An asymptotically optimal nonparametric adaptive controller

GUO Lei (郭 雷) & XIE Liangliang (谢亮亮)

Institute of Systems Science, Chinese Academy of Sciences, Beijing 100080, China
Correspondence should be addressed to Guo Lei (email: Lguo@control.iss.ac.cn)

Received October 28, 1999

Abstract For discrete-time nonlinear stochastic systems with unknown nonparametric structure, a kernel estimation-based nonparametric adaptive controller is constructed based on truncated certainty equivalence principle. Global stability and asymptotic optimality of the closed-loop systems are established without resorting to any external excitations.

Keywords: **adaptive control, nonparametric identification, stochastic systems, discrete-time, optimality.**

The primary reason of using feedback in a control system design is to reduce the effects of the system uncertainties on the control performance. The uncertainties of a system usually have two sources: structure uncertainties and external disturbances (noises). In general, the latter is easier to cope with than the former. Robust control and adaptive control are two existing means to deal with structure uncertainties of a system in a controller design. Robust control usually requires that the true system lie in a small ball with a known nominal model as the center, whereas adaptive control does not need such a prerequest.

For both stochastic[1] and deterministic linear systems with small unmodelled dynamics[2] the theory of adaptive control has been well-developed. This theory can easily be generalized to nonlinear systems with linear unknown parameters and with linearly growing nonlinearities[3]. But recently, it has been shown that the generalization of the existing theory to systems with nonlinearities having nonlinear growth rates is impossible in general[4]. This is a fundamental difference between adaptive control of discrete-time and continuous-time systems where no growth constraints on the nonlinearities are imposed[5].

For convenience, let $f(\cdot)$ be an unknown nonparametric function describing the nonlinear dynamics of a control system. There are various approximation techniques (e.g. Volterra series, neural nets, wavelets, etc.), to approximate nonparametric functions by parametric ones. To be precise, for x in a compact set, $f(\cdot)$ can be uniformly approximated by parametric functions of the form

$$g(\theta, x) \triangleq \sum_{i=1}^{N} a_i \sigma(b_i^{\tau} x),$$

where $\sigma(\cdot)$ is a known "basic" function, and a_i and b_i are unknown parameters.

Thus, instead of the original nonparametric model $f(\cdot)$ the above parametric model $g(\theta, x)$ can be used in adaptive control. This has attracted considerable attention[6]. But this approach has several limitations and difficulties. First, in order to ensure that x (which usually represents the system signals) lies in a compact set for reliable approximation, stability of the system must be established first, and the parametric model provides little (if any) help in this regard. Second, searching for the parameters a_i and b_i usually involves global nonlinear optimizations, of which no general efficient way is available up to now, and the on-line combination of the estimation and control (adaptive control) will further complicate the problem. Third, no matter how large the approximation complexity N is, there always exists an approximation error in the model, which will inevitably prevent the parametric-model-based control becoming optimal in general. Hence, it may be of advantages to consider the nonparametric model $f(\cdot)$ directly, and to use the nonparametric estimation methods which are well-developed in the mathematical statistics literature (see ref. [7] and the references therein).

To the best of our knowledge, the first concrete theoretical result on nonparametric adaptive control seems to be due to Oulidi, who proved that the diminishingly excited certainty equivalence nonparametric adaptive control is asymptotically optimal for a class of nonlinear systems with bounded noises[7].

In this paper, we shall show that for systems with Gaussian white noises, the use of external excitations in the controller design is not necessary. Our controller is designed based on a truncated certainty equivalence principle, which automatically sets the function estimate to be zero once it is too large.

1 Main results

Consider the following discrete-time nonlinear control model:

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

where y_t, u_t and ε_t are the d-dimensional system output signals, input signals 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 $\{y_i, \ i \leqslant t\}$ 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 in ref. [7].

Let $K(\cdot)$ be a nonnegative kernel function defined on $\mathbb{R}^d$ satisfying the following conditions:

$$K(0) > 0, \int_{\mathbb{R}^d} K(s)\mathrm{d}s = 1,$$

$$\int_{\mathbb{R}^d} K^2(s)\mathrm{d}s < \infty, \int_{\mathbb{R}^d} \| s \| K(s)\mathrm{d}s < \infty.$$

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

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

where $A > 0$ is a constant.

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

$$\delta_j(x,y) \triangleq K(j^a(x-y)), \quad \forall j > 0, \tag{2}$$

where $a \in \left(0, \frac{1}{2d}\right)$, d being the dimension of the system signals. Let $\delta_0 = 0$.

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

$$\hat{f}_t(y) = \begin{cases} N_t^{-1}(y)\sum_{j=1}^{t}\delta_{j-1}(y_{j-1},y)\times(y_j - u_{j-1}), \text{if } N_t(y) > 0; \\ 0 \qquad \text{otherwise,} \end{cases} \tag{3}$$

where

$$N_t(y) \triangleq \sum_{j=1}^{t}\delta_{j-1}(y_{j-1},y). \tag{4}$$

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

$$h_t = o(\sqrt{\log t}), \quad \text{as} \quad t \to \infty. \tag{5}$$

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

$$u_t = -\hat{f}_t(y_t) I_{(\| \hat{f}_t(y_t) \| \leqslant h_t)} + y_{t+1}^*, \tag{6}$$

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

With this control, the closed-loop system equation is

$$y_{t+1} = f(y_t) - \hat{f}_t(y_t) I_{(\| \hat{f}_t(y_t) \| \leqslant h_t)} + y_{t+1}^* + \varepsilon_{t+1}, \tag{7}$$

which is obviously a nonlinear dynamical system.

In order to analyze the properties of (7), we introduce the following assumptions on the system (1):

A1. The nonlinear function $f(\cdot)$ is Lipschitz continuous, and there exist two constants $\alpha \in (0,1)$ and $\beta \in (0,\infty)$ such that

$$\| f(x) \| \leqslant \alpha \| x \| + \beta, \quad \forall x \in \mathbb{R}^d.$$

A2. $\{\varepsilon_t\}$ is a Gaussian white noise sequence with mean zero and variance I.

A3. The reference signal $\{y_t^*\}$ is bounded.

The main result of this paper is stated as follows:

Theorem 1.1. Consider the control system (1) where the nonlinear function $f(\cdot)$ is completely unknown. Let the assumptions A1—A3 be fulfilled. Then the nonparametric adaptive tracking control defined by (6) is asymptotically optimal in the sense that

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

2 Proof of Theorem 1.1

We preface the proof of Theorem 1.1 with two lemmas.

Lemma 2.1. Under the conditions of Theorem 1.1,

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

where $\tilde{f} = \hat{f} - f$, $c \in \left(0, \frac{1}{2} - ad\right)$ and $\delta \in \left(0, \min\left\{\left(\frac{1}{2} - ad - c\right), (1 - ad - c)a\right\}\right)$.

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

$$\tilde{f}_t(y) = \hat{f}_t(y) - f(y) = \frac{M_t(y)}{N_t(y)} + \frac{L_t(y)}{N_t(y)}, \tag{8}$$

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 (7), Condition (5) and Assumptions A1 and A3, we have

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

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

$$\frac{1}{t} \sum_{j=1}^{t} \| y_j \|^2 = o(\log t). \tag{9}$$

Define

$$z_t \triangleq f(y_t) + u_t, \tag{10}$$

$$\mathscr{F}_{t-1} \triangleq \sigma\{(\varepsilon_j)_{j \leqslant t-1}, (y_j^*)_{j \leqslant t}\}.$$

Then $y_t = z_{t-1} + \varepsilon_t$ and by Assumption A2,

$$\begin{aligned} & E[K(j^a(y_j - y)) | \mathscr{F}_{j-1}] \\ & = \text{const.} \int_{\mathbb{R}^d} \exp\left(-\frac{\| x \|^2}{2}\right) K(j^a(z_{j-1} + x - y)) \mathrm{d}x \\ & = \text{const.} j^{-ad} \int_{\mathbb{R}^d} \exp\left(-\frac{1}{2} \| \lambda j^{-a} + y - z_{j-1} \|^2\right) K(\lambda) \mathrm{d}\lambda \\ & = \text{const.} j^{-ad} \exp\left(-\frac{1}{2} \| \lambda_0 j^{-a} + y - z_{j-1} \|^2\right) \int_{\mathbb{R}^d} K(\lambda) \mathrm{d}\lambda , \end{aligned} \tag{11}$$

where "const." stands for a certain positive constant, the last equality follows from the integral mean value theorem, and $|\lambda_0| \leqslant A$ since $K(\lambda) = 0$, $|\lambda| > A$.

Now, since

$$\| y + A - s \|^2 \leqslant 4 \| y \|^2 + 4 \| s \|^2 + 2A^2,$$

by choosing $0 < c < \frac{1}{2} - ad$, for any $t \geqslant 0$, we have

$$\inf\left\{\exp\left(-\frac{1}{2} \| y + A - s \|^2\right); \ \| s \|^2 \leqslant \frac{c}{4} \log t \text{ and } \| y \|^2 \leqslant \frac{c}{4} \log t\right\} \geqslant \text{const.}\, t^{-c}.$$

Then for $\| z_{j-1} \|^2 \leqslant \frac{c}{4} \log t$ and $\| y \|^2 \leqslant \frac{c}{4} \log t$, we have, by (11),

$$E[K(j^a(y_j - y)) | \mathscr{F}_{j-1}] \geqslant \text{const.}\, j^{-ad} t^{-c} \geqslant \text{const.}\, t^{-ad-c}.$$

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

$$\begin{aligned} & \frac{1}{t} \sum_{j=1}^{t} E[K(j^a(y_j - y)) | \mathscr{F}_{j-1}] \\ & \geqslant \text{const.}\, t^{-ad-c} \times \frac{1}{t} \sum_{j=1}^{t} I_{\left(\| z_{j-1} \|^2 \leqslant \frac{c}{4} \log t\right)} \\ & \geqslant \text{const.}\, t^{-ad-c} \times \frac{1}{t} \sum_{j=2}^{t} I_{\left(\| z_{j-1} \|^2 \leqslant \frac{c}{4} \log j\right)} . \end{aligned} \tag{12}$$

Now, we prove that

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

Actually,

$$1 = \frac{1}{t-1}\sum_{j=2}^{t}\left[I_{\left(\| z_{j-1} \|^2 \leqslant \frac{c}{4}\log j \right)} + I_{\left(\| z_{j-1} \|^2 > \frac{c}{4}\log j \right)} \right]$$

$$\leqslant \frac{t}{t-1}d_t + \frac{1}{t-1}\sum_{j=2}^{t}\frac{\| z_{j-1} \|^2}{\frac{c}{4}\log j}. \tag{13}$$

Let

$$S_j \triangleq \sum_{i=2}^{j} \| z_{i-1} \|^2,\ j \geqslant 2,\ S_1 \triangleq 0.$$

Then $\| z_{j-1} \|^2 = S_j - S_{j-1}$. By (9) and (10), we have

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

Thus,

$$\frac{1}{t}\sum_{j=2}^{t}\frac{\| z_{j-1} \|^2}{\log j} = \frac{1}{t}\sum_{j=2}^{t}\frac{S_j - S_{j-1}}{\log j}$$

$$= \frac{1}{t}\left\{ \sum_{j=2}^{t-1}\left(\frac{S_j}{\log j} - \frac{S_j}{\log(j+1)} \right) - \frac{S_1}{\log 2} + \frac{S_t}{\log t} \right\}$$

$$\leqslant \frac{1}{t}\left\{ \sum_{j=2}^{t-1}\frac{S_j}{j\log^2 j} - \frac{S_1}{\log 2} + \frac{S_t}{\log t} \right\}$$

$$\to 0. \tag{14}$$

On the other hand, apparently $d_t \leqslant 1$. Hence, by (13) and (14), $d_t \to 1$, a.s. Consequently, by (12)

$$\liminf_{t\to\infty} t^{ad+c-1} \times \inf\left\{ \sum_{j=1}^{t} E[K(j^a(y_j - y)) \mid \mathscr{F}_{j-1}] : \| y \|^2 \leqslant \frac{c}{4}\log t \right\} \geqslant \text{const}. \tag{15}$$

By the uniform law of large numbers (Theorem 6.4.34 in ref. [7]), we have

$$\sup\left\{ \left| N_t(y) - \sum_{j=1}^{t-1} E[K(j^a(y_j - y)) \mid \mathscr{F}_{j-1}] \right| : \| y \|^2 \leqslant \frac{c}{4}\log t \right\}$$

$$= o(t^{\sigma}),\ \text{a.s. for all } \sigma > \frac{1}{2}.$$

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

$$\liminf_{t\to\infty} t^{ad+c-1} \inf\left\{ N_t(y) :\ \| y \|^2 \leqslant \frac{c}{4}\log t \right\} \geqslant \text{const}. \tag{16}$$

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

$$\| L_t(y) \| \leqslant \text{const}. \sum_{j=1}^{t-1} K(j^a(y_j - y)) \| y_j - y \|$$

$$\leqslant \text{const}. \sum_{j=1}^{t-1} j^{-a} K(j^a(y_j - y))$$

$$\leqslant \text{const.} + \text{const.} \sum_{j=t_1+1}^{t-1} K(j^a(y_j - y))(N_j(y))^{-a}$$

$$= \text{const.} + \text{const.} \sum_{j=t_1+1}^{t-1} [N_{j+1}(y) - N_j(y)] \cdot (N_j(y))^{-a}$$

$$\leqslant \text{const.} + \text{const.} (N_t(y))^{1-a}/(1-a). \tag{17}$$

Again, applying the uniform law of large numbers (Theorem 6.4.34 in ref. [7]) by the definition of $M_t(y)$, we have

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

Hence, by (16) and (17), it follows from (8) that

$$\sup\left\{ \| \tilde{f}_t(y) \| : \| y \|^2 \leqslant \frac{c}{4}\log t \right\} = o(t^{-\delta}),$$

for all $\delta < \min\left\{\left(\frac{1}{2} - ad - c\right), (1 - ad - c)a\right\}$.

Lemma 2.2. Under the conditions of Theorem 1.1, for any $m \geqslant 1$, we have

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

Proof. By the closed-loop system equation (7), we have

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

For any integer $m \geqslant 1$

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

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

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

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

Hence, by Lemma 2.1,

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

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

Therefore,

$$\sum_{j=1}^{t}\|y_{j+1}\|^{m}\leqslant\alpha_{3}\sum_{j=1}^{t}\|y_{j}\|^{m}$$
$$+\lambda_{1}\lambda_{4}\sum_{j=1}^{t}\|\hat{f}_{j}(y_{j})\|^{m}\times I_{\{\|\hat{f}_{j}(y_{j})\|\leqslant h_{j},\ \|y_{j}\|^{2}>\frac{c}{4}\log j\}}+O(t)$$
$$\leqslant\alpha_{3}\sum_{j=1}^{t}\|y_{j}\|^{m}+\lambda_{1}\lambda_{4}\sum_{j=1}^{t}h_{j}^{m}\frac{\|y_{j}\|^{m}}{\left(\frac{c}{4}\log j\right)^{\frac{m}{2}}}+O(t).$$

Thus, using (5) again, we have

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

Proof of Theorem 1.1. By the closed-loop system equation (7), we have

$$\begin{aligned}&\frac{1}{t}\sum_{j=1}^{t}\|y_{j+1}-y_{j+1}^{*}-\varepsilon_{j+1}\|^{2}\\
&=\frac{1}{t}\sum_{j=1}^{t}\|f(y_{j})-\hat{f}_{j}(y_{j})\|^{2}I_{\{\|\hat{f}_{j}(y_{j})\|\leqslant h_{j}\}}+\frac{1}{t}\sum_{j=1}^{t}\|f(y_{j})\|^{2}I_{\{\|\hat{f}_{j}(y_{j})\|>h_{j}\}}\\
&=\frac{1}{t}\sum_{j=1}^{t}\|f(y_{j})-\hat{f}_{j}(y_{j})\|^{2}I_{\{\|\hat{f}_{j}(y_{j})\|\leqslant h_{j},\ \|y_{j}\|^{2}\leqslant\frac{c}{4}\log j\}}\\
&\quad+\frac{1}{t}\sum_{j=1}^{t}\|f(y_{j})-\hat{f}_{j}(y_{j})\|^{2}\times I_{\{\|\hat{f}_{j}(y_{j})\|\leqslant h_{j},\ \|y_{j}\|^{2}>\frac{c}{4}\log j\}}\\
&\quad+\frac{1}{t}\sum_{j=1}^{t}\|f(y_{j})\|^{2}I_{\{\|\hat{f}_{j}(y_{j})\|>h_{j},\ \|y_{j}\|^{2}\leqslant\frac{c}{4}\log j\}}\\
&\quad+\frac{1}{t}\sum_{j=1}^{t}\|f(y_{j})\|^{2}I_{\{\|\hat{f}_{j}(y_{j})\|>h_{j},\ \|y_{j}\|^{2}>\frac{c}{4}\log j\}}.\end{aligned}\tag{18}$$

Now, by using Lemma 2.1, we have

$$\frac{1}{t}\sum_{j=1}^{t}\|f(y_{j})-\hat{f}_{j}(y_{j})\|^{2}I_{\{\|y_{j}\|^{2}\leqslant\frac{c}{4}\log j\}}=o(1),\tag{19}$$

and

$$\|\hat{f}_{j}(y_{j})\|I_{\{\|y_{j}\|^{2}\leqslant\frac{c}{4}\log j\}}\leqslant\|f(y_{j})\|+o(1).\tag{20}$$

Then by (18) we have

$$\begin{aligned}&\frac{1}{t}\sum_{j=1}^{t}\|y_{j+1}-y_{j+1}^{*}-\varepsilon_{j+1}\|^{2}\leqslant o(1)+\frac{2}{t}\sum_{j=2}^{t}\|f(y_{j})\|^{2}I_{\{\|y_{j}\|^{2}>\frac{c}{4}\log j\}}\\
&+\frac{2}{t}\sum_{j=1}^{t}\|\hat{f}_{j}(y_{j})\|^{2}I_{\{\|\hat{f}_{j}(y_{j})\|\leqslant h_{j},\ \|y_{j}\|^{2}>\frac{c}{4}\log j\}}+\frac{1}{t}\sum_{j=1}^{t}\|f(y_{j})\|^{2}\frac{\|f(y_{j})\|+o(1)}{h_{j}}\\
&+\frac{1}{t}\sum_{j=1}^{t}\|f(y_{j})\|^{2}\frac{\|y_{j}\|^{2}}{\frac{c}{4}\log j}.\end{aligned}\tag{21}$$

To proceed further, by Lemma 2.2 and Assumption A1, we have

$$\frac{1}{t}\sum_{j=2}^{t}\frac{\|f(y_{j})\|^{3}}{h_{j}}=o(1),$$

and

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

Also, by (5),

$$\frac{1}{t}\sum_{j=2}^{t}\|\hat{f}_j(y_j)\|^2 I_{\{\|\hat{f}_j(y_j)\|\leqslant h_j,\ \|y_j\|^2>\frac{c}{4}\log j\}}=\frac{1}{t}\sum_{j=2}^{t}h_j^2\cdot\frac{\|y_j\|^2}{\frac{c}{4}\log j}=o(1).$$

Therefore, it follows from (19) that

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

which is just the conclusion of the theorem.

3 Concluding remarks

For discrete-time systems with dynamics described by an unknown nonparametric nonlinear function and with noises being Gaussian and white, we have shown that the truncated certainty equivalence nonparametric adaptive control is asymptotically optimal, without resorting to any external excitations. In a recent work[8] we have shown that the condition $\alpha\in(0,1)$ appearing in Assumption A1 cannot be replaced by $\alpha\geqslant\frac{3}{2}+\sqrt{2}$ in general. In further investigation, it would be of considerable interest to know whether or not the nonparametric adaptive control defined by (6) is still globally stable and asymptotically optimal for the case where $\alpha\in\left[1,\frac{3}{2}+\sqrt{2}\right)$.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No.69425003) and the National Key Project of China.

References

1. Chen, H. F., Guo, L., Identification and Stochastic Adaptive Control, Boston: Birkhäuser, 1991.
2. Ioannou, P. A., Sun, J., Robust Adaptive Control, NJ: Prentice-Hall, 1996.
3. Wei, C., Guo, L., Adaptive control of a class of nonlinear stochastic systems, in Proc. 1996, Chinese Control Conference, Qingdao, 508—512.
4. Xie, L. L., Guo, L., Fundamental limitations of discrete-time adaptive nonlinear control, IEEE Trans. Automat. Contr., 1999, 44(9): 1777.
5. Krstić, M. I., Kanellakopoulos K., P. V., Nonlinear and Adaptive Control Design, New York: John Wiley & Sons, 1995.
6. Narendra, K. S., Parthasarathey, K., Identification and control of dynamical systems using neural networks, IEEE Trans. Neural Networks, 1990, (1): 4-27.
7. Duflo, M., Random Iterative Models, Berlin: Springer-Verlag, 1997.
8. Xie, L. L., Guo, L., How much uncertainty can be dealt with by feedback? IEEE Trans, Automatic Control., to appear.