15th Triennial World Congress, Barcelona, Spain

STOCHASTIC ADAPTIVE SWITCHING CONTROL BASED ON MULTIPLE MODELS

Yanxia Zhang, Lei Guo

Institute of Systems Science, Chinese Academy of Sciences
Beijing, 100080, P.R.China
E-mail: Lguo@control.iss.ac.cn

Abstract: It is well known that the transient behaviors of the traditional adaptive control may be very poor in general, and that the adaptive control designed based on switching between multiple models is an intuitively appealing and practically feasible approach to improve the transient performances. This paper proves that for a typical class of linear systems disturbed by white noises, the multiple model based least-squares (LS) adaptive switching control is stable and convergent, and has the same convergence rate as that established for the standard least-squares-based self-tuning regulators. Moreover, the mixed case combining adaptive models with fixed models is also considered.

Keywords: Multiple models, Switching, Least-Squares, Adaptive control, Convergence rate, Optimality.

1. INTRODUCTION

In an uncertain and complex environment, the approach of "optimal" switching is often used for making decisions through predicting and comparing the effects of multiple schemes. In the area of control, the multiple model approach, which has been used to improve estimations and control accuracies, can be traced back at least to 1960s-1970s (see, e.g., Magill, 1965; Lainiotis, 1976). Some practical applications have also been reported (e.g., Moose *et al.*, 1979). In adaptive control, switching controller based on multiple models has also been used to reduce the dependence of the prior knowledge about the systems(cf. Martensson, 1985; Fu *et al.*, 1986). In (Morse, 1996,1997), the use of multiple fixed models was studied. By comparing the prediction errors of the fixed models and switching based on the "certainty equivalence principle", the author defined a supervisory controller, and proved the tracking performance and robustness of the control system; however, the multiple fixed models need to be chosen with care. Recently, (Narendra *et al.*, 1997) introduced and studied the adaptive control problem of the mixed case, where adaptive models are combined with fixed models.

All the above mentioned papers deal with continuous-time systems only. Lately, ((Narendra *et al.*, 2000) tried to extend the results of (Narendra *et al.*, 1997) to discrete-time case. However, when analyzing the RLS based controller, the authors either assume the persistent exciting condition as in (Goodwin *et al.*, 1984), or use essentially a stochastic gradient algorithm which has poor convergent rate in general. Furthermore, the proof of stability seems to be incomplete for the stochastic adaptive control based on mixed multiple models.

In this paper, we will consider a typical class of linear systems disturbed by white noises, and give a rigorous proof of stability and optimality for multiple-models-based minimum variance adaptive control.

2. PROBLEM FORMULATION

Consider the following SISO system:

$$A(z)y_t = B(z)u_{t-1} + w_t,\ t \geq 0 \tag{1}$$

where$\{y_t\}, \{u_t\}$, and $\{w_t\}$ are the system output, input and noise processes respectively. We assume that $y_t = u_t = w_t = 0, \forall t < 0$, $A(z)$ and $B(z)$ are polynomials in the backward-shift operator z:

$$A(z) = 1 + a_1 z + \cdots + a_p z^p,\ p \geq 0$$
$$B(z) = b_1 + b_2 z + \cdots + b_q z^{q-1},\ q \geq 1$$

with a_i, $1 \le i \le p$;and b_j, $1 \le j \le q$ unknown coefficients, p and q the upper bounds on the true orders.

Now, introduce the unknown parameter vector:

$$\theta = [-a_1, \cdots, -a_p, b_1, \cdots, b_q]^T \tag{2}$$

and the corresponding regressor:

$$\varphi_t = [y_t, \cdots, y_{t-p+1}, u_t, \cdots, u_{t-q+1}]^T \tag{3}$$

Then system (1) can be rewritten as

$$y_{t+1} = \varphi_t^T \theta + w_{t+1}, \ t \ge 0 \tag{4}$$

Our control objective is, at any instant t, to construct a feedback control u_t based on the past measurements $\{y_0, \cdots, y_t, u_0, \cdots, u_{t-1}\}$ so that the following averaged tracking error is asymptotically minimized:

$$J_t \triangleq \frac{1}{t} \sum_{i=1}^{t} (y_i - y_i^*)^2 \tag{5}$$

where $\{y_i^*\}$ is a known reference signal.

We need the following standard conditions:

(A.1) The noise sequence $\{w_t, \mathscr{F}_t\}$ is a martingale difference sequence with conditional variance σ^2, i.e.

$$E[w_{t+1}^2 | \mathscr{F}_t] = \sigma^2 > 0, \ a.s.$$

Moreover, there exists a constant $\beta > 2$ such that

$$\sup E\left[|w_{t+1}|^\beta | \mathscr{F}_t \right] < \infty, \ a.s.$$

(A.2) $B(z) \ne 0$, with $|z| \le 1$.
(A.3) $\{y_t^*\}$ is a bounded reference sequence independent of $\{w_t\}$.

We remark that if $\{d_t\}$ is a nondecreasing positive deterministic sequence such that

$$w_t{}^2 = O(d_t), \ a.s. \tag{6}$$

then under Condition (A.1), d_n can be taken as

$$d_t = t^\delta, \quad \forall \delta \in (\frac{2}{\beta}, 1) \tag{7}$$

where β is given by (A.1)(cf. Guo *et al.*, 1991).

Conventional adaptive control is based on single identification model (e.g. LS), which usually leads to large transient errors if the initial values of the algorithm are not properly chosen. In order to improve the transient behaviors of the control algorithms, a natural idea is to use parallel algorithms with multiple different initial values (cf. Narendra *et al.*, 1997). Compared with the previous results, the main contribution of this paper is the design and rigorous proof of stability, optimality and convergence rate of multiple-models-based stochastic adaptive switching control.

3. MULTIPLE MODELS BASED ON LS ALGORITHM

Let $I_1, I_2, \ldots, I_M$ be M predictive models described by

$$I_i: \quad \hat{y}_i(t+1) = \varphi_t^T \hat{\theta}_i(t),$$

$$i = 1, 2, \ldots, M, \ t = 1, 2, \cdots.$$

where φ_t is defined by (3), and $\hat{\theta}_i(t)$ is the estimation of θ given by the ith model at time t, corresponding to the ith initial value $\hat{\theta}_i(0)$. At any instant, one of the models is chosen according to a performance index, and the corresponding controller is used to control the system. The problem is considered in the following two cases.

3.1 Multiple Adaptive Models

First of all, we state some properties of the standard single LS algorithm for the estimation of the unknown parameter θ :

$$\theta_{t+1} = \theta_t + a_t P_t \varphi_t (y_{t+1} - \varphi_t^T \theta_t) \tag{8}$$

$$P_{t+1} = P_t - a_t P_t \varphi_t \varphi_t^T P_t \tag{9}$$

$$a_t = (1 + \varphi_t^T P_t \varphi_t)^{-1} \tag{10}$$

where the initial values θ_0 and $P_0 > 0$ can be chosen arbitrarily.

Let $\{j_t\}$ be a sequence of integers taking values in $\{0, 1, \cdots, d\}$, $d = p + q$, defined by

$$j_t = \underset{0 \le j \le d}{\operatorname{argmax}} |b_{1t} + e_{p+1}^T P_t^{\frac{1}{2}} e_j| \tag{11}$$

where $e_0 = 0$, e_j, $1 \le j \le d$ is the jth column of the $d \times d$ identity matrix, and b_{1t} is the estimate for b_1 given by θ_t.

To guarantee that the estimated "high frequency" gain b_1 is not too small in the minimum variance adaptive control, the LS algorithm can be modified as follows (Guo, 1995):

$$\hat{\theta}_t = \begin{cases} \theta_t, & \text{if } |b_{1t}| \ge \dfrac{\beta_0}{\sqrt{\log r_{t-1} + t}}, \\ \theta_t + P_t^{\frac{1}{2}} e_{j_t}, & \text{otherwise} \end{cases} \tag{12}$$

where $r_t \triangleq 1 + \sum_{i=0}^{t} ||\varphi_i||^2$, and β_0 is an arbitrary positive constant. In practice, β_0 may be taken as a lower bound to $|b_1|$ if it is available.

The following lemma states that the LS-based algorithm (8)-(12) has the same convergence rate as the standard LS. The proof of it is almost the same as Theorem 6.3 in (Guo, 1995).

Lemma 1 Under Conditions (A.1) and (A.2), for any initial values (θ_0, P_0), the estimation $\{\hat{\theta}_t\}$ given by LS-based algorithm (8)-(12) satisfies

$$(H.1) \quad \|\hat{\theta}_t\|^2 = O(\log r_{t-1}), \quad a.s.$$

$$(H.2) \quad \sum_{i=1}^{t} \frac{(\varphi_i^T \tilde{\theta}_i)^2}{1+\varphi_i^T P_i \varphi_i} = O(\log r_t), \quad a.s.$$

$$(H.3) \quad |\hat{b}_{1t}| \geq \frac{c_1}{\sqrt{\log(r_{t-1}+t)}}, \quad a.s.$$

where $c_1 > 0$ is a random variable, $\hat{b}_{1t}$ is the estimate for b_1 given by $\hat{\theta}_t$, and $\tilde{\theta}_t \triangleq \theta - \hat{\theta}_t$.

In the study of switching control using multiple adaptive models, the estimates of the unknown parameter $\hat{\theta}_i(t),\ i = 1, ...M$ are all given by LS-based algorithm (8)-(12). However, the initial values $(\theta_i(0), P_i(0)),\ i = 1, 2, \ldots, M$, are different for the each model I_i.

Denote

$$e_i(t) \triangleq y_t - \varphi_{t-1}^T \hat{\theta}_i(t-1)$$
$$J_i(t) \triangleq \frac{1}{t}\sum_{j=1}^{t} e_i^2(j), \quad i = 1, 2, \ldots, M \tag{13}$$
$$i_t \triangleq \operatorname*{argmin}_{1\leq i\leq M} J_i(t)$$

At any instant t, the model corresponding to the minimum of $J_i(t),\ i = 1, 2, \cdots, M$ is chosen to determine the input $u(t)$, i.e.

$$\hat{y}_{i_t}(t+1) = \varphi_t^T \hat{\theta}_{i_t}(t) = y_{t+1}^* \tag{14}$$

or

$$u_t = \frac{1}{\hat{b}_{1_{i_t}}(t)}\{\hat{a}_{1_{i_t}}(t) y_t + \cdots + \hat{a}_{p_{i_t}}(t) y_{t-p+1} - \hat{b}_{2_{i_t}}(t) u_{t-1} - \cdots - \hat{b}_{q_{i_t}}(t) u_{t-q+1} + y_{t+1}^*\}$$

where $\hat{a}_{j_{i_t}}(t)$ and $\hat{b}_{l_{i_t}}(t)$are the components of $\hat{\theta}_{i_t}(t)$.

Now introduce the following notations:

$$\tilde{\theta}_{i_t}(t) \triangleq \theta - \hat{\theta}_{i_t}(t), \quad \alpha_t \triangleq \frac{(\varphi_t^T \tilde{\theta}_{i_t}(t))^2}{1+\varphi_t^T P_{i_t}(t)\varphi_t} \tag{15}$$

$$\delta_t \triangleq \max_{1\leq i\leq M} \operatorname{tr}\{P_i(t) - P_i(t+1)\} \tag{16}$$

Theorem 1 For the system (1), let the conditions (A.1)-(A.3) be satisfied, and let the control law be defined by (13) and (14). Then the closed-loop system is globally stable, optimal and has the following rate of convergence

$$R_t = O(\log t + \varepsilon_t) \tag{17}$$

where

$$R_t \triangleq \sum_{j=1}^{t} (y_j - y_j^* - w_j)^2 \tag{18}$$

$$\varepsilon_t = (\log t) \max_{1\leq j\leq t} \{\delta_j j^{\varepsilon} d_j\},\ \forall \varepsilon > 0 \tag{19}$$

3.2 *Multiple Fixed-adaptive Mixed Models*

We now suppose, without loss of generality, that $\hat{\theta}_1(t)$ is given by adaptive algorithm and $\hat{\theta}_i(t) = \theta_i,\ i = 2, ...M$ are fixed estimates for the unknown parameter. For the adaptive model, we still use LS-based algorithm (8)-(12).

Denote

$$S_i(t) \triangleq \frac{1}{\log r_{t-1}} \sum_{j=0}^{t-1} \frac{(y_{j+1} - \hat{\theta}_i^T(j)\varphi_j)^2}{1+\varphi_j^T P_j \varphi_j} \tag{20}$$

where P_t is defined by (9). Let

$$I_i(t) \triangleq \max_{1\leq j\leq t}\{S_i(j) - S_1(j)\}, \quad i = 1, \ldots, M \tag{21}$$

$$\wedge_t \triangleq \{i : 1 \leq i \leq M | I_i(t) \leq K\} \tag{22}$$

where $K > 0$ is a constant. Since $I_1(t) \equiv 0$, it is obvious that $1 \in \wedge_t$, hence $\wedge_t \neq \emptyset$. Define

$$i_t \triangleq \operatorname*{argmin}_{i\in\wedge_t} S_i(t) \tag{23}$$

then at time t, $u(t)$ is determined by the following equation:

$$\hat{\theta}_{i_t}^T(t)\varphi_t = y_{t+1}^* \tag{24}$$

Theorem 2 For the system (1), let Conditions (A.1)-(A.3) be satisfied, and let the control law be defined by (23) and (24). Then the closed-loop system is globally stable, optimal and has the following rate of convergence

$$R_t = O(\log t) + (\varepsilon_t) \tag{25}$$

where R_t and ε_t are defined by (18), (19) respectively.

4. PROOFS OF THE MAIN THEOREMS

In this section, we first present two lemmas which are used in the proofs of the main theorems.

Lemma 2 Consider the closed-loop system (1)-(4) with the control given by (13)and (14). If Conditions (A.1)-(A.3) are satisfied, then there exists a positive random process $\{L_t\}$ such that

$$y_t^2 \leq L_t, \quad L_{t+1} \leq (\lambda + c f_t) L_t + \xi_t \quad \forall\, k$$

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

$$f_t = [\alpha_t \delta_t \log(t + r_t)]^2 + \alpha_t \delta_t \tag{26}$$

$$\xi_t = O(d_t \log^4(t + r_t)) \tag{27}$$

Lemma 3 Under the conditions of Lemma 2, the following result holds

$$\|\varphi_t\|^2 = O\{(t+r_t)^\varepsilon d_t\}, \quad a.s. \quad \forall \varepsilon > 0$$

The proofs of the above two lemmas are similar to those for Lemma 6.1 and Lemma 6.2 in (Guo, 1995) (the details will not be repeated here).

Proof of Theorem 1 By the definitions of α_t, it follows that

$$\begin{aligned}\sum_{j=1}^{t} \alpha_j \le \sum_{j=1}^{t} \frac{(\varphi_j^T \tilde{\theta}_1(j))^2}{1+\varphi_j^T P_1(j)\varphi_j} + \cdots \\ + \sum_{j=1}^{t} \frac{(\varphi_j^T \tilde{\theta}_M(j))^2}{1+\varphi_j^T P_M(j)\varphi_j} \\ = O(\log r_t) \end{aligned} \tag{28}$$

Hence

$$\begin{aligned} R_{t+1} &= \sum_{j=0}^{t} (y_{j+1} - y^*_{j+1} - w_{j+1})^2 \\ &= \sum_{j=0}^{t} (\varphi_j^T \tilde{\theta}_{i_j}(j))^2 = \sum_{j=0}^{t} \alpha_j (1+\varphi_j^T P_{i_j}(j)\varphi_j) \\ &= \sum_{j=0}^{t} \alpha_j [1+\varphi_j^T P_{i_j}(j+1)\varphi_j \\ &+ \varphi_j^T (P_{i_j}(j) - P_{i_j}(j+1))\varphi_j] \\ &= O(\log r_t) + O(\sum_{j=0}^{t} \alpha_j \delta_j \|\varphi_j^T\|^2) \end{aligned} \tag{29}$$

From Lemma 3, it is obvious that for every $\varepsilon > 0$,

$$R_{t+1} = O(\log r_t) + O\Big(\max_{1\le j\le t}\{\delta_j (j+r_j)^\varepsilon d_j\} \log r_t\Big) \tag{30}$$

Therefore, for (17), it suffices to prove that $r_t = O(t)$. By Conditions (A.1) and (A.3), it follows that

$$\sum_{j=0}^{t+1} y_j^2 = O(t) + R_{t+1} = O(t) + O\{(t+r_t)^\varepsilon d_t\} \tag{31}$$

By this and Condition (A.2), it follows from (1) that

$$\sum_{j=0}^{t} u_j^2 = O(t) + O\{(t+r_t)^\varepsilon d_t\} \tag{32}$$

Hence

$$\begin{aligned} r_t &= 1 + \sum_{j=0}^{t} \|\varphi_j\|^2 = O(t) + O\{(t+r_t)^\varepsilon d_t\} \\ &= O(t) + O\{(t+r_t)^\varepsilon t^\delta\}, \quad \forall \delta \in (\frac{2}{\beta}, 1) \end{aligned} \tag{33}$$

Take ε small enough such that $\varepsilon + \delta < 1$, then

$$\begin{aligned} \frac{r_t}{t} &= O(1) + O\left((\frac{r_t}{t})^\varepsilon \frac{1}{t^{1-\varepsilon-\delta}} \right) \\ &= O(1) + o\left((\frac{r_t}{t})^\varepsilon \right) \end{aligned} \tag{34}$$

From this, it is seen that $r_t = O(t)$. Hence

$$R_{t+1} = O(\log t) + O(\varepsilon_t), \quad a.s. \tag{35}$$

where ε_t is defined by (19). Obviously $R_t = o(t)$. Moreover, by the definition of J_t and Condition (A.1)(cf. Guo, 1994), it follows that

$$\lim_{t\to\infty} J_t = \sigma^2, \ a.s. \tag{36}$$

Hence the optimality of the control is also true.

Proof of Theorem 2 By the definition of $I_i(t)$, it is seen that for each $i = 1, 2, \ldots, M$, $I_i(t)$ is nondecreasing. If for some i, $\lim_{t\to\infty} I_i(t) > K$, then after a period of time, i will no longer belong to the set $\wedge_t$.

Denote

$$N \triangleq \{i : 1 \le i \le M | I_i(t) \le K, \forall t > 0\}, \tag{37}$$

then there exists some $t_1 > 0$ such that $\wedge_t = N, \ t > t_1$. Obviously, $1 \in N$.

$S_i(t)$ can be rewritten as the following

$$\begin{aligned} S_i(t) = \frac{1}{\log r_{t-1}} \Big(S_i^1(t) + S_i^2(t) + \\ \sum_{j=0}^{t-1} \frac{w_{j+1}^2}{1+\varphi_j^T P_j \varphi_j} \Big) \end{aligned} \tag{38}$$

where

$$S_i^1(t) \triangleq \sum_{j=0}^{t-1} \frac{(\tilde{\theta}_i^T(j)\varphi_j)^2}{1+\varphi_j^T P_j \varphi_j},$$

$$S_i^2(t) \triangleq 2\sum_{j=0}^{t-1} \frac{\tilde{\theta}_i^T(j)\varphi_j w_{j+1}}{1+\varphi_j^T P_j \varphi_j}$$

It follows from martingale convergence theorem that

$$S_i^2(t) = O\Big((S_i^1(t))^{\frac{1}{2}+\eta} \Big), \ \forall \eta > 0 \tag{39}$$

By (H.3) of Lemma 1 and (39), it is obvious that $S_1^1(t) = O(\log r_{t-1})$, and $S_1^2(t) = O(\log r_{t-1})$.

From (38), it follows that

$$\begin{aligned} S_i(t) - S_1(t) = \frac{1}{\log r_{t-1}} \Big(S_i^1(t) + S_i^2(t) \\ - S_1^1(t) - S_1^2(t) \Big) \end{aligned} \tag{40}$$

Therefore, for every $i \in N$, we have

$$
\begin{aligned}
&\frac{1}{\log r_{t-1}}\left(S_i^1(t)+S_i^2(t)\right)\\
&\leq I_i(t)+\frac{1}{\log r_{t-1}}\left(S_1^1(t)+S_1^2(t)\right)\\
&\leq K+O(1) \qquad (41)
\end{aligned}
$$

Combining this with (39), it follows that

$$
\sum_{j=0}^{t-1}\frac{(\tilde{\theta}_i^T(j)\varphi_j)^2}{1+\varphi_j^T P_j\varphi_j}=S_i^1(t)=O(\log r_{t-1}),\ \forall i\in N \quad (42)
$$

Similar to Lemma 2 and Lemma 3, it can be proven that

$$
\|\varphi_t\|^2=O\{(t+r_t)^{\varepsilon}d_t\},\ a.s.\ \forall\varepsilon>0 \qquad (43)
$$

Hence, for any $t>t_1$

$$
\begin{aligned}
&\sum_{j=0}^{t}(y_{j+1}-y_{j+1}^*-w_{j+1})^2\\
&=\sum_{j=0}^{t_1}(y_{j+1}-y_{j+1}^*-w_{j+1})^2+\sum_{t_1+1}^{t}(\tilde{\theta}_{i_j}^T(j)\varphi_j)^2\\
&\leq O(1)+\sum_{i\in N}\sum_{j=0}^{t}\frac{(\tilde{\theta}_i^T(j)\varphi_j)^2}{1+\varphi_j^T P_j\varphi_j}(1+\varphi_j^T P_j\varphi_j)\\
&=O(\log r_t)+\sum_{i\in N}\sum_{j=0}^{t}\frac{(\tilde{\theta}_i^T(j)\varphi_j)^2}{1+\varphi_j^T P_j\varphi_j}\delta_j\|\varphi_j\|^2\\
&=O(\log r_t)+O\Big(\max_{0\leq j\leq t}\{\delta_j(j+r_j)^{\varepsilon}d_j\}\log r_t\Big)
\end{aligned}
$$

The rest of the proof proceeds along the same lines as in Theorem 1.

5. SIMULATION RESULTS

In Sections 3 and 4, the convergence and optimality properties of stochastic adaptive control using multiple models are discussed. In this section, we use a simple example to test the effect of switching control on the performance of the adaptive systems.

5.1 The Problem

Consider the following linear time-invariant discrete-time plant described by

$$
\begin{aligned}
y(t+1)&=3y(t)+0.5y(t-1)+u(t)\\
&\quad+0.5u(t-1)+w(t+1)
\end{aligned}
$$

where $\{w(t)\}$ is a white noise sequence which is normally distributed with zero mean and variance $\sigma^2=0.04$.

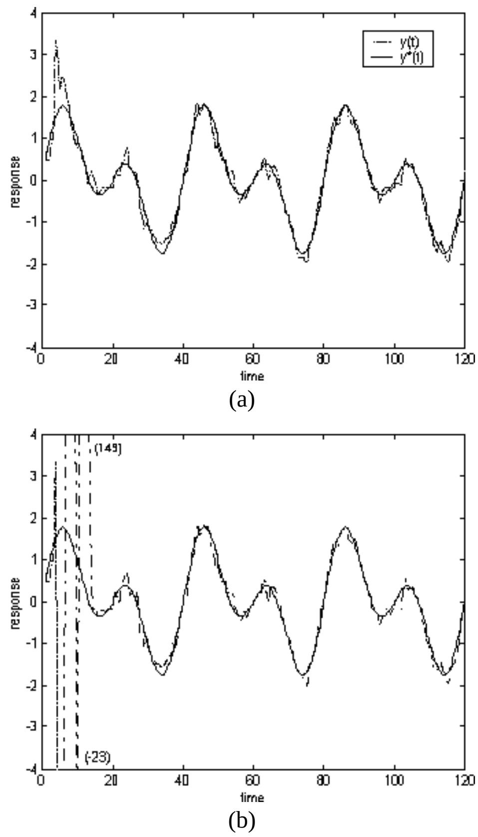


Fig.1 Comparison between multiple adaptive models and a single adaptive model

The above plant can also be written as

$$
y(t+1)=\theta^T\varphi(t)+w(t+1)
$$

where

$$
\begin{aligned}
\varphi(t)&=[y(t),y(t-1),u(t),u(t-1)]\\
\theta&=[3,\ 0.5,\ 1,\ 0.5]
\end{aligned}
$$

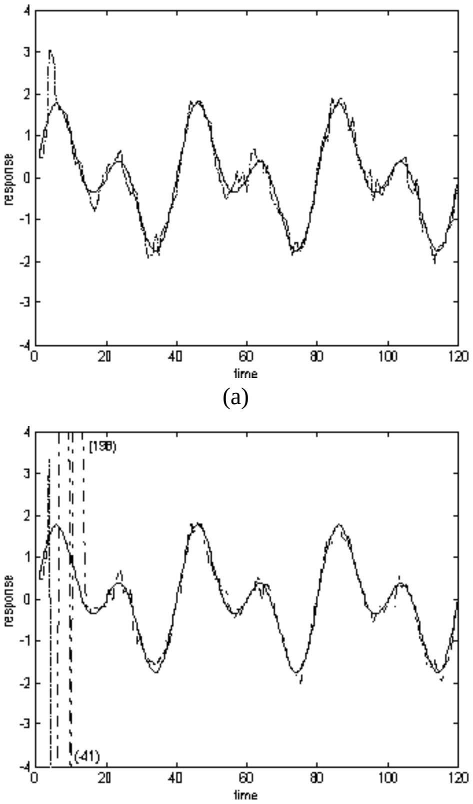
θ is an unknown parameter vector of the plant that has to be estimated. The objective of the control is to track a reference signal $y^*(t)$ described by (Narendra *et al.*, 2000)

$$
y^*(t)=\sin(\frac{\pi t}{20})+\sin(\frac{\pi t}{10})
$$

5.2 The Simulations

Simulation 1: The comparison of the transient responses between switching controller based on multiple adaptive models and controller based on single adaptive model is shown in Fig.1(a) and (b). For the switching controller, four adaptive models are used, which have the following initial values respectively:

$$
\begin{aligned}
\theta_1&=[4.5,\ 0.8,\ -0.7,\ 0.3]\\
\theta_2&=[3.15,\ 0.65,\ 1.13,\ 0.45]\\
\theta_3&=[3.53,\ 0.59,\ 1.65,\ 0.47]\\
\theta_4&=[4.2,\ 0.9,\ 1.6,\ 0.8]
\end{aligned}
$$



(a)

(b)

Fig.2 Comparison between mixed models and a single adaptive model

At each instant, the performance index $J_i(t) = (1/t)\sum_{j=1}^{t} e_i^2(j)$ is computed for all the models, and the model corresponding to the minimum of $J_i(t)$ is chosen to determine the control input. The response of switching controller based on multiple models is found to be satisfactory (see Fig.1(a)). However, the single adaptive model based controller with initial value $\theta_0 = \theta_1$, will result in large transient errors as shown in Fig.1(b).

Simulation 2: This experiment compares the transient response of switching controller based on fixed-adaptive mixed models with that of controller based on single adaptive model. Fig.2(a) corresponds to the switching controller, where the initial value of the adaptive model is

$$\bar{\theta}_1 = [4.5,\ 0.6,\ 0.4,\ 0.13],$$

and the fixed models are specified by the following parameter vectors $\bar{\theta}_i (i = 2, 3, 4)$:

$$\bar{\theta}_2 = [3.01,\ 0.54,\ 1.12,\ 0.56]$$
$$\bar{\theta}_3 = [4.1,\ -0.67,\ 2.3,\ 1.2]$$
$$\bar{\theta}_4 = [0.28,\ 0.52,\ 0.91,\ 0.48]$$

Fig.2(b) shows the response of the controller based on a single adaptive model, where the initial estimation of θ is $\bar{\theta}_0 = \bar{\theta}_1$. It is obvious that the response of switching controller is much more satisfactory.

References

Fu,M. and B.R.Barmish (1986). “Adaptive stabilization of linear systems via switching controls,” *IEEE Trans.Automat.Contr.*, Dec. pp.1079-1103.

Goodwin,G.C. and K.Sin (1984), *Adaptive Filtering, Prediction, and Control.* Englewood Cliffs, NJ: Prentice Hall.

Guo,L. and H.F.Chen (1991) “The Åström-Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers,” *IEEE Trans.Automat.Contr.*, vol.36,No.7, pp.802-812.

Guo,L. (1994), “Further results on least squares based adaptive minimum variance control,” *SIAM.J.Control Optimi.* vol.32(1), pp.187-211.

Guo,L. (1995), “Convergence and logarithm laws of self-tuning regulators,” *Automatica*,vol.31,No.3, pp.435-450.

Lainiotis,D.G. (1976), “Partitioning: A unifying framework for adaptive systems,Part I and II,” *Proc. IEEE*, vol.64, Part I, pp.1126-1142,Part II, pp.1182-1197.

Magill,D.T. (1965), “Optimal adaptive estimations of sampled stochastic processes,” *IEEE Trans. Automat. Contr.*, vol.AC-10, pp.434-439.

Martensson,B. (1985), “The order of any stabilizing regulator is sufficient information for adaptive stabilization,” *Syst. Contr. Lett.*, vol.6, pp.87-91.

Moose,R.L. *et al.*(1979), “Modeling and estimation for tracking maneuvering targets,” *IEEE Trans. Aerosp. Elec. Syst.*, vol.AES-15, pp.448-456.

Morse,A.S. (1996), “Supervisory control of families of linear set-point controllers-part I:Exact matching,” *IEEE Trans. Automat. Contr.*, vol.41, pp.1413-1431.

Morse,A.S. (1997), “Supervisory control of families of linear set-point controllers-part II:Robustness,” *IEEE Trans.Automat. Contr.*, vol.42, pp.1500-1515.

Narendra,K.S.and J.Balakrishnan(1997), “Adaptive control using multiple models and switching,” *IEEE Trans. Automat. Contr.*, vol.42, pp.171-187.

Narendra,K.S. and Cheng Xiang (2000), “Adaptive control of discrete-time systems using multiple models,” *IEEE Trans. Automat. Contr.*, vol.45, pp.1669-1686.