be differentiable with respect to $\boldsymbol{\theta}$. Then for dynamical system (1) with arbitrarily deterministic initial value ϕ_0, for $t \geqslant 0$, we have

$$E_x[f(\boldsymbol{\theta}, \boldsymbol{\phi}_t) - \hat{f}(\boldsymbol{\theta}, \boldsymbol{\phi}_t)]^2 \geqslant \frac{1}{2} E_x^\tau f'(\boldsymbol{\theta}, \boldsymbol{\phi}_t) P_t^{-1}(\boldsymbol{\theta}) E_x f'(\boldsymbol{\theta}, \boldsymbol{\phi}_t), \tag{12}$$

where $\boldsymbol{x} \triangleq \{y_1, \dots, y_t\}$.

The next proposition shows there is a set $D \in \Omega$, on which the outputs of system (1) have the desired dynamical lower bound.

Proposition 3.2. Under the conditions of Proposition 3.1 and Assumptions A3)–A4), if the regressor satisfies $\|\boldsymbol{\phi}_t\| \geqslant h\big(\frac{\delta}{(t+1)^2}\big)$ for all t, where δ is some constant to be defined later on, then there exists some set $D \subset \Omega$ with $P(D) > 0$ such that on this set,

$$y_{t+1}^2 \geqslant \frac{1}{(K_1(t+1)^4 + 4)} \left(\frac{\|E_x f'(\theta, \phi_t)\|^2}{2\lambda_{\max} P_t(\theta)} - 1 \right) - K_2, \qquad \text{a.s.} \tag{13}$$

holds for all $t \geqslant 0$, where $K_1, K_2 > 0$ are some random variables.

To prove Theorem 2.1, we need to prove Propositions 3.1 and 3.2. The proof of Proposition 3.1 involves several lemmas to be given below. The first lemma is a variation of the Cramér-Rao inequality (see [13, 16]).

Lemma 3.1. Let $\boldsymbol{x}$ be a random vector, and let $\boldsymbol{\theta}$ be a parameter vector with p.d.f. $p(\boldsymbol{\theta})$ defined by (9). Also, let $g(\boldsymbol{x}, \boldsymbol{\theta})$ be any measurable vector function having partial derivatives of first order w.r.t. θ, and let there be $E_{\boldsymbol{x}} g(\boldsymbol{x}, \boldsymbol{\theta})$ and $E_x \frac{\partial g(\boldsymbol{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$. Denote by $p(\boldsymbol{x}, \boldsymbol{\theta})$ the joint p.d.f. of $\boldsymbol{x}$ and $\boldsymbol{\theta}$. Then we have

$$\begin{aligned} & E_{\boldsymbol{x}}[g(\boldsymbol{x}, \boldsymbol{\theta}) - E_{\boldsymbol{x}} g(\boldsymbol{x}, \boldsymbol{\theta})]^2 \\ & \geqslant E_{\boldsymbol{x}} \frac{\partial g(\boldsymbol{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left\{ E_{\boldsymbol{x}} \left[\frac{\partial \log p(\boldsymbol{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \cdot \frac{\partial^\tau \log p(\boldsymbol{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \right\}^{-1} E_x^\tau \frac{\partial g(\boldsymbol{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}. \end{aligned}$$

Applying this lemma to the dynamical system defined by (1), we can further get the following result.

Lemma 3.2. Under the conditions of Proposition 3.1, for $t \geqslant 1$, we have

$$E_{\boldsymbol{x}}[f(\boldsymbol{\theta}, \boldsymbol{\phi}_t) - E_{\boldsymbol{x}} f(\boldsymbol{\theta}, \boldsymbol{\phi}_t)]^2 \geqslant \frac{1}{2} E_{\boldsymbol{x}}^\tau f'(\boldsymbol{\theta}, \boldsymbol{\phi}_t) [E_{\boldsymbol{x}} F_t(\boldsymbol{\theta})]^{-1} E_{\boldsymbol{x}} f'(\boldsymbol{\theta}, \boldsymbol{\phi}_t), \tag{14}$$

where $\boldsymbol{x} \triangleq \{y_1, \dots, y_t\}$ and

$$F_t(\boldsymbol{\theta}) \triangleq \sum_{k=1}^t \frac{\partial \log q_k(y_k - f_{k-1})}{\partial \boldsymbol{\theta}} \cdot \sum_{k=1}^t \frac{\partial^\tau \log q_k(y_k - f_{k-1})}{\partial \boldsymbol{\theta}} + KI,$$

with $p(\cdot)$ and $q_t(\cdot)$ being the p.d.f.s of the parameter $\boldsymbol{\theta}$ and noise w_t respectively, and where $K > 0$ is some random variable.

Proof. Directly applying Lemma 3.1, we have

$$\begin{aligned} & E_{\boldsymbol{x}}[f(\boldsymbol{\theta}, \boldsymbol{\phi}_t) - E_x f(\boldsymbol{\theta}, \boldsymbol{\phi}_t)]^2 \\ & \geqslant E_{\boldsymbol{x}}^\tau \frac{\partial f(\boldsymbol{\theta}, \boldsymbol{\phi}_t)}{\partial \boldsymbol{\theta}} \left\{ E_{\boldsymbol{x}} \left[\frac{\partial \log p(\boldsymbol{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \cdot \frac{\partial^\tau \log p(\boldsymbol{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] \right\}^{-1} E_{\boldsymbol{x}} \frac{\partial f(\boldsymbol{\theta}, \boldsymbol{\phi}_t)}{\partial \boldsymbol{\theta}} \\ & = E_{\boldsymbol{x}}^\tau \frac{\partial f(\boldsymbol{\theta}, \boldsymbol{\phi}_t)}{\partial \boldsymbol{\theta}} \left\{ E_{\boldsymbol{x}} \left[\frac{\partial [\log p(\boldsymbol{x}|\boldsymbol{\theta}) + \log p(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}} \cdot \frac{\partial^\tau [\log p(\boldsymbol{x}|\boldsymbol{\theta}) + \log p(\boldsymbol{\theta})]}{\partial \boldsymbol{\theta}} \right] \right\}^{-1} E_{\boldsymbol{x}} \frac{\partial f(\boldsymbol{\theta}, \boldsymbol{\phi}_t)}{\partial \boldsymbol{\theta}}. \end{aligned} \tag{15}$$

By the Bayes rule and the dynamical equation (1), we know that

$$\begin{aligned}p(\boldsymbol{x}|\boldsymbol{\theta}) &= p(y_1, y_2, \ldots, y_t|\boldsymbol{\theta})\\ &= p(y_1|\boldsymbol{\theta})p(y_2|\boldsymbol{\theta}, y_1)\cdots p(y_t|\boldsymbol{\theta}, \ldots, y_{t-1})\\ &= q_1(y_1 - f_0)\cdot q_2(y_2 - f_1)\cdots q_t(y_t - f_{t-1}).\end{aligned}$$

Then, by the matrix Schwarze inequality

$$\begin{aligned}&\left(\sum_{k=1}^{t}\frac{\partial \log q_k(y_k - f_{k-1})}{\partial\boldsymbol{\theta}} + \frac{\partial \log p(\theta)}{\partial\boldsymbol{\theta}}\right)\cdot\left(\sum_{k=1}^{t}\frac{\partial \log q_k(y_k - f_{k-1})}{\partial\boldsymbol{\theta}} + \frac{\partial \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^{\tau}\\ &\quad\leqslant 2\left(\sum_{k=1}^{t}\frac{\partial \log q_k(y_k - f_{k-1})}{\partial\boldsymbol{\theta}}\cdot\sum_{k=1}^{t}\frac{\partial^{\tau} \log q_k(y_k - f_{k-1})}{\partial\boldsymbol{\theta}} + \frac{\partial \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial^{\tau} \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right),\end{aligned}$$

and the following fact (e.g. [13])

$$E_{\boldsymbol{x}}\frac{\partial \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial^{\tau} \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = -E_{\boldsymbol{x}}\frac{\partial^2 \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2},$$

we know that to prove Lemma 3.2, it suffices to show that

$$-E_{\boldsymbol{x}}\frac{\partial^2 \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2} \leqslant KI, \quad \text{a.s.}, \tag{16}$$

where $K > 0$ is some random variable.

First, it is not difficult to show that $\frac{\partial^2 p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2}$ and $\frac{1}{p(\boldsymbol{\theta})}\left(\frac{\partial p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\left(\frac{\partial p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^{\tau}$ are bounded. Then with some simple manipulations, we have

$$\frac{\partial^2 \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2} = -\frac{1}{p^2(\boldsymbol{\theta})}\left(\frac{\partial p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\left(\frac{\partial p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^{\tau} + \frac{1}{p(\boldsymbol{\theta})}\frac{\partial^2 p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2} \geqslant -\frac{C}{p(\boldsymbol{\theta})}I,$$

where $C > 0$ is some constant. Hence

$$-E_{\boldsymbol{x}}\frac{\partial^2 \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^2} \leqslant CI\cdot E_{\boldsymbol{x}}\frac{1}{p(\theta)}. \tag{17}$$

Note that for any integrable random variable X, $E[X|\mathcal{F}_t^y]$ is a.s. bounded (see, e.g., [17, p.145]). Hence, we have $E_x\frac{1}{p(\theta)}$ a.s. bounded since $E\frac{1}{p(\boldsymbol{\theta})} = 1$, which gives (16).

Remark 3.1. Lemma 3.2 still holds for $F_0(\boldsymbol{\theta}) \triangleq KI$ at time $t = 0$. This is because in (15), the following equality

$$E_{\boldsymbol{x}}\left[\frac{\partial \log p(\boldsymbol{x}, \boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial^{\tau} \log p(\boldsymbol{x}, \boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right] = \frac{\partial \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\cdot\frac{\partial^{\tau} \log p(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}, \quad \text{a.s.},$$

will be true for $t = 0$, which together with (16) shows our claim at $t = 0$.

Lemma 3.3. Under the conditions of Proposition 3.1, for any $t \geqslant 1$, we have

$$F_t(\boldsymbol{\theta}) \leqslant M_1 t^4 \sum_{k=0}^{t-1} f'(\boldsymbol{\theta}, \boldsymbol{\phi}_k) f'^{\tau}(\boldsymbol{\theta}, \boldsymbol{\phi}_k) + KI,$$

where $F_t(\boldsymbol{\theta})$ is defined in Lemma 3.2 and $M_1 > 0$ is some random variable.

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

$$\frac{\partial \log q_k(y_k - f_{k-1})}{\partial\boldsymbol{\theta}} = \frac{\partial}{\partial\boldsymbol{\theta}}\left\{-\frac{k^2}{2}(y_k - f_{k-1})^2\right\} = k^2 f'(\boldsymbol{\theta}, \boldsymbol{\theta}_{k-1})w_k.$$

Moreover, by the definition of the p.d.f.s $q_t(z)$ of $\{w_t\}$, we know that, for some random $M_1 > 0$, $\sum_{k=1}^t w_k^2 \leqslant M_1$. Hence, by the matrix Schwarze inequality, we have

$$\left(\sum_{k=1}^t k^2 f'(\boldsymbol{\theta},\boldsymbol{\phi}_{k-1})w_k\right)\left(\sum_{k=1}^t k^2 f'(\boldsymbol{\theta},\boldsymbol{\phi}_{k-1})w_k\right)^\tau \leqslant \left(\sum_{k=1}^t k^4 f'(\boldsymbol{\theta},\boldsymbol{\phi}_{k-1})f^{'\tau}(\boldsymbol{\theta},\boldsymbol{\phi}_{k-1}^\tau)\right)\left(\sum_{k=1}^t w_k^2\right) \leqslant M_1 t^4 \sum_{k=1}^t f'(\boldsymbol{\theta},\boldsymbol{\phi}_{k-1})f^{'\tau}(\boldsymbol{\theta},\boldsymbol{\phi}_{k-1}^\tau),$$

which gives the lemma by the definition of $F_t(\theta)$.

Proof of Proposition 3.1. By Lemmas 3.2, 3.3 and Remark 3.1, Proposition 3.1 is true.

To prove Proposition 3.2, we first prove the following lemma.

Lemma 3.4. Under the conditions of Proposition 3.2, there exists a set $D \subset \Omega$ with positive probability such that on D, $E[y_{t+1}^2|\mathcal{F}_t^y] \leqslant \left(4+K_1(t+1)^4\right)(y_{t+1}^2+K_2)+1$ holds for any $t \geqslant 0$, where $K_1, K_2 > 0$ are some random variables.

Proof. Define

$$\Delta_t \triangleq \left\{\boldsymbol{\theta}\in\Theta_0 : |f(\boldsymbol{\theta},\boldsymbol{\phi}_t)| < \frac{\delta}{(t+1)^2}f^*(\boldsymbol{\phi}_t)\right\}, \quad t\geqslant 0,$$

where δ is defined by A4) and $f^*(\boldsymbol{\phi}_t) \triangleq \max_{\boldsymbol{\theta}\in\Theta_0}\|f'(\boldsymbol{\theta},\boldsymbol{\phi}_t)\|$. Now, we we will prove that

$$0 < \delta < \frac{1}{MP\sum_{t=0}^\infty \frac{1}{(t+1)^2}}, \tag{18}$$

where $P \triangleq \sup_{\boldsymbol{\theta}\in\Theta} p(\boldsymbol{\theta})$, and $M>0$ is defined in A4). In fact, by (9), $P=\frac{cR^2}{2}$. To estimate c, noteing that $\int_{\|\boldsymbol{\theta}\|\leqslant\frac{R}{2}} p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} < 1$, by the definition of c and (9), we immediately have $c\left(\frac{R^2}{2}-\frac{R^2}{4}\right)\frac{\pi R^2}{4} < 1$, and hence $c < \frac{16}{\pi R^4}$. Consequently, $P < \frac{8}{\pi R^2}$. Since $\sum_{t=0}^\infty \frac{1}{(t+1)^2} \leqslant 1+\int_{t=1}^\infty \frac{1}{t^2}\mathrm{d}t \leqslant 2$, we have $MP\sum_{t=0}^\infty\frac{1}{(t+1)^2} \leqslant \frac{16M}{\pi R^2}$. As a result, (3) holds by A4).

Recursively define $\Theta_{t+1} \triangleq \Theta_t - \Delta_t, t=0,1,\ldots,$ where Θ_0 is defined by Assumption A1). Let $\Theta_\infty \triangleq \lim_{t\to\infty}\Theta_t, D \triangleq \{\omega : \boldsymbol{\theta}\in\Theta_\infty\}$.

So, by (3), Assumption A4) and the condition of this lemma, we have $L(\Delta_t) \leqslant \frac{M\delta}{(t+1)^2}$, and hence

$$P\left(\left\{\omega : \boldsymbol{\theta}\in\bigcup_{t=0}^\infty \Delta_t\right\}\right) \leqslant \sum_{t=0}^\infty P(\{\omega : \boldsymbol{\theta}\in\Delta_t\}) = \sum_{t=0}^\infty \int_{\Delta_t} p(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \leqslant PM\sum_{t=0}^\infty \frac{\delta}{(t+1)^2} < 1,$$

which implies

$$P(D) \geqslant 1 - P\left(\left\{\omega : \boldsymbol{\theta}\in\bigcup_{t=0}^\infty \Delta_t\right\}\right) > 0.$$

Now, let $\omega^*\in D$ be a fixed point, and let $\{\boldsymbol{\theta}_t\}$ be a sequence of random variables such that $|f(\boldsymbol{\theta}_t,\boldsymbol{\phi}_t)| = \max_{\boldsymbol{\theta}\in\Theta}|f(\boldsymbol{\theta},\boldsymbol{\phi}_t)|$. Then by the definitions of D and Δ_t, we have

$$\begin{aligned}[f(\boldsymbol{\theta}_t,\boldsymbol{\phi}_t) - f(\boldsymbol{\theta}(\omega^*),\boldsymbol{\theta}_t)]^2 &= [(\boldsymbol{\theta}_t-\boldsymbol{\theta}(\omega^*))^\tau f'(\xi_t,\boldsymbol{\theta}_t)]^2\\ &\leqslant \|\boldsymbol{\theta}_t-\boldsymbol{\theta}(\omega^*)\|^2\cdot\max_\xi\|f'(\boldsymbol{\xi},\boldsymbol{\theta}_t)\|^2 \leqslant \frac{2R(t+1)^4}{\delta}f^2(\boldsymbol{\theta}(\omega^*),\boldsymbol{\phi}_t),\end{aligned} \tag{19}$$

where ξ_t is some random variable, and R is defined in Assumption A1). Consequently, by noting that $w_t^2 \leqslant K_2$, a.s. for some random constant $K_2>0$, and the fact that $\max_{\theta\in\Theta_0} f^2(\boldsymbol{\theta},\boldsymbol{\phi}_t)$ is measurable $\mathcal{F}_t^y$, for any $\omega^*\in D$, we have

$$E_x y_{t+1}^2 = E_x f^2(\boldsymbol{\theta},\boldsymbol{\phi}_t) + Ew_{t+1}^2 \leqslant \max_{\boldsymbol{\theta}\in\Theta_0} f^2(\boldsymbol{\theta},\boldsymbol{\phi}_t)+1$$

$$\leqslant 2f^2(\boldsymbol{\theta}(\omega^*),\boldsymbol{\phi}_t)+2[f(\boldsymbol{\theta}_t,\boldsymbol{\phi}_t)-f(\boldsymbol{\theta}(\omega^*),\boldsymbol{\phi}_t)]^2+1$$
$$\leqslant \left(2+\frac{2R(t+1)^4}{\delta^2}\right)f^2(\boldsymbol{\theta}(\omega^*),\boldsymbol{\phi}_t)+1.$$

Hence,

$$\begin{aligned}[E_x y_{t+1}^2](\omega^*) &\leqslant \left(2+\frac{2R(t+1)^4}{\delta^2}\right)f^2(\boldsymbol{\theta},\boldsymbol{\phi}_t)(\omega^*)+1\\ &=\left(2+\frac{2R(t+1)^4}{\delta^2}\right)[y_{t+1}(\omega^*)-w_{t+1}(\omega^*)]^2+1\\ &=\left(4+\frac{4R(t+1)^4}{\delta^2}\right)(y_{t+1}^2(\omega^*)+K_2)+1\\ &\leqslant(4+K_1(t+1)^4)(y_{t+1}^2(\omega^*)+K_2)+1,\end{aligned}$$

where $K_1=\frac{4R}{\delta^2}$ is a constant. Hence the proof is completed.

Proof of Proposition 3.2. First of all, it is easy to see that $E[w_{t+1}|\mathcal{F}_t^y]=Ew_{t+1}=0$ by (10). By (1) we know that

$$y_{t+1}=[f(\boldsymbol{\theta},\boldsymbol{\phi}_k)-\hat f(\boldsymbol{\theta},\boldsymbol{\phi}_k)]+\hat f(\boldsymbol{\theta},\boldsymbol{\phi}_k)+w_{t+1}, \tag{20}$$

where $\hat f(\boldsymbol{\theta},\boldsymbol{\phi}_k)$ are defined as in Proposition 3.1. Consequently, from the fact that $E[f(\boldsymbol{\theta},\boldsymbol{\phi}_k)–\hat f(\boldsymbol{\theta},\boldsymbol{\phi}_k)|\mathcal{F}_t^y]=0$ and $E[w_{t+1}|\mathcal{F}_t^y]=0$, it follows that, for any $u_t\in\mathcal{F}_t^y$,

$$\begin{aligned}E[y_{t+1}^2|\mathcal{F}_t^y]&=E\{[f(\boldsymbol{\theta},\boldsymbol{\phi}_k)-\hat f(\boldsymbol{\theta},\boldsymbol{\phi}_k)]^2|\mathcal{F}_t^y\}+\hat f^2(\boldsymbol{\theta},\boldsymbol{\phi}_k)+E[w_{t+1}^2|\mathcal{F}_t^y]\\ &\geqslant E\{[f(\boldsymbol{\theta},\boldsymbol{\phi}_k)-\hat f(\boldsymbol{\theta},\boldsymbol{\phi}_k)]^2|\mathcal{F}_t^y\}.\end{aligned}\tag{21}$$

Then by Proposition 3.1, on D, we have

$$E_{\boldsymbol{x}}y_{t+1}^2\geqslant\frac{1}{2}E_{\boldsymbol{x}}^{\tau}f'(\boldsymbol{\theta},\boldsymbol{\phi}_t)P_t^{-1}(\boldsymbol{\theta})E_{\boldsymbol{x}}f'(\boldsymbol{\theta},\boldsymbol{\phi}_t)\geqslant\frac{1}{2}\frac{\|E_{\boldsymbol{x}}^{\tau}f'(\boldsymbol{\theta},\boldsymbol{\phi}_t)\|^2}{\lambda_{\max}P_t(\boldsymbol{\theta})}.$$

This together with Lemma 3.4 shows that Proposition 3.2 holds true.

Proof of Theorem 2.1. By Proposition 3.2, there is a $\omega\in D$, which corresponds to a value $\boldsymbol{\theta}$, a sequence $\{w_t\}$ and some constants M_1,K,K_1,K_2 such that

$$y_{t+1}^2\geqslant\frac{1}{(K_1(t+1)^4+4)}\left(\frac{\|E_{\boldsymbol{x}}f'(\boldsymbol{\theta},\boldsymbol{\phi}_t)\|^2(\omega)}{2\lambda_{\max}P_t(\boldsymbol{\theta})(\omega)}-1\right)-K_2. \tag{22}$$

Note that for any $t\geqslant 0$, $\|\boldsymbol{\phi}_j\|>h((j+1)^{-2}\delta)\geqslant h(\delta), 0\leqslant j\leqslant t$. Since $f_i'(\boldsymbol{\theta},\boldsymbol{\phi}_j)$ is continuous, $i=1,2,\ldots,p$, it is thus not equal to zero for all $\boldsymbol{\theta}\in\Theta_0$ by Assumption A4). This implies that $f_i'(\boldsymbol{\theta},\boldsymbol{\phi}_j)$ is either positive or negative for all $\boldsymbol{\theta}\in\Theta_0$. As a result, for any $\omega\in\Omega$, we have

$$\|E_{\boldsymbol{x}}f'(\boldsymbol{\theta},\boldsymbol{\phi}_t)\|^2=\sum_{i=1}^{p}E_{\boldsymbol{x}}^2f_i'(\boldsymbol{\theta},\boldsymbol{\phi}_t)\geqslant\sum_{j=1}^{p}\min_{\|\boldsymbol{\theta}\|\leqslant R}|f_j'(\boldsymbol{\theta},\boldsymbol{\phi}_t)|^2. \tag{23}$$

Moreover, since $P_t(\boldsymbol{\theta})\geqslant 0$, for any $t\geqslant 1$, we have

$$\lambda_{\max}P_t(\boldsymbol{\theta})\leqslant trP_t(\boldsymbol{\theta})\leqslant M_1t^4\max_{\|\boldsymbol{\theta}\|\leqslant R}\left[\sum_{i=0}^{t-1}f'^{\tau}(\boldsymbol{\theta},\boldsymbol{\phi}_i)f'(\boldsymbol{\theta},\boldsymbol{\phi}_i)\right]+Kp. \tag{24}$$

Obviously, $\lambda_{\max}P_0(\boldsymbol{\theta})=K$. Taking (23) and (24) into (22), we immediately obtain the theorem.

4 Concluding remarks

We have presented a dynamical inequality on the output sequence of a wide class of uncertain nonlinear control systems, and have illustrated how to use this inequality to derive some fundamental limitations to the capability of the feedback principle. Such an inequality for uncertain control systems may play an important role, like the well-known Cramér-Rao bound, in mathematical statistics. More results are to be derived along this line of research.

Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant No. 60821091), and Committee on Research and Conference Grants of The University of Hong Kong (HKU CRCG) (Grant No. 201007176047).

References

1 Aström K J, Wittenmark B. Adaptive Control. 2nd ed. Reading, MA: Addison-Wesley, 1995
2 Chen H, Guo L. Identification and Stochastic Adaptive Control. Boston, MA: Birkhauser, 1991
3 Ioannou P A, Sun J. Robust Adaptive Control. Englewood Cliffs, NJ: Prentice-Hall, 1996
4 Krstić M, Kanellakopoulos I, Kokotović P V. Non-linear and Adaptive Control Design. New York: Wiley, 1995
5 Tao G, Kokotovic P V. Adaptive Control of Systems with Actuator and Sensor Nonlinearities. New York: Wiley, 1996
6 Zames G, Wang L Y. Adaptive vs. robust control: information based concepts. Adapt Syst Contr Signal Proces, 1993, 8: 533–536
7 Zhou K, Doyle J C, Glove K. Robust and Optimal Control. New Jersey: Prentice Hall, 1996
8 Guo L. On critical stability of discrete-time adaptive nonlinear control. IEEE Trans Automat Contr, 1997, 42: 1488–1499
9 Li C Y, Guo L. A new critical theorem for adaptive nonlinear stabilization. Automatica, 2010, 46: 999–1007
10 Li C Y, Guo L. On feedback capability in a class of nonlinearly parameterized uncertain systems. IEEE Trans Automat Contr, 2011, 56: 2946–2951
11 Li C, Xie L-L. On robust stability of discrete-time adaptive nonlinear control. Syst Control Lett, 2006, 55: 452–458
12 Li C, Xie L-L, Guo L. A Polynomial criterion for adaptive stabilizability of discrete-time nonlinear systems. Commun Inform Syst, 2006, 6: 273–298
13 Xie L-L, Guo L. Adaptive control of discrete-time nonlinear systems with structural uncertainties. In: Lectures on Systems, Control, and Information. AMS/IP, 2000. 49–90
14 Xie L-L, Guo L. How much uncertainty can be dealt with by feedback? IEEE Trans Automat Contr, 2000, 45: 2203–2217
15 Xue F, Guo L. On limitations of the sampled-data feedback for nonparametric dynamical systems. J Syst Sci Complex, 2002, 15: 225–250
16 Van Trees H L. Detection, Estimation, and Modulation Thorey: Part I. New York: John Wiley & Sons, 1968
17 Chow Y S, Teicher H. Probability Theory. 3rd ed. New York: Springer, 1997