The Rationale for Second Level Adaptation

Kumpati S. Narendra, Yu Wang and Wei Chen

Center for Systems Science, Yale University

Abstract—Recently, a new approach to the adaptive control of linear time-invariant plants with unknown parameters (referred to as second level adaptation), was introduced by Han and Narendra in [1]. Based on $N(\geq m+1)$ fixed or adaptive models of the plant, where m is the dimension of the unknown parameter vector, an unknown parameter vector $\alpha \in R^N$ is estimated in the new approach, and in turn, is used to control the overall system. Simulation studies were presented in [1] to demonstrate that the new method is significantly better than those that are currently in use.

In this paper, we undertake a more detailed examination of the theoretical and practical advantages claimed for the new method. In particular, the need for many models, the proof of stability, and the improvement in performance and robustness are explored in depth both theoretically and experimentally.

I. Introduction

For over two decades, there has been a great deal of interest in developing new methods for rapidly identifying and controlling linear time-invariant plants with large parametric uncertainties in a stable fashion. It is argued that such methods can also be used to adaptively control linear systems with time-varying parameters without losing stability. Recently, second level adaptation was proposed in [1], where the information generated by multiple adaptive models can be used at a second level to identify the plant rapidly. If the unknown plant parameter vector is of dimension m, and belongs to a compact set, it is well known that $(m + 1)$ points can be chosen in parameter space so that the unknown vector lies in their convex hull. This is the starting point of second level adaptation. If $m+1$ models are chosen with these parameters as their initial values, it was argued in [1] that the information derived from them could be used to derive an equivalent parameterization of the plant which converges much more rapidly. The basic ideas involved in the new approach can be summarized as follows, for use throughout the paper:

1) The set $S_\theta \subset R^m$ in parameter space in which the unknown parameter θ_p can lie, is convex.
2) Parameters $\theta_1(t_0), \theta_2(t_0), \cdots, \theta_{m+1}(t_0)$ can be chosen such that θ_p lies in the convex hull of $\theta_i(t_0), \{i = 1, 2, \cdots, m+1\}$. This implies that constants $\alpha_1, \alpha_2, \cdots, \alpha_{m+1}$ with $0 \leq \alpha_i \leq 1$ and $\sum_{i=1}^{m+1} \alpha_i = 1$ exits such that $\sum_{i=1}^{m+1} \alpha_i \theta_i(t_0) = \theta_p, \theta_p \in \Omega$.
3) If $(m+1)$ adaptive models are used to identify θ_p at every instant, it can be shown using arguments of linearity and convexity that θ_p also lies in the convex hull of $\theta_i(t)$, (*i.e.* $\sum_{i=1}^{m+1} \alpha_i \theta_i(t) = \theta_p$ and $\theta_i(t)(i = 1, 2, \cdots, n+1)$ are the estimates of θ_p generated by the adaptive models) provided the initial states of the plant and models are

identical. It is worth noting that when the state variables of the plant are accessible, the initial conditions of the states of the models can be chosen to be identical to that of the plant.
4) If the $(m + 1)$ models are non-adaptive but constant (with θ_i constant), $\sum_{i=1}^{m+1} \alpha_i e_i(t) \equiv 0$, where $e_i(t)$ are the $(m + 1)$ identification errors.
5) Since $\theta_i(t)$ in (4) and $e_i(t)$ in (5) are accessible, the conditions

$$\sum_{i=1}^{m+1} \alpha_i \theta_i(t) = \theta_p$$
$$\text{or} \quad \sum_{i=1}^{m+1} \alpha_i e_i(t) = 0 \tag{1}$$

can be used to estimate the values of the constants α_i. Since by equation (1), the vector α can be considered as an alternative parameterization of the unknown plant, it can be used to estimate θ_p and hence the control input $u(\cdot)$.
6) The estimate $\hat{\alpha}(t)$ of α can be determined either algebraically, or by using a differential equation as described later.

In this paper we will be interested in the following two principal questions:

(i) under what conditions would we expect second level adaptation to result in better performance than first level (or conventional) adaptation?
(ii) why is the overall system stable even when control is based on the estimate $\hat{\alpha}(t)$, rather than the true value of α of the unknown parameter vector? (*i.e* proof of stability)

In addition, we will also be interested in questions related to design and performance such as:

(iii) can the performance and robustness of second level adaptation be improved by increasing the number of models? If additional models are to be used, where should they be located for maximal marginal benefit?
(iv) are fixed models or adaptive models to be preferred? Can rational reasons be provided for combining such models?

Starting with a brief review of the theory of second level adaptation based on both adaptive and fixed models, we shall discuss the above questions in detail and compare both theoretically and using simulations, the new approach with

conventional first level adaptive control. The paper will conclude with simulations studies of adaptation in rapidly time-varying environments and a discussion of future directions for research.

II. MATHEMATICAL PRELIMINARIES

The principal ideas involved in second level adaptation using multiple adaptive models and fixed models are discussed briefly in this section. These are used in all the discussions throughout the paper.

A. Second Level Adaptation (Adaptive models)

We first consider the simplest case where the plant to be controlled adaptively is in companion form, and all the state variables of the plant are accessible. The plant is described by the state equations:

$$\Sigma_p : \quad \dot{x}_p = A_p x_p(t) + bu(t) \tag{2}$$

where $x_p(t) \in r^n, u(t) \in R$ and A_p and b are in companion form and defined as

$$A_p = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \end{bmatrix}, b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \tag{3}$$

The parameters $\{a_i\}$ are unknown and it is assumed that the plant is unstable. If $\theta_p^T = [a_1, a_2, \cdots, a_n]$, we shall refer to θ_p as the unknown plant parameter vector. We further assume that θ_p belongs to a convex set S_θ, where S_θ is a hypercube in parameter space defined by

$$S_\theta = \{\theta_p | \underline{\theta}_p \leq \theta_p \leq \bar{\theta}_p\} \tag{4}$$

The reference model Σ_m is also in companion form and defined by the differential equation

$$\Sigma_m : \quad \dot{x}_m = A_m x_m(t) + br \tag{5}$$

As in equation (3), the last row of A_m is θ_m^T and is chosen by the designer to make A_m stable. The reference input $r(t)$ is uniformly bounded and specified.

B. Statement of the Problem

The problem of adaptive control is to choose the input $u(t)$ in the presence of parameter uncertainty so that the output $x_p(t)$ of the plant follows asymptotically the output $x_m(t)$ of the reference model (*i.e.* $\lim_{t \to \infty} \|x_p(t) - x_m(t)\| = 0$). Direct and indirect control are two distinct methods for achieving the above objective.

<u>Direct Control</u>: In this case, the input $u(t)$ is chosen as $u(t) = k^T(t)x_p(t) + r(t)$, where $k(t)$ is adjusted adaptively. Since it is known a prior that a constant gain vector k^* exists such that $\theta_p + k^* = \theta_m$, $k(t)$ has to be adjusted such that

$\lim_{t \to \infty} k(t) = k^*$. If the equations describing the plant and the reference model are respectively

$$\dot{x}_p = A_p x_p + bu$$
$$\dot{x}_m = A_m x_m + br$$
$$A_m = A_p + bk^{*T}$$

the input $u(t)$ is chosen as

$$u(t) = r(t) + k^T(t)x_p(t)$$

where $\dot{\tilde{k}} = -e_c^T P b x_p(t), \tilde{k} = k - k^*, A_m^T P + P A_m = -Q < 0$. $e_c(t) = x_p(t) - x_m(t)$ is the control error. If $V(e_c, \tilde{k}) = e_c^T(t)Pe_c(t) + \tilde{k}^T\tilde{k}$ is a Lyapunov function candidate, it follows that

$$\frac{dV(t)}{dt} = -e_c^T Q e_c \leq 0 \tag{6}$$

Hence, the system is stable and $e_c(t)$ and $k(t)$ are bounded. Further, since $\dot{e}_c(t)$ is bounded, it follows that $\lim_{t \to \infty} e_c(t) = 0$ and $\lim_{t \to \infty} \tilde{k}(t) = 0$ if the input $r(t)$ is persistently exciting.

<u>Comment</u>: In direct control, the feedback parameter $k(t)$ is directly adjusted using the control error and system identification (or parameter estimation) is not involved.

<u>Indirect Control</u>: In contrast to direct control described earlier, in indirect control, the unknown parameter vector θ_p is estimated at every instant as $\theta_1(t)$ and in turn used to control the system. To assure the stability of the estimator, the following model is used:

$$\dot{\hat{x}}_p(t) = A_m \hat{x}_p(t) + [A_1(t) - A_m]x_p(t) + bu \tag{7}$$

where A_m is a stable matrix, A_m and A_1 (like A_p) are in companion form and the last rows of the two matrices are respectively θ_m^T and θ_1^T. The later is an estimate of the unknown plant parameter vector θ_p. At every instant, $\theta_1(t)$ is used to determine the control parameter vector $k(t)$ as

$$k(t) = \theta_m - \theta_1(t)$$
$$= \theta_m - [\theta_p + \phi_1(t)]$$
$$= k^* - \phi_1(t)$$

where $\phi_1(t)$ is the error in the estimate of θ_p.

Note that in equation (7), A_m is chosen to be the same as the matrix in the reference model (5). However, it can be any stable matrix. It is chosen as in equation (5) to make the control problem simple.

Defining $e_1(t) = \hat{x}_p(t) - x_p(t)$, we have

$$\dot{e}_1 = A_m e_1(t) + b\phi_1^T(t)x_p(t) \tag{8}$$

Using standard arguments of adaptive control, the adaptive law

$$\dot{\theta}_1(t) = \dot{\phi}_1(t) = -e_1^T P b x_p(t) \tag{9}$$

follows readily. The adaptive law is seen to be very similar to that derived from direct control, except that identification rather than control error is used in (9).

Note that the control error is not used in the adjustment of the control parameter $k(t)$. The parametric erorr $\phi_1(t)$ tends to zero when the input $r(t)$ is persistently exciting, and converges to a constant value (if $\phi_1^T x_p(t)$ tends to zero) when it is not. In both cases, the control error tends to zero as seen from equation (8).

Multiple Models (Adaptive): We now consider N models used simultaneously to estimate the unknown plant parameter vectors θ_p. All of them have identical structures

$$\dot{x}_i = A_m x_i + [A_i(t) - A_m]x_p(t) + bu$$

with $x_i(t_0) = x_p(t_0)$, $A_i(t)$ in companion form, and the last row of $A_i(t)$ equal to $\theta_i^T(t)$. This implies that the N parameter error vectors $\phi_i(t)$ and identification error vectors $e_i(t)$ ($i = 1, 2, \cdots, N$) are generated at every instant with the equations

$$\cdot \begin{cases} \dot{e}_i(t) = A_m e_i(t) + b\phi_i^T(t)x_p(t), & e_i(t_0) = 0 \quad \text{(10a)} \\ \dot{\theta}_i(t) = \dot{\phi}_i(t) = -e_i^T(t)Pbx_p(t), & \theta_i(t_0) = \theta_{i0} \text{(10b)} \end{cases}$$

Since $\sum_{i=1}^N \alpha_i \theta_i(t_0) = \theta_p$ or equivalently $\sum_{i=1}^N \alpha_i \phi_i(t_0) = 0$ from equation (10a), it follows that $\sum_{i=1}^N \alpha_i \theta_i(t) \equiv \theta_p$ and from (10b) that $\sum_{i=1}^N \alpha_i e_i(t) = o$. Thus we have the two relations

$$\cdot \begin{cases} \sum_{i=1}^N \alpha_i \theta_i(t) = \theta_p & \text{(11a)} \\ \sum_{i=1}^N \alpha_i e_i(t) = 0 & \text{(11b)} \end{cases}$$

from which α_i can be computed. Further, any m of the $(m+1)$ signals $e_i(t)$ are linearly independent.

Comment: The rapidity with which α can be estimated from equation (11) determines the extent to which second level adaptation will be better than first level adaptation. We note that while θ_p is estimated using a nonlinear equation (10b), α can be estimated using linear equations (11).

From equation (11) it is seen that α_i can be estimated using either equation (11a) or equation (11b). Since we treat the latter in great detail in the context of fixed models, we shall confine our attention to the first approach based on equation (11a):

$$[\Theta(t)]\alpha = \theta_p \quad \text{(12)}$$

where $\Theta(t)$ is the matrix whose columns are the estimates of the parameter vectors at any instant t as given by the N adaptive models. $\Theta(t)$ is an $[m \times (m+1)]$ matrix, $\alpha \in R^{m+1}$ and $\theta_p \in R^m$. The equation can also be expressed as

$$\begin{bmatrix} \theta_1(t) - \theta_{m+1}(t), \theta_2(t) - \theta_{m+1}(t),, \cdots, \theta_m(t) - \theta_{m+1}(t) \end{bmatrix}$$
$$\cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = \theta_p - \theta_{m+1}$$
$$\Phi(t)\bar{\alpha} = \bar{l}(t)$$
$$\text{(13)}$$

where $\Phi(t) \in R^{m \times m}, \bar{\alpha} \in R^m$ and $\bar{l}(t) = \theta_p - \theta_{m+1}(t) \in R^m$. $\theta_i(t) - \theta_{m+1}(t)$ are linearly independent for all t and $i \in \{1, 2, \cdots, m\}$ and in equation (13), both α and θ_p are unknown. However, if $\Phi(t)$ is known at two instants of time t_1 and t_2, we can solve $\bar{\alpha}$ and hence α in equation (14). This, in turn, can be used to compute θ_p and the feedback control vector

$$[\Phi(t_1) - \Phi(t_2)]\bar{\alpha} = \bar{l}(t_1) - \bar{l}(t_2) = [\theta_{m+1}(t_1) - \theta_{m+1}(t_2)]$$
$$\text{(14)}$$

An alternative approach would be to consider the derivative of $\Phi(t)$ which yields

$$\begin{aligned} & \dot{\Phi}(t)\bar{\alpha} = -\dot{\theta}_{m+1}(t) \\ \text{or} \quad & -b^T P[e_{1(2)}, e_2(t), \cdots, e_M(t)]\bar{\alpha} = -b^T Pe_{m+1}(t) \\ \text{or} \quad & E(t)\bar{\alpha} = -e_{m+1}(t) \end{aligned}$$
$$\text{(15)}$$

which is the same as that obtained in the following section.

Multiple Models (Fixed): The principal results to come out of the previous analysis, which are contained in equation (11), are that $\sum_{i=1}^N \alpha_i \theta_i(t) = \theta_p$ at every instant of time t, and that $\sum_{i=1}^N \alpha_i e_i(t) = 0$. A brief examination reveals that the latter result is also true even if the models are fixed (i.e θ_i are constant and not estimated online). Hence once again, α_i can be determined by observing the output error vector $e_i(t)$ of the N models.

If α_i are known, the unknown parameter vector θ_p can be computed using equation (11), and in turn used to compute the control parameter vector.

Note: When multiple fixed models are used, the vectors θ_i are constant. Hence equation (11a) cannot be used to compute the vector α, and only equation (11b) can be used.

III. STABILITY ANALYSIS

Estimation of α: From the proceeding discussion, it is clear that the speed of adaptation is directly proportional to the speed with which α can be estimated. As in conventional adaptive control, the estimate $\hat{\alpha}$ of α can be obtained either algebraically or using a differential equation. In both cases, the starting point is the equation $\sum_{i=1}^{m+1} \alpha_1 e_i(t) = 0$. This can be represented as a matrix equation as shown below:

$$\begin{bmatrix} e_1(t), e_2(t), \cdots, e_{m+1}(t) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{m+1} \end{bmatrix} = 0 \quad \text{(16)}$$

with $0 \leq \alpha_i \leq 1, \sum_{i=1}^{m+1} \alpha_i = 1$. Since $\alpha_{m+1} = 1 - \sum_{i=1}^m \alpha_i$, we have the matrix equation

$$\begin{bmatrix} e_1(t) - e_{m+1}(t), \cdots, e_m(t) - e_{m+1}(t) \end{bmatrix}$$
$$\cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = -e_{m+1}(t)$$
$$E(t)\bar{\alpha} = -e_{m+1}(t)$$
$$\text{(17)}$$

Properties of the Matrix E(t): The identification errors $e_i(t)$ of the m models $(i = 1, 2, \cdots, m)$ were defined as

$$e_i(t) = x_i(t) - x_p(t) \qquad (18)$$

Since all the columns of the matrix $E(t)$ are of the form $e_i(t) - e_{m+1}(t)$, it follows that the ith column of $E(t)$ is merely $x_i(t) - x_{m+1}(t)$ (i.e the difference in the outputs of two fixed models). Hence, for any input to the $m + 1$ models, the entire matrix $E(t)$ is known, and is independent of the actual plant output $x_p(t)$. It is only the right-hand side of equation (19), given by $-e_{m+1}(t)$ that depends explicitly on the plant output since

$$- e_{m+1}(t) = x_p(t) - x_{m+1}(t) \qquad (19)$$

Matrix Inversion: Since $\theta_i - \theta_{m+1}$ are linearly independent vectors for $i = 1, 2, \cdots, m$. It follows that the columns of $E(t)$ are linearly independent for any finite time t. Hence $E^{-1}(t)$ exists at every instant, so that theoretically α can be estimated as $\alpha = -E^{-1}(t)e_{m+1}$ in an arbitrarily short interval of time. If the plant is unstable, control action with the estimated value of α can be initiated at the same time instant.

Simulation 1 *In view of the importance of this concept in second level adaptation, we show the errors $e_1(t), e_2(t)$ and $e_3(t)$ of three fixed models of an unstable second order system (with no control) over a finite period of time T and the corresponding value of $\alpha_i, (i = 1, 2, 3)$ such that $\sum_{i=1}^{3} \alpha_i e_i(t) \equiv 0, t \in [0, T]$. A stable reference model and an unstable plant are described by the parameter vectors $\theta_m = [-1, -3]^T$ and $\theta_p = [2, 1]$. We estimate $E^{-1}(t)$ as*

$$and \begin{bmatrix} -71.5919 & 5.6376 \\ 132.9577 & -10.8885 \end{bmatrix}, \begin{bmatrix} -118.0.91 & 9.7363 \\ 199.8165 & -16.6202 \end{bmatrix},$$
$$\begin{bmatrix} -334.5461 & 27.8656 \\ 5553.3886 & -46.1118 \end{bmatrix} \qquad (20)$$

at three instants of time $t = 0.5, 1, 1.5$ and compute the estimate $\hat{\alpha}(0.5) = [0.3471, 0.2526, 0.4003]^T$, $\hat{\alpha}(1) = [0.3474, 0.252, 0.4006]^T$, $\hat{\alpha}(1.5) = [0.3492, 0.2491, 0.4017]^T$. Using these values, the feedback parameters can be computed. The errors in the first state variable between plant and reference model, when estimation is carried out at $t = 0.5, 1, 1.5$ and used to control the plant at that instant, are shown in Fig. (1).
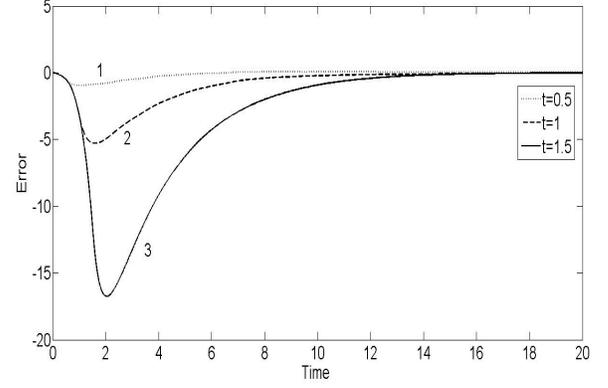

Fig. 1

Note: The above exercise is merely to indicate that the procedure is practically feasible. In practice, the estimates used are derived continuously from differential equations as shown in what follows.

Estimation Using a Differential Equation: For well known reasons related to robustness, α is estimated using a differential equation rather than algebraically. In such a case $\hat{\alpha}(t)$, the estimate of α is determined by the differential equation

$$\dot{\hat{\alpha}} = -E^T(t)E(t)\hat{\alpha} - E^T(t)e_{m+1}(t) \qquad (21)$$

Since the constant vector α satisfies the algebraic equation

$$E^T(t)E(t)\alpha + E^T(t)e_{m+1}(t) = 0 \qquad (22)$$

it follows that the error in the estimate $\tilde{\alpha} = \hat{\alpha} - \alpha$ satisfies the differential equation

$$\dot{\tilde{\alpha}} = -E^T(t)E(t)\tilde{\alpha} \qquad (23)$$

Using the Lyapunov function $V(\tilde{\alpha}) = \frac{1}{2}\tilde{\alpha}^T\tilde{\alpha}$, it follows that

$$\frac{dV(\tilde{\alpha})}{dt} = -\|E(t)\tilde{\alpha}\|^2 < 0 \qquad (24)$$

or equation (23) is asymptotically stable and $\tilde{\alpha} \to 0$. The rate of convergence depends upon the positive definite matrix $E^T(t)E(t)$.

When α is known, the plant together with the feedback gain (control) vector $\theta_m^T - \sum_{i=1}^{m} \alpha_i \theta_i$ is asymptotically stable. With $\hat{\alpha}$ in the feedback path, the overall system is represented by the differential equation

$$\dot{x}_p = A_m x_p + [\Theta\tilde{\alpha}]^T x_p + br \qquad (25)$$

where Θ is a time-invariant matrix and $\lim_{t\to\infty}\tilde{\alpha}(t) = 0$. Hence, $x_p(t)$ tracks $x_m(t)$ asymptotically with zero error. <u>Note</u>: The convergence of $\tilde{\alpha}$ to zero depends only on $E^T(t)E(t)$ and hence on the location of the m models represented by $\theta_1, \theta_2, \cdots, \theta_m$. This provides a rationale for choosing the location of the models.

Comparison of First and Second Level Adaptation: The discussion in the proceeding section brings into better focus the difference between first and second level adaptation.

4

In first level (or conventional) adaptive control, the emphasis is primarily on stability of the overall system. The control error is described by the equation

$$\dot{e}_c = A_m e_c + b\tilde{\theta}^T x_p \tag{26}$$

In attempting to stabilize the system using an approach based on the existence of a Lyapunov function, the adaptive law

$$\dot{\tilde{\theta}} = -e^T P b x_p \tag{27}$$

is arrived at which is nonlinear. As is well known, the analysis of the system becomes complicated when the errors are large. In particular, when the initial conditions $\theta(t_0)$ and $e_c(t_0)$ are large, the solution of the equations (26) and (27) is far from simple. The only fact that has been exploited extensively is that $e_c^T P e_c + \tilde{\theta}^T \tilde{\theta}$ is a Lyapunov function so that e_c and $\hat{\theta}$ are bounded and $lim_{t \to \infty} e_c(t) = 0$. However, very little can be concluded about the transient behavior of the system. Hence, additional methods, as well as arguments, are needed to justify how satisfactory response can be achieved.

In the new approach, based on multiple models, and second level adaptation, scores significantly in such a case, since the differential equation describing the behavior of $\tilde{\alpha}$ continues to be linear. Further, the designer also has considerable prior knowledge of the matrix $E(t)$, which depends only upon the identification models which are known, rather than the plant which is unknown. This in turn, provides much greater control over the performance of second level adaptation.

Example 1 *To compare the performance of first level (conventional) adaptive control and second level adaptive control, simulation studies were carried out in the following example*

$$\theta_m^T = [-24, -8]; \qquad \theta_p^T = [5, 3]$$

and it is known that θ_p lies in the convex hull of the three models

$$\theta_1^T = [-10, -10]; \qquad \theta_2^T = [15, -10]; \qquad \theta_3^T = [5, 15]$$

Due to space limitations, the comparison is made only when the uncertainty is large.
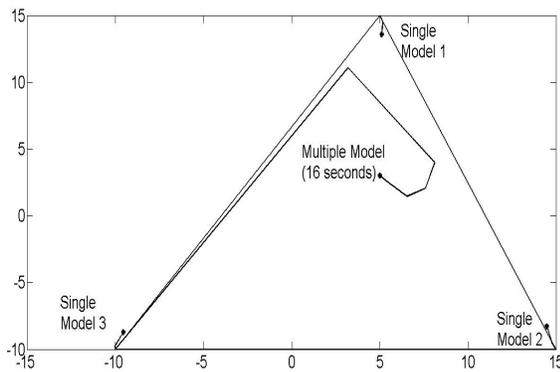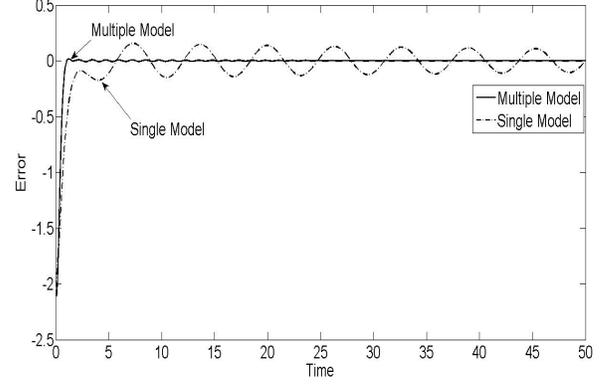

Fig. 2: Parameter Convergence

Fig. 3: Tracking Error

<u>Comment</u>: Fig. (2) demonstrates that the speed of convergence of second level adaptation is orders of magnitude faster than first level adaptation. While the former converges in 16 seconds, it is seen that adaptation of three first level models has barely commenced. In Fig. (3), it is seen that the tracking error (using multiple models) converges in 3 seconds, but the convergence time for the first level is considerably longer.

<u>Time Varying Parameters</u>: Thus far, the analysis has been that of the adaptive control of a time-invariant plant, where θ_p, the plant parameter vector is unknown but constant. We have discussed the advantage of using second level adaptation over conventional adaptive control when the region of uncertainty is large. In such a case, the nonlinear effects in the adaptive algorithm (first level) begin to play a major role in the convergence of the parameter estimate. In this section, we deal with time-varying plant parameters, which are becoming increasingly important in adaptive control.

Let $\theta_p(t)$ be a time-varying plant parameter vector, which lies in the convex hull of the fixed models $\theta_1, \theta_2, \cdots, \theta_{m+1}$. Since $\sum_{i=1}^{m+1} \alpha_i \theta_i = \theta_p(t)$, it follows that $\alpha(t)$ must be time-varying. Since $\theta_i, (i = 1, 2, \cdots, m+1)$ are constant,

$$\left[\theta_1 - \theta_{m+1}, \theta_2 - \theta_{m+1}, \cdots, \theta_m - \theta_{m+1} \right] \bar{\alpha}(t)$$
$$= \theta_p(t) - \theta_{m+1} \tag{28}$$
$$\text{or} \quad \Theta\bar{\alpha}(t) = \theta_p(t) - \theta_{m+1}$$

While the equation (28) cannot be used to determine α, it nevertheless indicates that $\alpha(t)$ can be estimated at least as rapidly as the variations in $\theta_P(t)$.

<u>First Level Adaptation</u>: The behavior of adaptive systems with time-varying parameters has been the subject of the research for several decades. However, useful results are available only when θ_p has rapid variation with small amplitude or slow but large variations. Since $\tilde{\theta} = \hat{\theta}(t) - \theta_p(t)$, $\dot{\tilde{\theta}}(t) = \dot{\hat{\theta}}(t) - \dot{\theta}_p(t)$. In other words, $\dot{\theta}_p(t)$ is a forcing function of the nonlinear error differential equation. This accounts for the difficulty encountered in analyzing adaptive systems with time-varying parameters.

<u>Second Level Adaptation</u>: The simple equations (11a) and (11b), which can be used to determine α, were discussed assuming linearity and time-invariance of the error equations.

These in turn were used to generate the differential equation (21) for determining $\hat{\alpha}$ and consequently the control parameter.

In the time-varying case, the analysis is considerably more complex, but arguments can be provided as to why $\sum_{i=1}^{m+1} \alpha_i(t)e_i(t) = 0$ is still a valid approximation of the equation (11b). In the following analysis, we assume that equation $\sum_{i=1}^{m+1} \alpha_i(t)e_i(t) = 0$ is a valid approximation of the equation (11a). Once again, we use $\hat{\alpha}(t)$, the solution of the differential equation

$$\dot{\alpha}(t) = -E^T(t)E(t)(t)\hat{\alpha} - E^T(t)e_{m+1} \qquad (29)$$

to approximate the time varying $\alpha(t)$. Since $E(t)$ depends only on the output of the fixed identification models, it is not affected by the time variations in the plant parameters. This is only reflected in the term $E^T e_{m+1}(t)$ where $e_{m+1}(t)$ is the output error between the $(m+1)_{th}$ model and the plant. Since $-E^T(t)E(t)$ is an asymptotically stable matrix, the effect of parameters variations are reflected as bounded variations in $\hat{\alpha}(t)$, the estimate of $\alpha(t)$. From the qualitative description of the time-varying problem given above, it is evident that the analysis of the effect of time-varying parameters is considerably simpler than in conventional adaptive control. As in all linear systems, the convergence of the adaptive algorithm can be adjusted using a simple gain parameter.

Example 2 *Using both first and second level adaptation, the following adaptive control problems with time-varying parameters were simulated. The plant is of third order with parameter $\theta_p(t)$. Two specific cases were considered as described below:*

$$\text{Experiment 1}: \quad \theta_p = [3 + sin(0.5t), 4 + cos(0.5t), 3]$$
$$\text{Experiment 2}: \quad \theta_p = [3 + f(t), 4 + f(t), 3]$$

where $f(t)$ is a square wave with mean value zero, amplitude 1 and period of 40 units.

In Experiment 1 the time-varying parameters vary sinusoidally, while in Experiment 2 they vary discontinuously over intervals of 10 units. In both cases, the objective is to track a reference signal, which is the output of a reference model with parameter $\theta_m^T = [-15, -23, -9]$.

The variation of $\theta_p(t)$ in R^2 is shown Fig. (4a); the variations of $f(t)$ as a function of time in Experiment 2 is shown in Fig. (4b).

The response due to first level adaptation and second level adaptation in Experiment 1 are shown in Fig. (5a) & (5b) and that for Experiment 2 are shown in Fig. (6a) & (6b). Fig. (7) show the tracking of the true value of parameter θ_p in Experiment 2 using first level adaptation and second level adaptation. In all cases, second level adaptation is seen to result in significantly better response that first level adaptation.
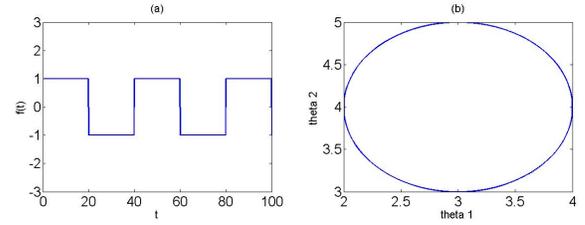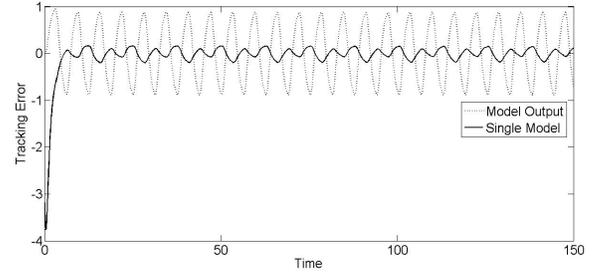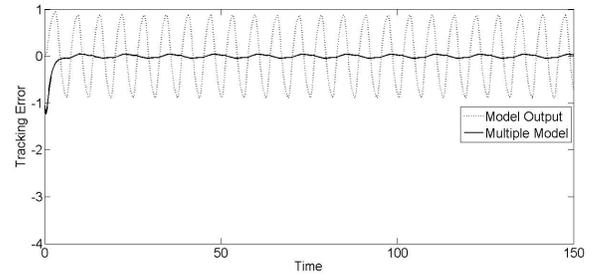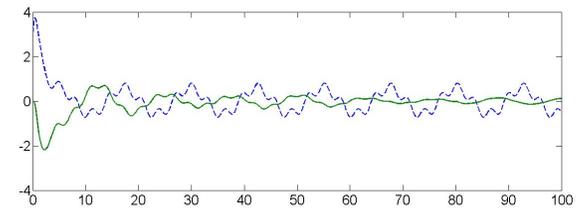

Fig. 4: The parameter variation
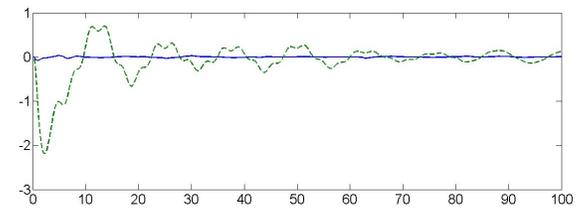

(a) Single Model

(b) Multiple Model

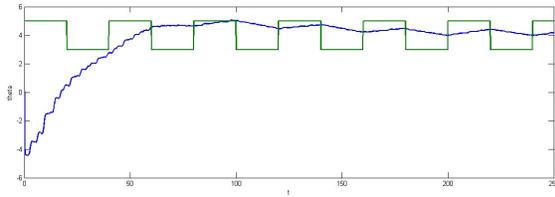Fig. 5: Output Tracking Error of Experiment 1
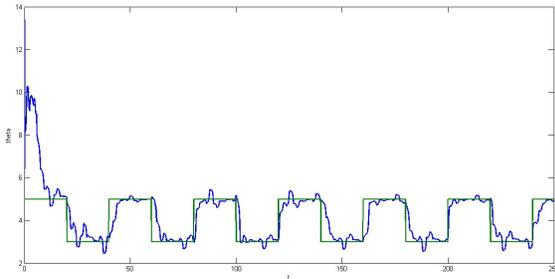

(a) Single Model

(b) Multiple Model

Fig. 6: Output Tracking Error of Experiment 2

(a) Single Model

(b) Multiple Model

Fig. 7: Parameter Tracking Error of Experiment 2

IV. FIXED MODEL AND ADAPTIVE MODELS

It was seen from the discussions in earlier sections that either fixed models or adaptive models can be used for identifying and controlling a plant. Extensive simulation studies have shown that multiple adaptive models (rather than multiple fixed models) result in a smoother response. However, the convergence times are comparable in the two cases. If the plant parameter is not constant but time-varying, the approach using fixed models is found to be better, since in the long run, all adaptive models tend to converge to a single point in parameter space. This make the approach almost equivalent to conventional adaptive control. Hence, depending on the application either fixed or adaptive models can be used in second level adaptation.

The use of fixed and adaptive models for switching and tuning is a well investigated approach for adapting to large uncertainty. The experience gained from those studies are currently being used to determine how the two types of models can be combined for maximum effect in second level adaptation.

V. ROBUSTNESS OF SECOND LEVEL ADAPTATION

The behavior of adaptive systems with different types of perturbations has been studied for decades under the title of "Robustness". In particular,

(i) the effect of input output disturbance

(ii) tine variation in unknown parameters

(iii) the effect of unmodeled dynamics

on the performance of the adaptive system have to be studied in this context.

We have already seen that second level adaptation performs significantly better than first level, both when the region of uncertainty is large and when the plant parameter varies rapidly with time.

When the output of the plant is corrupted with noise, it is seen that the matrix $E(t)$ is unaffected while the signal $e_{m+1}(t)$ contains noise. This results in $\tilde{\alpha}(t)$ having zero mean value due to the noise and hence smaller error in the response of the adaptive system. Hence, second level adaptation is found to be more robust than first level adaptation in two out of three cases of interest. The behavior of second level adaptation in the presence of unmodeled dynamics is currently being investigated.

VI. CONCLUSION

Second level adaptation based on multiple identification models appears to have advantages over conventional or first level adaptive control in regard to every feature of interest to the designer. It is significantly faster with large uncertainties, and/or rapid time-variation of parameters, and more robust in the presence of disturbances. All the advantages are due to the fact that α, which is an alternative parameterization of θ_p can be estimated using linear equations while $\hat{\theta}_p(t)$ is invariable determined by nonlinear equations, due to the importance of stability in adaptive systems.

All the analysis carried out in the paper deals with the plant in companion form, whose state variables are accessible for measurement. Extensions to more general representations, and adaptation using only the input-output of the plant are currently being investigated.

REFERENCES

[1] K. S. Narendra and J. Balakrishnan. Performance Improvement in Adaptive Control Systems Using Multiple Models and Switching. In *Proc. Seventh Yale Workshp. Adaptive Learning Syst.*, 27-33, New Haven, CT, May. 1992. Center for Systems Science, Yale University .

[2] A. S. Morse, D. Q. Mayne, and G. C. Goodwin. Applications of Hysteresis Switching in Parameter Adaptive Control. *IEEE Trans. Automat. Contr .*, 37: 1343-1354, Sep. 1992.

[3] A. S. Morse. Supervisory Control of Families of Linear Set Point Controllers. In *Proc. IEEE 32nd Conf. Decision Contr .*, San Antonio, TX, Dec. 1993.

[4] K. S. Narendra and J. Balakrishnan. Improving Transient Response of Adaptive Control Systems Using Multiple Models and Switching. *IEEE Trans. Automat. Contr .*, 39: 1861-1866, Sep. 1994.

[5] K. S. Narendra and J. Balakrishnan. Adaptive Control Using Multiple Models. *IEEE Trans. Automat. Contr .*, 42:171-187, Feb. 1997

[6] Kumpati S. Narendra and Zhuo Han. Location of Models In Multiple-Model Based Adaptive Control For Improved Performance. In *Proceedings of the 2010 American Control Conference*, 117-122, Baltimore, MD, June 2010

[7] K. S. Narendra and A. M. Annaswamy . *Stable Adaptive Systems.* Prentice-Hall Inc., Englewood Cliffs, NJ 07632, 1989

[8] Zhuo Han and K. S. Narendra. New Concepts in Adaptive Control Using Multiple Models. *IEEE Trans. Automat. Contr .*, 57:78-89, Jan. 2013.