A Self-Optimizing Adaptive LQG Control Scheme for Input-Output Systems[1]

M. Prandini and M.C. Campi
Dipartimento di Elettronica per l'Automazione
Università degli Studi di Brescia – Via Branze 38, 25123 Brescia, Italy
E-mail: {prandini,campi}@ing.unibs.it

Abstract

In this paper, we consider the optimal control problem of an unknown linear system in input-output form based on the linear quadratic Gaussian (LQG) control design method. A self-tuning LQG control scheme is proposed which is shown to be stable and self-optimizing. Optimality is achieved by using a new identification algorithm which incorporates a cost-biasing term favoring the parameters with smaller LQG optimal cost and a second term aiming at moderating the time-variability of the estimate.

1 Introduction

It is well known that, in general, a self-tuning control system is not guaranteed to obtain the same performance as the one achievable under complete knowledge of the true plant (*self-optimization*). In particular, self-optimization result does not hold true for general control laws based on the minimization of multistep performance indexes (see *e.g.* [1, 2, 3]). In absence of suitable excitation conditions, the interplay between identification and control in a certainty equivalence adaptive control scheme may in fact result in the convergence of the parameter estimate to a parameterization different from the true one (see *e.g.* [1, 4]). When a cost criterion other than the output variance is considered, this identifiability problem results in a strictly suboptimal performance. In particular, the identifiability problem is significant in infinite-horizon LQG control and, in fact, in [3] it is proven that for a state space system subject to Gaussian noise the set of parameterizations leading to optimality of LQG control is strictly contained in the set of the potential convergence points.

Two main approaches have been proposed in the literature to deal with the self-optimality issue.
A first approach consists in achieving optimality as a side result of parameter consistency. This is typically obtained by introducing an appropriately vanishing dither noise in the control system, which is sufficiently exciting so that consistency is obtained, and - at the same time - mild enough in order not to degrade the control system performance, [5, 6, 7]. This approach is then useful only in the case when noise injection is feasible.
A second approach - adopted in this paper - is based on the so-called *cost-biased method* originally introduced in [8] where controlled Markov chains with a finite parameter set are considered. The basic idea is adding a suitable cost-biasing term to the least squares cost function that favors parameters with lower optimal cost, while preserving the closed-loop identification properties of the least squares algorithm. The results of this paper have been extended to Markov chains with an infinite parameter set in [9] and to systems with a general state space but still with a finite parameter set in [10].
Linear systems in a state space representation are dealt with in [1] and [11]. In these papers, the restrictive assumption that the state is fully accessible is made. Moreover, it is assumed that the noise system affects all state variables. This assumption is crucial for the correct functioning of the proposed identification procedure. As a matter of fact, the presence of a full-range noise sheds light on the existing difference between the true system and the estimated model and this helps the identification task. In [11], it is in fact shown that this mechanism is effective enough so as to counteract the effect of the cost-biasing term, thus guaranteeing the closed-loop identification property. Unfortunately, the assumption that the noise is full-range is so restrictive that it is not verified in many situations of interest, such as in the case of state space realizations of input-output systems.

In the present paper, an optimal adaptive control scheme still based on the cost-biasing idea, but for input-output systems, is presented. Extending the cost-biased approach to input-output systems is not trivial. On the other hand, it is important in that they are largely used in adaptive control applications. Moreover, assuming only the input and output measurability is much more realistic than assuming full state accessibility. As a side remark we also note that, in contrast with [1] and [11], our approach does not require the noise to be Gaussian.

[1] Research supported by MURST under the projects on "Identification and Control of Industrial Systems" and "Synthesis of Adaptive and Robust Controllers".

The paper is organized as follows: in Section 2, we introduce the dynamic systems we consider in the following and describe the cost-biased adaptive LQG control scheme. Some relevant properties of the standard least squares (LS) estimates are recalled in the same section, whereas the study of the cost-biased identification algorithm is presented in Section 3. Section 4 is devoted to the analysis of the closed-loop stability and the characterization of the self-tuning LQG control performance. Finally, Section 5 presents conclusions and suggestions for future research.

2 The cost-biased adaptive LQG control law

We consider dynamic systems in input-output form described by the following equation:

$$\mathcal{A}(\vartheta^\circ; q^{-1})\, y_t = \mathcal{B}(\vartheta^\circ; q^{-1})\, u_{t-1} + n_t, \qquad (1)$$

where $\mathcal{A}(\vartheta^\circ; q^{-1}) = 1 - \sum_{i=1}^{n} a_i^\circ q^{-i}$ and $\mathcal{B}(\vartheta^\circ; q^{-1}) = \sum_{i=1}^{m} b_i^\circ q^{-i+1}$ are polynomials in the unit-delay operator q^{-1}, and $\vartheta^\circ = [\, a_1^\circ\ a_2^\circ \ldots a_n^\circ\ b_1^\circ\ b_2^\circ \ldots b_m^\circ\,]^T$ is the *unknown* system parameter vector. The control objective is to determine the control law that minimizes the quadratic cost

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} [\, y_t^2 + \beta\, u_t^2],$$

where the control weight β is strictly positive.

2.1 The LQG optimal control problem

In this section, we summarize some facts on infinite-horizon LQG control for *known* systems which are relevant for the subsequent developments. This is also useful in order to introduce the assumptions and the notations we shall use throughout the paper.

Signal n_t in equation (1) is a stochastic disturbance precisely described in the following

Assumption 1 *$\{n_t\}$ is a martingale difference sequence with respect to a filtration $\{\mathcal{F}_t\}$, satisfying the following conditions*

1. $\sup_t E[|n_t|^p / \mathcal{F}_{t-1}] < \infty$, *almost surely (a.s) for all* $p > 0$*;*
2. $\displaystyle\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} n_t^2 = \sigma^2 > 0$, *a.s.*

Note that Assumption 1 is satisfied for example when $\{n_t\}$ is an i.i.d. Gaussian sequence, but it includes many other situations.
We make the assumption on system (1) that $n > 0$ (non trivial autoregressive part). Note that if $n = 0$ the trivial control law $u_t = 0$, $t \geq 0$, is obviously optimal irrespective of the value of ϑ°.

We further assume that system (1) belongs to a known set of stabilizable models according to

Assumption 2 *$\vartheta^\circ \in \Theta$, where Θ is a compact set such that $\Theta \subset \mathcal{C} = \{\vartheta \in \Re^{n+m} : q^s \mathcal{A}(\vartheta; q^{-1})$ and $q^{s-1}\mathcal{B}(\vartheta; q^{-1})$ do not present unstable pole-zero cancellations$\}$, $s = \max\{n, m\}$ being the order of the system.*

System (1) is initialized with $y_t = u_{t-1} = 0$, $t \leq 0$.

For the determination of an optimal control law, it is convenient to represent system (1) in a state space form such that the state is accessible, and then apply the well-known solution to the optimal LQG control problem for full state accessible state space systems (see *e.g.* [5], [12]).
Defining $x_t := [y_t \ldots y_{t-(n-1)}\ u_{t-1} \ldots u_{t-(m-1)}]^T$, system (1) can be given the following state space representation of order $\bar{s} := n + m - 1$

$$\begin{cases} x_{t+1} = A_{\vartheta^\circ} x_t + B_{\vartheta^\circ} u_t + C n_{t+1} \\ y_t = H x_t, \end{cases} \qquad (2)$$

initialized with $x_0 = [\, 0 \ldots 0\,]^T$, with matrices

$$A_\vartheta = \left[\begin{array}{cccc|cccc} a_1 & \ldots & a_{n-1} & a_n & b_2 & \ldots & b_{m-1} & b_m \\ 1 & 0 & \ldots & & 0 & \ldots & & 0 \\ & \ddots & \ddots & & & \ddots & & 0 \\ & & 1 & 0 & & & & 0 \\ \hline 0 & \ldots & \ldots & 0 & 0 & \ldots & \ldots & 0 \\ 0 & \ldots & \ldots & 0 & 1 & 0 & & \\ & \ddots & \ddots & & & \ddots & \ddots & \\ & & 0 & 0 & & & 1 & 0 \end{array}\right],$$

$$B_\vartheta = [\; b_1 \quad 0 \quad \ldots \quad 0 \;|\; 1 \quad 0 \quad \ldots \quad 0 \;]^T,$$

$$C = H^T = [\; 1 \quad 0 \quad \ldots \quad 0 \;|\; 0 \quad 0 \quad \ldots \quad 0 \;]^T.$$

In this way, the LQG regulation problem for the system in input-output representation (1) is reformulated as a complete state information control problem where the performance index to be minimized is given by $\limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} [\, x_t^T T x_t + \beta u_t^2]$, with $T = H^T H \geq 0$ and $\beta > 0$.
Note that, in the case when $n > 1$ and $m > 1$, the state space representation (2) of system (1) is non minimal (the order of system (1) is $s = \max\{n, m\}$, whereas the dimension of matrix A_{ϑ° is $\bar{s} = n + m - 1$). However, from the block triangular matrix structure of A_{ϑ° it is easily seen that the added eigenvalues are identically equal to zero. Then from Assumption 2 it follows that $(A_{\vartheta^\circ}, B_{\vartheta^\circ})$ is stabilizable and (A_{ϑ°, H) is detectable, and hence the standard approach based on the solution to a Riccati equation can be used to determine the control law.

Specifically, the solution to the original LQG control problem has the following expression ([5])

$$u_t = \mathcal{S}(\vartheta^\circ; q^{-1})\, y_t + \mathcal{R}(\vartheta^\circ; q^{-1})\, u_t, \tag{3}$$

with $\mathcal{S}(\vartheta^\circ; q^{-1}) = \sum_{i=0}^{n-1} s_i(\vartheta^\circ) q^{-i}$ and $\mathcal{R}(\vartheta^\circ; q^{-1}) = \sum_{i=1}^{m-1} r_i(\vartheta^\circ) q^{-i}$, where the coefficients $\{s_i(\vartheta^\circ)\}$ and $\{r_i(\vartheta^\circ)\}$ are computed as follows.
Set $L_{\vartheta^\circ} := [\, s_0(\vartheta^\circ) \dots s_{n-1}(\vartheta^\circ) \;\; r_1(\vartheta^\circ) \dots r_{m-1}(\vartheta^\circ)\,]$. Then,

$$L_{\vartheta^\circ} = -(B_{\vartheta^\circ}^T P_{\vartheta^\circ} B_{\vartheta^\circ} + \beta)^{-1} B_{\vartheta^\circ}^T P_{\vartheta^\circ} A_{\vartheta^\circ},$$

where P_{ϑ° is the unique positive semidefinite solution to the discrete time algebraic Riccati equation

$$P = A_{\vartheta^\circ}^T \left[P - P B_{\vartheta^\circ} (B_{\vartheta^\circ}^T P B_{\vartheta^\circ} + \beta)^{-1} B_{\vartheta^\circ}^T P\right] A_{\vartheta^\circ} + T.$$

Moreover, the optimal LQG cost is given by $J^\star(\vartheta^\circ) = \sigma^2 trace(P_{\vartheta^\circ} C C^T)$, almost surely.

Remark 1 *Since the solution P_ϑ to*

$$P = A_\vartheta^T \left[P - P B_\vartheta (B_\vartheta^T P B_\vartheta + \beta)^{-1} B_\vartheta^T P\right] A_\vartheta + T$$

is analytic as a function of the parameter vector ϑ in the set $\mathcal{C}$ (see [13]), it is easily seen that $s_i(\vartheta)$, $r_i(\vartheta)$, and $J^\star(\vartheta)$ are analytic functions of ϑ, $\vartheta \in \mathcal{C}$, as well.

2.2 The cost-biased identification algorithm

Introducing the observation vector $\varphi_t := [\, y_t \dots y_{t-(n-1)}\, u_t \dots u_{t-(m-1)}]^T$, system (1) can be given the regression-like form

$$y_t = \varphi_{t-1}^T \vartheta^\circ + n_t,$$

and the LS index for estimating ϑ° is

$$V_t(\vartheta) = \sum_{s=1}^{t} (y_s - \varphi_{s-1}^T \vartheta)^2. \tag{4}$$

In the theorem below, we recall a fundamental result for the LS estimate $\hat{\vartheta}_t^{LS}$ proven in [14], Theorem 1.

Theorem 1 *Suppose that u_t is $\mathcal{F}_t$-measurable. Then,*

$$\begin{aligned} &(\vartheta^\circ - \hat{\vartheta}_t^{LS})^T \textstyle\sum_{s=1}^t \varphi_{s-1}\varphi_{s-1}^T (\vartheta^\circ - \hat{\vartheta}_t^{LS}) \\ &= O\big(\log \lambda_{max}(\textstyle\sum_{s=1}^t \varphi_{s-1}\varphi_{s-1}^T)\big) \quad a.s.. \end{aligned} \tag{5}$$

In particular, this implies that under the conditions $\lambda_{min}(\sum_{s=1}^t \varphi_{s-1}\varphi_{s-1}^T) \to \infty$ and $\log \lambda_{max}(\sum_{s=0}^t \varphi_{s-1}\varphi_{s-1}^T) = o(\lambda_{min}(\sum_{s=1}^t \varphi_{s-1}\varphi_{s-1}^T))$ a.s., the least squares estimate is a.s. consistent.

In adaptive control, identification is performed in closed-loop. Therefore, one cannot ensure the satisfaction of the conditions for consistency, and hence the true parameter vector is generally not consistently estimated. Nevertheless, property (5) still provides a valuable bound on the discrepancy between the estimated parameter and the true parameter. We call this property "closed-loop identification property" to emphasize that it holds even in closed-loop. On the other hand, the LS identification algorithm generally provides estimates with an optimal LQG cost larger than the optimal cost associated with the true system. This because if, as expected, the behavior of the adaptively controlled system is the same as the one of the estimated system, at least in the long run, then, the optimal LQG cost for the system with parameter $\hat{\vartheta}_t^{LS}$ is the same as the actual incurred cost, which obviously cannot be lower than the optimal cost for the true system.
Motivated by these considerations, we introduce a cost-biased identification algorithm with the twofold objective of preserving the LS property (5) and forcing the estimates to lie asymptotically in the parameter region with an optimal cost not larger than the optimal cost one of the true system.

Consider the estimate $\hat{\vartheta}_t$ computed through the following algorithm:

$$\hat{\vartheta}_t = \begin{cases} \arg\min\limits_{\vartheta \in \Theta} D_t(\vartheta), & \text{if } t = t_i,\ i = 0, 1, \dots \\ \hat{\vartheta}_{t-1}, & \text{otherwise,} \end{cases} \tag{6}$$

where the time instants $\{t_i\}$ are obtained by the recursive equation $t_{i+1} = t_i + T_i$ initialized with $t_0 = 0$, and the cost-biased identification index $D_t(\vartheta)$ is given by

$$D_t(\vartheta) = V_t(\vartheta) + \alpha_t J^\star(\vartheta) + \gamma_t \|\vartheta - \hat{\vartheta}_{t-1}\|,\ \hat{\vartheta}_{-1} = 0,$$

where $V_t(\vartheta)$ is the LS cost (4) and $J^\star(\vartheta)$ is the optimal LQG cost for system (1) with parameter ϑ.

The identification algorithm is completely defined by specifying the sequences of freezing time intervals $\{T_i\}$, cost-biasing weights $\{\alpha_t\}$, friction parameters $\{\gamma_t\}$. We discuss hereafter the meaning of these parameters, while their actual choice is relegated to the following section.

The freezing parameter T_i is used to ensure stability of the closed-loop system. Since the parameter estimate changes with time and the control law is tuned to such an estimate, the adaptive control system is time-varying. On the other hand, it is well known that guaranteeing a stability property at each time instant for the "frozen dynamics" does not imply that the overall time-varying system has a stable dynamics. This problem can be solved by updating the estimate at a slower rate than the updating of the system variables, and this is achieved by a suitable choice of T_i. This same approach is for instance exploited in [15] and [16].
The cost-biasing term $\alpha_t J^\star(\vartheta)$ is introduced with the objective of penalizing those parameterizations with high optimal LQG cost. The weight α_t has to be appropriately selected so as to balance the contrasting

objectives of preserving the closed-loop identification property (5) and forcing the asymptotic estimate to correspond to a model with value of the optimal LQG performance index not larger than the optimal performance value for the true system.
Finally, the friction term $\gamma_t\|\vartheta - \hat{\vartheta}_{t-1}\|$ is introduced so as to avoid that the estimate $\hat{\vartheta}_t$ is subject to undesired jumps in the time instants t_i when it is updated. This is necessary to prove optimality of the adaptive scheme.

2.3 Cost-biased adaptive LQG control

According to the certainty equivalence principle, the adaptive control law is given by the optimal control law (3) with the estimate $\hat{\vartheta}_t$ in place of ϑ°

$$u_t = \mathcal{S}(\hat{\vartheta}_t; q^{-1})\, y_t + \mathcal{R}(\hat{\vartheta}_t; q^{-1})\, u_t.$$

The system:

$$\begin{cases} y_{t+1} = [1 - \mathcal{A}(\hat{\vartheta}_t; q^{-1})]\, y_{t+1} + \mathcal{B}(\hat{\vartheta}_t; q^{-1})\, u_t \\ u_t = \mathcal{S}(\hat{\vartheta}_t; q^{-1})\, y_t + \mathcal{R}(\hat{\vartheta}_t; q^{-1})\, u_t \end{cases} \tag{7}$$

is then given the name of *autonomous estimated system*. We will select T_i so as to stabilize system (7) and later on (Section 4) we will see that this leads to the stability of the true closed-loop system. Letting $x_t := [y_t \dots y_{t-(n-1)} u_{t-1} \dots u_{t-(m-1)}]^T$, this system can be given the state space representation $x_{t+1} = F_{\hat{\vartheta}_t}\, x_t$, with

$$F_\vartheta = A_\vartheta + B_\vartheta L_\vartheta,$$

where the matrices A_ϑ, B_ϑ and L_ϑ have been introduced in Section 2.1.
Choose now a constant $\mu < 1$ (*contraction* constant). The time interval T_i is then defined as

$$T_i := \inf\{\tau \in Z_+ : \|(F_{\hat{\vartheta}_{t_i}})^\tau\| \le \mu\} \tag{8}$$

(note that such a T_i exists since $\hat{\vartheta}_{t_i}$ belongs to Θ and therefore corresponds to a stabilizable system). In this way, the time-varying system (7) is kept constant until its dynamics is contracted by a factor μ, whence guaranteeing its stability. The following proposition makes this precise (the proof is omitted due to space limitations. The interested reader is referred to [17].).

Proposition 1 *The sequence of freezing time intervals $\{T_i\}$ given in (8) is bounded, i.e., $\sup_{i\ge 0} T_i < \infty$. Moreover, the autonomous estimated system $x_{t+1} = F_{\hat{\vartheta}_t}\, x_t$ is a.s. exponentially stable, uniformly in time.*

3 Properties of the cost-biased estimate $\hat{\vartheta}_t$

We now show that by suitably choosing $\{\alpha_t\}$ and $\{\gamma_t\}$ we can obtained the desired properties for $\hat{\vartheta}_t$, while preserving the closed-loop identification property of $\hat{\vartheta}_t^{LS}$.

Theorem 2 *Suppose that u_t is $\mathcal{F}_t$-measurable. Given $\delta > 0$, select*

$$\alpha_t := \log^{1+\delta} \lambda_{max}(\sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T) \tag{9}$$

and $\{\gamma_t\}$ to be a positive diverging sequence of real numbers satisfying $\gamma_t = o(\alpha_t)$. Then,

i) $(\vartheta^\circ - \hat{\vartheta}_{t_i})^T \sum_{s=1}^{t_i} \varphi_{s-1}\varphi_{s-1}^T (\vartheta^\circ - \hat{\vartheta}_{t_i}) = O(\alpha_{t_i})$ *a.s.,*

ii) $\limsup_{t\to\infty} J^\star(\hat{\vartheta}_t) \le J^\star_{\vartheta^\circ}$, *a.s.,*

iii) if $\sum_{t=1}^{N} \|\varphi_{t-1}\|^2 = O(N)$ *a.s., then* $\sum_{t=1}^{N} \|\hat{\vartheta}_t - \hat{\vartheta}_{t-1}\| = o(N)$ *a.s.*

Proof: Point *i)*: $D_t(\vartheta) - V_t(\hat{\vartheta}_t^{LS})$ can be written as

$$\begin{aligned} D_t(\vartheta) - V_t(\hat{\vartheta}_t^{LS}) &= \vartheta^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T \vartheta - 2\vartheta^T \sum_{s=1}^{t} \varphi_{s-1} y_s \\ &+ \alpha_t J^\star(\vartheta) + \gamma_t\|\vartheta - \hat{\vartheta}_{t-1}\| - (\hat{\vartheta}_t^{LS})^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T \hat{\vartheta}_t^{LS} \\ &+ 2(\hat{\vartheta}_t^{LS})^T \sum_{s=1}^{t} \varphi_{s-1} y_s. \end{aligned} \tag{10}$$

The LS estimate $\hat{\vartheta}_t^{LS}$ minimizing $V_t(\vartheta)$ satisfies $\sum_{s=1}^{t} \varphi_{s-1} y_s = \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T \hat{\vartheta}_t^{LS}$. Substituting this last expression in equation (10), we obtain

$$\begin{aligned} D_t(\vartheta) - V_t(\hat{\vartheta}_t^{LS}) =& (\vartheta - \hat{\vartheta}_t^{LS})^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T (\vartheta - \hat{\vartheta}_t^{LS}) \\ &+ \alpha_t J^\star(\vartheta) + \gamma_t\|\vartheta - \hat{\vartheta}_{t-1}\|. \end{aligned} \tag{11}$$

Set $\vartheta_t := \arg\min_{\vartheta\in\Theta} D_t(\vartheta)$. By definition of ϑ_t we have $D_t(\vartheta_t) - V_t(\hat{\vartheta}_t^{LS}) \le D_t(\vartheta) - V_t(\hat{\vartheta}_t^{LS})$, $\vartheta \in \Theta$. By choosing $\vartheta = \vartheta^\circ$ and using expression (11), we then get

$$\begin{aligned} &(\vartheta_t - \hat{\vartheta}_t^{LS})^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T (\vartheta_t - \hat{\vartheta}_t^{LS}) + \alpha_t J^\star(\vartheta_t) \\ &+ \gamma_t\|\vartheta_t - \hat{\vartheta}_{t-1}\| \le (\vartheta^\circ - \hat{\vartheta}_t^{LS})^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T (\vartheta^\circ - \hat{\vartheta}_t^{LS}) \\ &+ \alpha_t J^\star(\vartheta^\circ) + \gamma_t\|\vartheta^\circ - \hat{\vartheta}_{t-1}\| = O(\alpha_t), \quad a.s., \end{aligned} \tag{12}$$

where the last equality follows from Theorem 1, the definition (9) of α_t, the fact that $\|\vartheta^\circ - \hat{\vartheta}_{t-1}\|$ is bounded and the relation $\gamma_t = o(\alpha_t)$. Since $\alpha_t J^\star(\vartheta_t) + \gamma_t\|\vartheta_t - \hat{\vartheta}_{t-1}\| \ge 0$, we have $(\vartheta_t - \hat{\vartheta}_t^{LS})^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T (\vartheta_t - \hat{\vartheta}_t^{LS}) = O(\alpha_t)$, a.s. From definition (9) of α_t and Theorem 1, we then have

$$(\vartheta_t - \vartheta^\circ)^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T(\vartheta_t - \vartheta^\circ)$$
$$\leq 2\Big[(\vartheta_t - \hat{\vartheta}_t^{LS})^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T(\vartheta_t - \hat{\vartheta}_t^{LS}) +$$
$$(\hat{\vartheta}_t^{LS} - \vartheta^\circ)^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T(\hat{\vartheta}_t^{LS} - \vartheta^\circ)\Big] = O(\alpha_t),$$

a.s., thus concluding the proof of point i), since $\hat{\vartheta}_t = \vartheta_t$, for $t = t_i$, $i = 0, 1, \ldots$.

Point *ii)*: A simple elaboration of (12) shows that

$$J^\star(\vartheta_t) \leq \frac{(\vartheta^\circ - \hat{\vartheta}_t^{LS})^T \sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T(\vartheta^\circ - \hat{\vartheta}_t^{LS})}{\alpha_t}$$
$$+ J^\star(\vartheta^\circ) + \frac{\gamma_t}{\alpha_t}\|\vartheta^\circ - \hat{\vartheta}_{t-1}\|$$
$$= \frac{O(\log \lambda_{max}(\sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T))}{\log^{1+\delta} \lambda_{max}(\sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T)} + \frac{o(\alpha_t)}{\alpha_t} + J^\star(\vartheta^\circ),$$

a.s., where in the second equation we have used the definition (9) of α_t and the fact that $\gamma_t = o(\alpha_t)$. To conclude the proof, it suffices to show that $\lim_{t\to\infty} \log \lambda_{max}(\sum_{s=1}^{t} \varphi_{s-1}\varphi_{s-1}^T) = \infty$. The easy proof of this fact is omitted.

Point *iii)*: By definition (6) of $\hat{\vartheta}_t$, then $V_t(\hat{\vartheta}_t) + \alpha_t J^\star(\hat{\vartheta}_t) + \gamma_t\|\hat{\vartheta}_t - \hat{\vartheta}_{t-1}\| \leq V_t(\hat{\vartheta}_{t-1}) + \alpha_t J^\star(\hat{\vartheta}_{t-1})$, which implies

$$\sum_{t=1}^{N} \gamma_t\|\hat{\vartheta}_t - \hat{\vartheta}_{t-1}\| \leq \sum_{t=1}^{N}[V_t(\hat{\vartheta}_{t-1}) - V_t(\hat{\vartheta}_t)]$$
$$+ \sum_{t=1}^{N} \alpha_t[J^\star(\hat{\vartheta}_{t-1}) - J^\star(\hat{\vartheta}_t)]. \quad (13)$$

The first term in the right-hand-side of equation (13) can be bounded as follows

$$\sum_{t=1}^{N}[V_t(\hat{\vartheta}_{t-1}) - V_t(\hat{\vartheta}_t)]$$
$$\leq V_1(\hat{\vartheta}_0) - V_N(\hat{\vartheta}_N) + \sum_{t=1}^{N-1}[V_{t+1}(\hat{\vartheta}_t) - V_t(\hat{\vartheta}_t)]$$
$$\leq V_1(\hat{\vartheta}_0) + \sum_{t=1}^{N-1}[\varphi_t^T(\vartheta^\circ - \hat{\vartheta}_t) + n_{t+1}]^2$$
$$\leq k_1[1 + \sum_{t=1}^{N}\|\varphi_{t-1}\|^2 + \sum_{t=1}^{N-1} n_{t+1}^2],$$

k_1 being a suitable constant, where we used the boundedness of $\hat{\vartheta}_t$.
By Remark 1, the second term in the right-hand-side of equation (13) can be bounded as follows

$$\sum_{t=1}^{N} \alpha_t[J^\star(\hat{\vartheta}_{t-1}) - J^\star(\hat{\vartheta}_t)]$$
$$= \alpha_1 J^\star(\hat{\vartheta}_0) - \alpha_N J^\star(\hat{\vartheta}_N) + \sum_{t=1}^{N-1}(\alpha_{t+1} - \alpha_t)J^\star(\hat{\vartheta}_t)$$
$$\leq \alpha_1 J^\star(\hat{\vartheta}_0) + \max_{\vartheta\in\Theta} J^\star(\vartheta) \sum_{t=1}^{N-1}(\alpha_{t+1} - \alpha_t)$$
$$= k_2[1 + \alpha_N],$$

where k_2 is a suitable constant.
Substituting these bounds in equation (13), we get

$$\frac{1}{N}\sum_{t=1}^{N} \gamma_t\|\hat{\vartheta}_t - \hat{\vartheta}_{t-1}\| \leq \bar{k}\Big[\frac{1}{N} + \frac{\alpha_N}{N} + \frac{1}{N}\sum_{t=1}^{N}\|\varphi_{t-1}\|^2$$
$$+ \frac{1}{N}\sum_{t=1}^{N-1} n_{t+1}^2\Big]. \quad (14)$$

with $\bar{k}$ = suitable constant. Observe now that all the terms in the right-hand-side of equation (14) are $O(1)$. This in particular follows from $\sum_{t=1}^{N}\|\varphi_{t-1}\|^2 = O(N)$ and Assumption 1, point 2. Then, $\frac{1}{N}\sum_{t=1}^{N}\gamma_t\|\hat{\vartheta}_t - \hat{\vartheta}_{t-1}\| = O(1)$. Since γ_t tends to infinity, this last equation implies $\frac{1}{N}\sum_{t=1}^{N}\|\hat{\vartheta}_t - \hat{\vartheta}_{t-1}\| = o(1)$, that is the thesis. ■

By definition (9), $\{\alpha_t\}$ is chosen to be an increasing sequence of real numbers adaptively selected on the basis of the data generated by the controlled system. According to result *ii)*, this selection is effective in pushing the estimate towards the region where the optimal LQG cost is not larger than $J^\star(\vartheta^\circ)$. In turn, result *i)* shows that the closed-loop identification property (5) is preserved with two slight differences: 1) the exponent $1+\delta$ appears in the right-hand-side, 2) the rate of divergence in point *i)* of Theorem 2 only concerns with the time instants t_i when the estimate $\hat{\vartheta}_t$ is updated, while the original closed-loop identification property refers to all t's. Nevertheless by a suitable manipulation of the sole result *i)* in Theorem 2, we are able to prove that the estimation error $e_t := \varphi_t^T[\vartheta^\circ - \hat{\vartheta}_t]$ is small compared to the signal involved in the closed-loop suystem. The technical proof of this result is omitted due to space limitations. The interested reader is referred to [17] for a proof of this result.

Proposition 2 *The estimation error* $e_t = \varphi_t^T[\vartheta^\circ - \hat{\vartheta}_t]$ *satisfies the following equation*

$$\sum_{t=0,\ t\notin\mathcal{B}_N}^{N} |e_t|^p = o(\sum_{t=0}^{N}\|\varphi_t\|^p + N), \quad p \geq 2,\ a.s.,$$

where $\mathcal{B}_N$ *is a set of instant points which depends on* N*, whose cardinality is bounded:* $|\mathcal{B}_N| \leq C_{\mathcal{B}}, \forall N$.

4 Stability and optimality

In this section we state the stability and optimality results for the cost-biased adaptive LQG control scheme introduced in Section 2. The proof of the theorem is omitted due to space limitations (see [17]).

The closed-loop system

$$\begin{cases} y_{t+1} = [1 - \mathcal{A}(\vartheta^\circ; q^{-1})]\, y_{t+1} + \mathcal{B}(\vartheta^\circ; q^{-1})\, u_t + n_{t+1} \\ u_t = \mathcal{S}(\hat{\vartheta}_t; q^{-1})\, y_t + \mathcal{R}(\hat{\vartheta}_t; q^{-1})\, u_t \end{cases} \tag{15}$$

can be represented as a variation system with respect to the estimated system of equation (7) as follows

$$\begin{cases} y_{t+1} = [1 - \mathcal{A}(\hat{\vartheta}_t; q^{-1})]\, y_{t+1} + \mathcal{B}(\hat{\vartheta}_t; q^{-1})\, u_t + n_{t+1} + e_t \\ u_t = \mathcal{S}(\hat{\vartheta}_t; q^{-1})\, y_t + \mathcal{R}(\hat{\vartheta}_t; q^{-1})\, u_t. \end{cases}$$

The uniform stability property of the estimated system (7) (Proposition 1) and the property of e_t stated in Proposition 2 are exploited in the next theorem to prove the stability of system (15). Properties *ii)* and *iii)* in Theorem 2 are needed to show that system (15) is self-optimizing.

Theorem 3 (stability & optimality) *The adaptive LQG control scheme (15) is L^p-stable:*

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} [|y_t|^p + |u_t|^p] < \infty, \ \textit{a.s., for all } p > 0.$$

Moreover, it is self-optimizing:

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} \left[y_t^2 + \beta\, u_t^2\right] = J^\star(\vartheta^\circ), \ \textit{a.s.}$$

5 Conclusions

We introduced a new LQG adaptive control scheme based on the certainty equivalence principle able to ensure both stability and optimality results irrespectively of the excitation characteristics of the involved signals. This is obtained by adopting a cost-biased approach, which is effective in overcoming the identifiability problems usually arising in certainty equivalence adaptive control and causing sub-optimality.

Interesting research issues has still to be addressed. First, in this paper we deal with the case when the noise affecting the system is white. This hypothesis is necessary mainly for the applicability of the proposed cost-biased least squares identification method, whose properties are in fact derived on the basis of the least squares estimate properties. As a consequence of this fact, the extension to the ARMAX system case is not trivial. An encouraging starting point is represented by the fact that the extended least squares algorithm satisfies closed-loop properties similar to those valid for the least squares algorithm (see *e.g.* [5]).
Moreover, the proposed identification method is non-recursive. The cost-biased identification index has, in general, multiple local minima and its minimization is not straightforward. Therefore, it should be minimized by resorting to some global optimization algorithm. This limitation must be removed by introducing some recursive way to minimize our performance index so as to retain all the properties relevant to control.

References

[1] P. R. Kumar, "Optimal adaptive control of linear-quadratic-gaussian systems", *SIAM J. Control and Optim.*, vol. 21, pp. 163–178, 1983.

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

[3] J. H. van Schuppen, "Tuning of Gaussian stochastic control systems", *IEEE Trans. on Automatic Control*, vol. AC-39, pp. 2178–2190, 1994.

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

[5] H. F. Chen and L. Guo, *Identification and Stochastic Adaptive Control*, Birkhäuser, Boston, 1991.

[6] L. Guo, "Self-convergence of weighted least-squares with applications to stochastic adaptive control", *IEEE Trans. on Automatic Control*, vol. 41, pp. 79–89, 1996.

[7] M. Prandini, *Adaptive LQG control: optimality analysis and robust controller design*, Ph.D. Thesis, Feb. 1998.

[8] P. R. Kumar and A. Becker, "A new family of optimal adaptive controllers for Markov chains", *IEEE Trans. on Automatic Control*, vol. AC-27, pp. 137–146, 1982.

[9] W. Lin and P.K. Kumar, *Stochastic control of a queue with two servers of different rates*, chapter 44, Analysis and Optimization of Systems (A. Bensoussan and J.L. Lions, eds.) Springer-Verlag. Lecture Notes in Control and Information Sciences., 1982.

[10] P. R. Kumar, "Simultaneous identification and adaptive control of unknown systems over finite parameter sets", *IEEE Trans. on Automatic Control*, vol. AC-28, pp. 68–76, 1983.

[11] M. C. Campi and P. R. Kumar, "Adaptive linear quadratic Gaussian control: the cost-biased approach revisited", *SIAM J. Control and Optim.*, vol. 36, pp. 1890–1907, 1998.

[12] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. I and II, Athena Scientific, Belmont, Massachusetts, 1995.

[13] D. F. Delchamps, "Analytic feedback control and the algebraic Riccati equation", *IEEE Trans. on Automatic Control*, vol. AC-29, pp. 1031–1033, 1984.

[14] T. L. Lai and C. Z. Wei, "Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems", *Ann. Statist.*, vol. 10, pp. 154–166, 1982.

[15] R. Ortega, R. Kelly, and R. Lozano-Leal, "On global stability of adaptive systems using an estimator with parameter freezing", *IEEE Trans. on Automatic Control*, vol. AC-34, pp. 343–346, 1989.

[16] M. Prandini, S. Bittanti, and M. C. Campi, "A penalized identification criterion for securing controllability in adaptive control", *J. of Math. Syst., Estim. and Control*, vol. 8, pp. 491–494, 1998, Retrival code for the electronic version: 29460.

[17] M. Prandini and M. C. Campi, "Adaptive LQG control of input-output systems - A cost-biased approach", *SIAM J. Control and Optim.*, to appear.