Achieving Optimality in Adaptive Control: the "Bet On the Best" approach [1]

M. C. Campi
Dept. of Electrical Engineering and Automation
University of Brescia
Via Branze 38, 25123 Brescia, Italy
E-Mail: CAMPI@ELET.POLIMI.IT
Phone: ++39.30.3715458
Fax: ++39.30.380014

Abstract

Over the last two decades, the certainty equivalence principle has been the fundamental paradigm in the design of adaptive control laws. It is well known, however, that for general control criterions the performance achieved through its use is strictly suboptimal.

In this paper we introduce a new general philosophy - still based on the certainty equivalence idea - so as to ensure optimality in adaptive control problems under general conditions. Rather than focusing on a particular control scheme, we present the method in a general control setting. Specific control algorithms to cope with different situations can be derived from this general method.

Keywords Adaptive control; stochastic systems; certainty equivalence principle; long-term average cost; optimality.

1 Introduction

An adaptive control problem is a control problem in which some parameter describing the system is known with uncertainty. During the operation of the control system, the controller collects information on the system behavior, therefore reducing the level of uncertainty regarding the value of the parameter. In turn, as the level of uncertainty is reduced, the controller is tuned more accurately on the system parameter so as to obtain a better control result. In this procedure it is essential that the controller chooses the control actions so as to minimize the performance index, as well as probe the system so that uncertainty is reduced to better select future control actions.

In this paper we consider adaptive long-term average optimal control problems and introduce a general philosophy for their solution. In adaptive control, due to the uncertainty affecting the true value of the system parameter, the control law cannot be expected to be optimal in finite time. When the cost criterion is of the long-term average type, however, the control performance in finite time does not affect the asymptotic value of the control cost. Hence, even in an adaptive context, there is a hope to achieve optimality, i.e. to drive the long-term average cost to the value which would have been obtained under complete knowledge of the system. When this happens, we say that the adaptive control law meets the *ideal objective.*

The most common solution methods to adaptive control problems rely on the so-called *certainty equivalence principle*, [1], [2]. The unknown parameter is estimated via some reliable estimation method and the estimate is used as if it were the true value of the unknown parameter.

Certainty equivalent adaptive control schemes have been studied by many authors. In [3] it is proven that a certainty equivalent controller based on the stochastic approximation algorithm achieves the ideal objective for minimum output variance costs. This result has been extended to least squares - minimum output variance adaptive control in [4]. A complete analysis of a minimum output variance self-tuning regulator equipped with the extended least squares algorithm can be found in [5]. Again, the main result is that this adaptive scheme achieves the ideal objective.

The fact that the ideal objective is met in the situations described in the above mentioned papers is due

[1]The author would like to acknowledge the financial support of MURST under the 60% project "Adaptive identification, prediction and control" and the 40% project "Model identification, system control and signal processing".

0-7803-4187-2 4671-4676

to the special properties of the minimum output variance cost criterion. On the other hand, it is well known that the certainty equivalence principle suffers from a general identifiability problem, namely the parameter estimate can converge with positive probability to a false value, e.g. [6], [7], [8], [9]. When a cost criterion other than the output variance is considered, this identifiability problem leads to a strictly suboptimal performance. See e.g. [10], [11], [12], and [13] for a discussion on this problem in different contexts.

This paper goes beyond the straightforward certainty equivalence principle. We show that the ideal objective can be pursued in a general control setting through a control-directed estimation approach aiming at falsifying parameters with an associated progressively increasing optimal cost (Bet On the Best (BOB) philosophy, see Section 3). We prove that a careful use of this approach leads to optimality under general conditions. The theory developed in this paper is not devoted to a particular control scheme. Instead, the proposed approach is intended to be a basic inspirational paradigm. Based on it, many ad-hoc control techniques for specific situations can be designed.

The paper is organized as follows. In Section 2 we present our general control setting and the relevant notations. The BOB-principle is introduced in Section 3 and better formalized in Section 4. Section 5 is devoted to the application of the BOB-principle to a scalar LQG control problem.

2 A general adaptive control setting

An adaptive control problem can be described in terms of four sets $\mathcal{U}$, Θ, $\mathcal{X}$, and $\mathcal{Y}$, a function $c(\cdot,\cdot) : \mathcal{U} \times \mathcal{Y} \to \mathbf{R}$ and two conditional probabilities $\mathcal{P}_{x'|x,u,\theta}$ and $\mathcal{P}_{y|x,u,\theta}$:

- $\mathcal{U}$ — the set of control inputs
- Θ — the set of unknown system parameters
- $\mathcal{X}$ — the set of states
- $\mathcal{Y}$ — the set of outputs
- $c(\cdot,\cdot)$ — the cost function
- $\mathcal{P}_{x'|x,u,\theta} := Pr\{x_{t+1} \leq x' \mid x_t = x, u_t = u, \theta\}$ the probability that the state at time $t+1$ is less than or equal to x' given $x_t = x$, $u_t = u$, and the parameter $\theta \in \Theta$ (here, we assume that $\mathcal{X}$ is a set with an order relation).
- $\mathcal{P}_{y|x,u,\theta} := Pr\{y_t \leq y \mid x_t = x, u_k = u, \theta\}$ the probability that the output at time t is less than or equal to y given $x_t = x$, $u_t = u$, and the parameter $\theta \in \Theta$ (here we also assume that $\mathcal{Y}$ has an order relation).

The adaptive control process takes place as follows. At time t the adaptive controller has access to the observations $o_t = \{u_1, u_2, \ldots, u_{t-1}, y_1, y_2, \ldots, y_t\}$. Based on this, it selects the control input $u_t \in \mathcal{U}$. As a consequence of this control action, the state transits from x_t to x_{t+1} according to the probability distribution $\mathcal{P}_{x'|x,u,\theta}$, a new output y_{t+1} generated according to the probability distribution $\mathcal{P}_{y|x,u,\theta}$ becomes available and the cost $c(u_t, y_t)$ is paid. Then, the observation set is updated to $o_{t+1} = o_t \cup \{u_t, y_{t+1}\}$ and the controller selects the subsequent control input.

The control objective is to minimize the long-term average cost criterion

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} c(u_t, y_t).$$

In doing so, however, the controller has to take care of the problem posed by the fact that the system parameter θ is not known in advance and, therefore, information regarding its value must be accrued through time via the observations u_t and y_t (*adaptive control problem*). A control law is a sequence of functions $l_t : \mathcal{U}^{t-1} \times \mathcal{Y}^t \to \mathcal{U}$, and $l_t(o_t)$ is the corresponding control input after we have observed $o_t = \{u_1, u_2, \ldots, u_{t-1}, y_1, y_2, \ldots, y_t\}$. The set of admissible control laws is denoted by $\mathcal{L}$. Moreover, the optimal control law for the system with parameter θ is denoted by $\{l^*_{\theta,t}\}$ (we assume that $\mathcal{L}$ is large enough such that $\{l^*_{\theta,t}\} \in \mathcal{L}, \forall \theta \in \Theta$). The corresponding optimal long-term average cost is $\mathcal{J}^*_\theta$.

3 The "Bet On the Best" (BOB) principle

The most common approach to adaptive control is to resort to the so-called certainty equivalence principle. The unknown parameter is estimated via some estimation method and the estimate is used in the control law as if it were the true value of the unknown parameter. Unfortunately, the certainty equivalence principle may lead to an estimability problem which results in a strictly suboptimal performance. Here, we present a very simple example which clarifies what can go wrong with the use of this principle. This example will be used as a start for the subsequent discussion.

Example 1 Consider the system

$$x_{t+1} = a^\circ x_t + b^\circ u_t + w_{t+1},$$

where $\{w_t\}$ is an i.i.d. $N(0,1)$ noise process and state x_t is accessible: $y_t = x_t$. Vector $[\ a^\circ \quad b^\circ\ ]$ is unknown but we know that it belongs to a compact set $\Theta = \{[\ a \quad b\ ] : b = 8a/5 - 3/5, a \in [0,1]\}$. Our objective is to minimize the long-term average cost $\limsup_{N\to\infty} 1/N \sum_{t=1}^{N}[qx_t^2 + u_t^2]$, where $q = 25/24$.

In order to determine an estimate of $[\ a^\circ \quad b^\circ\ ]$ the standard least squares algorithm is used. This amounts to selecting at time t the vector $[\ a_t^{LS} \quad b_t^{LS}\ ]$ which minimizes the index $\sum_{k=1}^{t-1}(x_{k+1} - ax_k - bu_k)^2$. Once estimate $[\ a_t^{LS} \quad b_t^{LS}\ ]$ has been determined, according to the certainty equivalence principle the optimal control law for parameter $[\ a_t^{LS}, \quad b_t^{LS}\ ]$ is applied.

Suppose now that at a certain instant point t the least squares estimate is $[\ a_t^{LS} \quad b_t^{LS}\ ] = [\ 1 \quad 1\ ]$. Since the corresponding optimal control law is given by $u_t = -5/8\ x_t$, the squared error at time $t+1$ turns out to be $(x_{t+1} - ax_t - bu_t)^2 = (x_{t+1} - ax_t - (8a/5 - 3/5)(-5/8\ x_t))^2 = (x_{t+1} - 3/8\ x_t)^2, \forall [\ a \quad b\ ] \in \Theta$. The important feature of this last expression is that it is independent of parameter $[\ a \quad b\ ] \in \Theta$. Hence, the term added at time $t+1$ to the least squares index does not influence the location of its minimizer and the least squares estimate remains unchanged at time $t+1$: $[\ a_{t+1}^{LS} \quad b_{t+1}^{LS}\ ] = [\ a_t^{LS} \quad b_t^{LS}\ ] = [\ 1 \quad 1\ ]$. As the same rationale can be repeated in the subsequent instant points, we can conclude that the estimate sticks at $[\ 1 \quad 1\ ]$.

Now, the important fact is that the least squares estimates can in fact take value $[\ 1 \quad 1\ ]$ with positive probability, even when the true parameter is different from $[\ 1 \quad 1\ ]$. Moreover, the optimal cost for the true parameter may be strictly lower than the incurred cost obtained by applying the optimal control law for parameter $[\ 1 \quad 1\ ]$. To see that this is the case, suppose that $[\ a^\circ \quad b^\circ\ ] = [\ 0 \quad -3/5\ ]$ and assume that the system is initialized with $x_1 = 1$ and $u_1 = 0$. Then, at time $t = 2$ the least squares estimate minimizes the cost $(x_2 - a)^2 = (w_2 - a)^2$. Thus, $[\ a_2^{LS} \quad b_2^{LS}\ ] = [\ 1 \quad 1\ ]$ whenever $w_2 > 1$, which happens with positive probability. In addition, it is easily seen that the cost associated with the optimal control law for parameter $[\ 1 \quad 1\ ]$ is $5/3$ whereas the optimal cost for the true parameter $[\ a^\circ \quad b^\circ\ ] = [\ 0 \quad -3/5\ ]$ is $25/24$. □

A careful analysis of the example above reveals where the trouble comes from in a straightforward use of the certainty equivalence principle. When the suboptimal control $u_t = -5/8\ x_t$ is selected based on the current estimate $[\ a_t^{LS} \quad b_t^{LS}\ ] = [\ 1 \quad 1\ ]$, the resulting observation is $y_{t+1} = x_{t+1} = 3/8\ x_t + w_{t+1}$. This observation is in perfect agreement with the one which would have been obtained if $[\ a_t^{LS} \quad b_t^{LS}\ ] = [\ 1 \quad 1\ ]$ were the true parameter. Therefore, there is no reason for having doubts as to the correctness of the estimate $[\ a_t^{LS} \quad b_t^{LS}\ ]$ and thus this estimate is kept unchanged at the next time point.

This is just a single example of a general estimability problem arising in adaptive control problems. This general estimability problem can be described as follows:

- applying to the true system a control which is optimal for the estimated system may result in observations which concur with those that would have been obtained if the estimated system were the true system;

if the estimation method drives the estimate to a value such that the above happens, then

- there is no clue that the system is incorrectly estimated and, consequently, the estimate remains unchanged;

however,

- the adopted control law is optimal for the estimated system, while it may be *strictly* suboptimal for the true system.

A way out of this pernicious mechanism is to employ a more fine grained estimation method based on the optimal long-term average cost for the different systems with parameters $\theta \in \Theta$. Developing this idea will lead us to the formulation of the "Bet On the Best principle".

We start by observing the following elementary fact:

- suppose you apply to the true system a control law which is optimal for another system. If the long-term average cost you pay is different from the optimal cost for this second system, then this system is falsified by the observations and it can be dropped from the set of possible true systems.

Suppose now that at a certain instant point, you select among the systems which are still unfalsified the one with lower optimal cost. Then,

- if you pay a cost different from the expected one, you can falsify this system. In the opposite, you cannot falsify it, but then you are paying a cost which is minimal over the set of possible true systems. Indeed, this implies that you are actually paying the optimal cost for the true system.

These considerations can be summarized as follows: selecting a control law which is optimal for the best unfalsified system (i.e. the system with lower optimal cost among those that are as far unfalsified by the observations) may lead to an estimability problem only when you are achieving optimality. This is in contrast with what happens with the straightforward certainty equivalence principle, where an estimability problem may arise and, yet, the incurred cost may be strictly suboptimal.

The above observations suggest that a very natural way to overcome the estimability problem posed by the certainty equivalence principle is simply to iteratively select among the unfalsified systems the one with minimal optimal cost and then apply the optimal control law for it. Doing so will lead either to falsifying it or to getting optimality. In this way, we arrive to the following principle of general validity:

The "Bet On the Best" (BOB) principle

Consider an adaptive control problem with long-term average cost criterion. At the generic instant point t, do the following:

1. determine the set of unfalsified systems;
2. select the system in the unfalsified set with lower optimal cost;
3. apply the decision which is optimal for the selected system. □

4 Putting the BOB-principle into practice

In this section we better formalize the concept of unfalsified system and point out the properties of the unfalsified set such that applying the BOB-principle leads to optimality.

Let $\mathcal{U}_t$ denote the unfalsified set at time t. Clearly, this set will depend on the observations $o_t = \{u_1, u_2, \ldots, u_{t-1}, y_1, y_2, \ldots, y_t\}$ available at time t, and so it is in fact a stochastic set. Moreover, we note that set $\mathcal{U}_t$ depends through $u_1, u_2, \ldots, u_{t-1}$ on the control law l_k applied from time $k = 1$ to time $k = t-1$. Once the control law $\{l_t\}$ has been fixed, the processes $\{u_t\}$ and $\{y_t\}$ are completely determined and so is the sequence of unfalsified sets $\{\mathcal{U}_t\}$. A (stochastic) sequence of parameters $\{\theta_t\}$, $\theta_t \in \Theta$, $t = 1, 2, \ldots$, is said to be *feasible* if $\theta_t \in \mathcal{U}_t$ $a.s., \forall t$.

In view of the discussion in Section 3, the unfalsified set sequence $\{\mathcal{U}_t\}$ is expected to satisfy the two conditions described below.

i) Consider a (stochastic) sequence of parameters $\{\theta_t\}$ and suppose that the optimal control law $\{l^\star_{\theta_t,t}\}$ for the system with parameter θ_t is admissible. Select the control action to be optimal for θ_t: $u_t = l^\star_{\theta_t,t}(o_t)$. If

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} c(u_t, y_t) > \limsup_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \mathcal{J}^\star_{\theta_t},$$

with positive probability, then $\{\theta_t\}$ is not feasible.

Condition i) says that a sequence of parameters $\{\theta_t\}$ has to be falsified at some instant point whenever the long-term average cost paid by applying the optimal control law for it is strictly larger than the expected average cost.

ii) For any control law $\{l_t\} \in \mathcal{L}$ we have

$$\theta^\circ \in \cup_t \cap_{k>t} \mathcal{U}_k \quad a.s.$$

Condition ii) simply says that the falsification procedure must not be overselective so as to also falsify the true system (note that considering $\cup_t \cap_{k>t} \mathcal{U}_k$ rather than the straightforward $\mathcal{U}_t$ allows for transient phenomena due to stochastic fluctuations).

The following theorem points out the effectiveness of the BOB-principle when conditions i) and ii) are met.

Theorem 1 *Define* $\theta_t^{min} := \arg\min_{\theta\in\mathcal{U}_t} \mathcal{J}^\star_\theta$, *and assume that* $\{l^\star_{\theta_t^{min},t}\}$ *is admissible. Under conditions i) and ii), the BOB-procedure defined by points 1 through 3 in the BOB-principle achieves the ideal objective, i.e.*

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} c(u_t, y_t) = \mathcal{J}^\star_{\theta^\circ} \quad a.s., \qquad \forall \theta^\circ \in \Theta.$$

Proof: Condition ii) implies that $\theta^\circ \in \mathcal{U}_t$ $a.s., \forall t \geq \bar{t}$, where $\bar{t}$ is a suitable instant point. From this, $\inf_{u\in\mathcal{U}_t} \mathcal{J}^\star_\theta \leq \mathcal{J}^\star_{\theta^\circ}$ $a.s., \forall t \geq \bar{t}$. Since, according to the BOB-procedure, at each instant point t we select in $\mathcal{U}_t$ the parameter θ_t^{min} with lower optimal cost $\mathcal{J}^\star_{\theta_t^{min}}$, we have $\limsup_{N\to\infty} 1/N \sum_{t=1}^{N} \mathcal{J}^\star_{\theta_t^{min}} \leq \limsup_{N\to\infty} 1/N \sum_{t=1}^{N} \mathcal{J}^\star_{\theta^\circ} = \mathcal{J}^\star_{\theta^\circ}$. Finally, recalling that $\{\theta_t^{min}\}$ is feasible, in view of condition i) we obtain

$$\begin{aligned}\limsup_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}c(u_t,y_t) &\leq \limsup_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\mathcal{J}^{\star}_{\theta_t^{min}}\\ &\leq \mathcal{J}^{\star}_{\theta^{\circ}}.\end{aligned}$$

□

Theorem 1 formalizes the conditions under which the BOB-principle leads to optimality. As it has been argued in Section 3, the observations falsify a given system when the incurred cost obtained by applying the optimal control law for it is larger than the optimal cost for this system. From this we see that condition i) is a natural condition which should be generally satisfied. Similarly, condition ii) is very natural since it only requires that the falsification procedure is not overselective. In conclusion, Theorem 1 delivers a very natural formalization of the BOB-principle introduced in the previous section. Clearly, different formalization are also possible.

Once the general BOB-philosophy has been translated into a precise mathematical statement such as the one in Theorem 1, the problem becomes merely technical in the need of satisfying the corresponding mathematical conditions. Obviously, this matter can only be discussed with regard to the particular application at hand. In the next section, a very simple example of application of the BOB principle is given, namely the study of a scalar LQG control scheme.

5 Adaptive LQG control - scalar case

5.1 Problem position

In this section the BOB-principle is applied to an adaptive LQG control problem under the assumption that the system state is scalar and noiselessly accessible. All the proofs are omitted.

Set $\mathcal{U}=\mathbf{R}$, $\mathcal{X}=\mathbf{R}$ and assume that the state evolution is governed by the equation

$$x_{t+1}=a^{\circ}x_t+b^{\circ}u_t+w_{t+1}, \qquad (1)$$

where $\{w_t\}$ is a noise process described as an i.i.d. Gaussian sequence with zero mean and unitary variance. The true parameter $\theta^{\circ}=[\ a^{\circ}\ \ b^{\circ}\ ]$ is unknown and belongs to a known compact set $\Theta\subset\mathbf{R}^2$ such that $b\neq 0, \forall[\ a\ \ b\ ]\in\Theta$ (controllability condition). The system state is observed without noise, i.e. $y_t=x_t$. Finally, the long-term cost criterion is given by

$$\limsup_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}[qx_t^2+u_t^2], \qquad q>0. \qquad (2)$$

In the case in which the true parameter θ° is known, it is a standard matter to compute the optimal control law that minimizes criterion (2). For future use, we name $\mathcal{J}^{\star}_{(a^{\circ},b^{\circ})}$ the corresponding optimal cost.

In the adaptive case where θ° is not known, we set the following

Adaptive control problem

Find a control law $\{l_t\}$ such that, with the position $u_t = l_t(o_t)$, we achieve the ideal objective, i.e. $\limsup_{N\to\infty}1/N\sum_{t=1}^{N}[qx_t^2+u_t^2]=\mathcal{J}^{\star}_{(a^{\circ},b^{\circ})}$ a.s., $\forall[\ a^{\circ}\ \ b^{\circ}\ ]\in\Theta$. □

5.2 Solving the adaptive control problem via the BOB-principle

To attack the adaptive control problem with the BOB-principle we need to find a suitable falsification criterion. The resulting unfalsified sets should satisfy conditions i) and ii) in Theorem 1.

A broad hint on how to select the unfalsified sets so as to satisfy condition ii) in Theorem 1 is provided by Lemma 1 below.

Name $[\ a_t^{LS}\ \ b_t^{LS}\ ]$ the least squares estimate of $[\ a^{\circ}\ \ b^{\circ}\ ]$:

$$[\ a_t^{LS}\ \ b_t^{LS}\ ]:=\arg\min_{[a\ \ b]\in\mathbf{R}^2}\sum_{k=1}^{t-1}(x_{k+1}-ax_k-bu_k)^2,$$

and define $\phi_k:=[\ x_k\ \ u_k\ ]$, and $V_t:=\sum_{k=1}^{t-1}\phi_k^T\phi_k$.

Lemma 1 *Fix any control law $\{l_t\}$ and choose a function μ_t such that $\log\sum_{k=1}^{t-1}x_k^2=o(\mu_t)$. Define the unfalsified set sequence through equation*

$$\begin{aligned}\mathcal{U}_t:=\{[\ a\ \ b\ ]\in\Theta:\ &([\ a\ \ b\ ]-[\ a_t^{LS}\ \ b_t^{LS}\ ])V_t\\ &([\ a\ \ b\ ]-[\ a_t^{LS}\ \ b_t^{LS}\ ])^T\leq\mu_t\}.\end{aligned} \qquad (3)$$

Then,

$$[\ a^{\circ}\ \ b^{\circ}\ ]\in\cup_t\cap_{k>t}\mathcal{U}_k\quad a.s.$$

□

Lemma 1 delivers a lower bound for μ_t, the fulfillment of which implies that condition ii) is satisfied. The next lemma gives an upper bound for μ_t such that condition i) in Theorem 1 is also satisfied.

Lemma 2 *Choose a function* μ_t *such that* $\mu_t = o(\log^2 \sum_{k=1}^{t-1} x_k^2)$ *and set* $u_t = K(a_t, b_t)x_t$*, where* $[\ a_t \quad b_t\ ]$ *belongs almost surely to set* $\mathcal{U}_t$ *defined through equation (3) and* $K(a_t, b_t)$ *denotes the optimal gain associated with parameter* $[\ a_t \quad b_t\ ]$*. Then*

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} [qx_t^2 + u_t^2] \leq \limsup_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \mathcal{J}^\star_{(a_t,b_t)} \quad a.s.$$

□

Basically, condition $\mu_t = o(\log^2 \sum_{k=1}^{t-1} x_k^2)$ requires that region $\mathcal{U}_t$ is not too spread around the least squares estimate.

By selecting the unfalsified set at time t as $\mathcal{U}_t$ in definition (3) with the bounds on μ_t as suggested by Lemma 1 and 2, the BOB-principle immediately leads to the following

Adaptive control method

At time t, do the following:

1. determine $\mathcal{U}_t$ as in definition (3) with $\mu_t = \log^r \sum_{k=1}^{t-1} x_k^2, \quad r \in (1, 2)$;
2. compute $[\ a_t^{min} \quad b_t^{min}]$ as the minimizer of $\mathcal{J}^\star_{(a,b)}$ in $\mathcal{U}_t$:

$$[\ a_t^{min} \quad b_t^{min}] := \arg\min_{[a \ \ b] \in \mathcal{U}_t} \mathcal{J}^\star_{(a,b)};$$

3. compute u_t by applying the optimal control law for $[\ a_t^{min} \quad b_t^{min}]$:

$$u_t = K(a_t^{min}, b_t^{min})x_t.$$

□

In view of Lemma 1 and 2, the effectiveness of this adaptive control method is a consequence of Theorem 1. This leads to the following

Theorem 2 *With the control law chosen according to the adaptive control method, we achieve the ideal objective, i.e.* $\limsup_{N\to\infty} 1/N \sum_{t=1}^{N} [qx_t^2 + u_t^2] = \mathcal{J}^\star_{(a^\circ, b^\circ)}$ *a.s.,* $\forall [\ a^\circ \quad b^\circ\ ] \in \Theta$.

□

6 Conclusions

In this paper, we have presented a broad solution method to achieve optimality in adaptive control problems. It is our belief that other researchers will be able to solve many specific problems by resorting to this general approach.

References

[1] Y. Bar-Shalom and E. Tse. Dual effort, certainty equivalence and separation in stochastic control. *IEEE Trans. on Automatic Control*, 19:494–500, 1974.

[2] Y. Bar-Shalom and K.D. Wall. Dual adaptive control and uncertainty effects in macroeconomic modelling. *Automatica*, 16:147–156, 1974.

[3] G.C. Goodwin, P.J. Ramadge, and P.E. Caines. Discrete-time stochastic adaptive control. *SIAM J. Control and Optim.*, 19:829–853, 1981.

[4] K.S. Sin and G.C. Goodwin. Stochastic adaptive control using a modified least squares algorithm. *Automatica*, 18:815–321, 1982.

[5] L. Guo and H.F. Chen. The astrom-wittenmark self-tuning regulator revisited and els-based adaptive trackers. *IEEE Trans. on Automatic Control*, AC-36:802–812, 1991.

[6] K. Aström and B. Wittenmark. On self-tuning regulators. *Automatica*, 9:185–199, 1973.

[7] A. Becker, P.R. Kumar, and C.Z. Wei. Adaptive control with the stochastic approximation algorithm: Geometry and convergence. *IEEE Trans. on Automatic Control*, AC-30:330–338, 1985.

[8] M.C. Campi and P.R. Kumar. Optimal adaptive control of an lqg system. In *Proc. 35th Conf. on Decision and Control*, page to appear, December 1996.

[9] M.C. Campi. The problem of pole-zero cancellation in transfer function identification and application to adaptive stabilization. *Automatica*, 32:849–857, 1996.

[10] P.R. Kumar W. Lin and T.I. Seidman. Will the self-tuning approach work for general cost criteria? *Syst. Contr. Lett.*, 6:77–85, 1985.

[11] J.W. Polderman. On the necessity of identifying the true parameter in adaptive lq control. *Syst. Contr. Lett.*, 8:87–91, 1986.

[12] J.W. Polderman. A note on the structure of two subsets of the parameter space in adaptive control problems. *Syst. Contr. Lett.*, 8:25–34, 1986.

[13] J.H. van Schuppen. Tuning of gaussian stochastic control systems. *IEEE Trans. on Automatic Control*, AC-39:2178–2190, 1994.