Journal of Mathematical Systems, Estimation, and Control
Vol. 7, No. 1, 1997, pp. 1–22

Information Inequalities and Limiting Possibilities of Adaptive Control Strategies in ARX models with a General Quadratic Criterion*

Alexander Poznyak Roberto Salas Zúñiga

Abstract

This paper deals with the analysis of the convergence rate of adaptive asymptotically optimal control strategies when applied to linear, stationary, multidimensional objects belonging to some class which might include a moving average control term. A general type of performance index, which is the sum of a quadratic form in the output signal plus a quadratic form in the input signal, is considered. Finally, it is shown that for any adaptive control scheme, the corresponding state space trajectories do not differ less than some lower bound, which is sharp, from those corresponding to an optimal control scheme (where full information on the parameters is available). Single input - single output (SISO) and two dimensional case (MIMO) examples are presented.

1 Introduction

Many different papers have been devoted to the synthesis and analysis of adaptive control strategies for the class of stationary linear objects, perturbed with stationary (in the wide sense) stochastic noise [1], [2], [3], [4], [7], [8]. It has been shown that different adaptive control algorithms have different convergence rates in some functional sense. In other words, they guarantee different adaptation rates. So two algorithms of adaptive control have been studied in [6].

The convergence of the outputs of a system controlled using an adaptive scheme to the outputs of the system whose parameters are the true ones, has

*Received February 28, 1994; received in final form March 1, 1995. Summary appeared in Volume 7, Number 1, 1997.

been considered. However, the convergence rate analysis of these outputs has not be included. Other publications [1]-[5] also did not touch this problem.

It is clear that any characteristic of the adaptation process is dependent on the particular adaptive strategy employed, on the external perturbations and on the class of objects (plants) considered. Hence, if we desire to suggest any characteristic, which is a general one for some groups of objects, we must consider a class of objects and formulate a problem for this class. So, in [8] an approach for finding the lower bound for the optimal convergence rate of adaptive strategies was suggested for the class of linear multidimensional stationary plants without moving average terms in the input, considering a simple performance functional which contains only the losses of the output signal.

This paper is devoted to the generalization of the approach given in [8], for the class of linear controllable systems with moving average terms in the input and with a general criterion that includes losses of the output and input signals. The paper is organized as follows:

after the notation, a description of the class of controllable objects as well as the statement of the problem on the computation of the lower bound of the converge rate are presented;

the second part contains the main results which consist on two statements concerning some information inequality for determining the limiting boundary value for the rate of adaptation of any adaptive control strategy belonging to some class;

the next section is devoted to the consideration of two partial cases which are important in practical applications: regulation and tracking problems;

the last section deals with three examples for computing information inequalities in some concrete SISO and MIMO ARX systems;

the conclusions contain some discussion concerning the use of nonlinear identificators for achieving of this bound.

2 Description of a Class of Adaptive Control Strategies and Statement of the Problem

Let us consider the following sequences of random variables:

$$\{y_n\}, \quad \{u_n\}, \quad \{\xi_n\}$$

which are defined on the probability space $(\Omega, \mathcal{F}, P)$ and are connected by the following recurrent equation

$$y_n = -\sum_{i=1}^{N_a} A_i y_{n-i} + \sum_{j=0}^{N_b} B_j u_{n-j} + \xi_n \tag{2.1}$$

where

$\{y_n\}$ $(n = 1, 2, \cdots)$ is a sequence of outputs of the system (2.1), $y_n \in R^m$, and the initial conditions $y_0, y_{-1}, \cdots, y_{1-N_a}$ are fixed vectors;

$\{u_n\}$ is a sequence of inputs of the system (2.1), $u_n \in R^K$, and the initial conditions are given by: $u_0 = u_{-1} = \cdots = u_{1-N_b} = 0$ (for the simplicity);

$\{\xi_n\}$ is a sequence of external disturbances of the system (2.1), $\xi_n \in R^m$

$A_i \in R^{m\times m}$ $(i = 1, \cdots, N_a)$, and $B_i \in R^{m\times K}$ $(i = 0, \cdots, N_b)$ are constant matrices describing the system (2.1).

Here we assume that the sequences $\{y_n\}$ and $\{u_n\}$ are *observable*; i.e. we can use these values for constructing control strategies.

If we denote

$$N := N_a + N_b$$

$$\theta := [A_1 \cdots A_{N_a}; B_1 \cdots B_{N_b}]$$

$$z_n := \left[y_{n-1}^T \cdots y_{n-Na}^T; -u_{n-1}^T \cdots - u_{n-Nb}^T\right]^T,$$

then we can rewrite the given ARX (auto-regression with exogenous inputs) model (2.1) in the *standard form*:

$$y_n = -\theta z_n + B_0 u_n + \xi_n. \tag{2.2}$$

Consider the following performance index (objective function) J which includes losses in the control and the output of the system (2.1):

$$J(\{u_n\}) := \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E\{y_t^T Q y_t + u_t^T R u_t\} \tag{2.3}$$

where the matrices Q and R are given.

Definition 1 *Let us say that any sequence of $\mathcal{F}$-measurable Borel functions*

$$u_n := u_n(u_1, y_1, \cdots, u_{n-1}, y_{n-1})$$

is **an $\mathbf{L}_p$-realizable strategy** *$(p > 1)$ if for any random trajectory $\{\xi_n\}$ such that*

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E\{\|\xi_t\|^p\} < \infty \tag{2.4}$$

the following inequality holds

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E\{\|u_t\|^p + \|y_t\|^p\} < \infty. \tag{2.5}$$

The next lemma gives a lower bound estimation for the performance index $J = J(\{u_n\})$ (2.3).

Lemma 2.1 . *Let us assume that:*

1. *the sequence $\{\xi_n\}$ of independent random vectors ξ_n is a stationary random process (in wide sense), i.e.*

$$E\{\xi_n\} = 0;\ E\{\xi_n \xi_n^T\} = D > 0 \qquad \forall\, n\, \in \mathbf{N};$$

2. *the operators $\Lambda(z^{-1})$* of the feedback

$$\Lambda(z^{-1}) := [R + B_0^T Q B(z^{-1})]^{-1} B_0^T Q [A(z^{-1}) - I]$$

and $A_c(z^{-1})$ of the closed loop system defined by

$$A_c(z^{-1}) := \left[A(z^{-1}) - B(z^{-1})\Lambda(z^{-1})\right]^{-1}$$

are stable; here:

$$A(z^{-1}) := \sum_{i=0}^{N_a} A_i z^{-i},\ A_0 := I,\ B(z^{-1}) := \sum_{i=0}^{N_b} B_i z^{-i},\ z^{-i} y_n := y_{n-i};$$

3. *given matrices $Q \in R^{m\times m}$ and $R \in R^{K\times K}$ such that*

$$Q = Q^T \geq 0,\ R = R^T > 0.$$

Then for any ARX model (2.1) and for any L_p-realizable strategy $(p > 2)$ $\{u_n\}$, the following inequality holds

$$J(\{u_n\}) = \liminf_{n\to\infty} \overline{J_n}(\{u_n\}) \overset{a.s.}{\geq} J^* \tag{2.6}$$

where

$$\overline{J_n}(\{u_n\}) := \frac{1}{n}\sum_{t=1}^{n}(\ y_t^T Q\ y_t + u_t^T R\ u_t)$$

and

$$J^* := Tr(QD_{y*}) + Tr(RD_{u*})$$

$$D_{y*} = \frac{1}{2\pi}\int_{-\pi}^{\pi} A_c(e^{i\omega})D\ A_c^T(e^{-i\omega})d\omega$$

$$D_{u*} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \Lambda(e^{i\omega})A_c(e^{i\omega})D\ \left[\Lambda(e^{-i\omega})A_c(e^{-i\omega})\right]^T d\omega$$

correspond to the locally optimal strategy $\{u_n^*\}$ *defined in the following way:*

$$u_n^* := \underset{u_n \in L_P}{argmin}\ J_n(u_n)$$

$$J_n(u_n) := E\{y_n^T Q y_n + u_n^T R u_n / \mathcal{F}_{n-1}\} \tag{2.7}$$

where

$$\mathcal{F}_{n-1} := \sigma(u_1, \xi_1, \ldots, u_{n-1}, \xi_{n-1})$$

is the σ*-algebra generated by* $(u_1, \xi_1, \ldots, u_{n-1}, \xi_{n-1})$.

Proof: From the equation

$$\nabla J_n(u_n^*) \overset{a.s.}{=} 0,$$

we obtain:

$$[R + B_0^T Q B(z^{-1})]\ u_n^* = B_0^T Q[A(z^{-1}) - I]\ y_n$$

or

$$u_n^* = \Lambda(z^{-1})\ y_n. \tag{2.8}$$

Substituting this relation in equation (2.1), we derive that:

$$y_n^* :=\ y_n = \left[A(z^{-1}) - B(z^{-1})\Lambda(z^{-1})\right]^{-1}\xi_n + O\left(\lambda^n\right),\ |\lambda| < 1 \tag{2.9}$$

$$u_n^* = \Lambda(z^{-1})\left[A(z^{-1}) - B(z^{-1})\Lambda(z^{-1})\right]^{-1}\xi_n + O\left(\lambda^n\right). \tag{2.10}$$

The result of this lemma follows from these identities and Parseval's theorem [12]. The lemma is proved. □

Comment 1: The loss of function, which has been investigated in [6] for two types of concrete adaptive strategies, has the following form:

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E\{\|y_t - y_t^*\|^2 / \mathcal{F}_{t-1}\}$$

and hence, use the operation of conditional mathematical expectation. It has been shown that this criterion reaches a minimal possible value γ^2, but the convergence rate has not been studied. Criterion (2.6), considered here, does not use any averaging on the sample space Ω, in comparison with the corresponding functional in other publications [1]-[6], [8].

Comment 2: Condition 2) of the stability of the operator $\Lambda\left(z^{-1}\right)$ practically presents some generalization of "*the minimal phase property*" for the considered class of systems as for $R = 0$; this condition is equivalent to the stability of the operator $B\left(z^{-1}\right)$.

Definition 2 *A sequence* $\{u_n\}$ *is called an* **almost surely asymptotically optimal** L_p**-realizable strategy** *if it is an* L_p*-realizable strategy and*

$$\limsup_{n\to\infty}\ \overline{J_n}\left(\{u_n\}\right) \overset{a.s.}{=} \limsup_{n\to\infty}\ \overline{J_n}\left(\{u_n^*\}\right) = J^*.$$

Hereafter we will denote by $\mathbf{U}_p^*\left(\mathbf{\Theta}\right)$ *the set of all almost surely asymptotically optimal* L_p*-realizable strategies* $\{u_n\}$.

Comment 3:

Notice that the L_p-realizable optimal strategy $\{u_n^*\}$, defined in (2.8), is an almost surely asymptotically optimal L_p-realizable strategy, i.e.

$$\{u_n^*\} \in \mathbf{U}_p^*\left(\mathbf{\Theta}\right).$$

Lemma 1 gives the possibility to formulate the following problem.

Statement of the problem:

For any realizable strategy $\{u_n\} \in U_p^*\left(\mathbf{\Theta}\right)$, *estimate the sharp (reachable) lower bound of the value of a possible convergence rate of the trajectories* $\{y_n, u_n\}$ *to the optimal trajectories* $\{y_n^*, u_n^*\}$ *which use information on parameters of the controlled system.*

This bound presents some generalization of the well-known Cramer-Rao inequality in statistics [13]:

$$E\left\{\|c_n - c^*\|^2\right\} \geq \frac{tr\left\{I^{-1}(q)\right\}}{n} + o(\frac{1}{n^2})$$

where c_n is any "*regular*" estimation of the vector c^*, which uses only available information , i.e.

$$c_n = c_n(y_{1,\dots,}y_{n-1}).$$

The random variables $(y_{1,\ldots}, y_n)$ are suggested to be distributed with a regular density $p(y_{1,\ldots}, y_n; c^*)$:

$$0 < I(q) = E\left\{\nabla \ln p(y_n; c^*)\nabla^T \ln p(y_n; c^*)\right\} < \infty.$$

Here $I(q)$ is the *Fisher information* matrix of a single independent measurement. For the nonstationary case (when unknown parameters c_t^* change in time) such inequality can be rewritten in "a *more general form"* :

$$\liminf_{n \to \infty} \frac{1}{\ln(n)} \sum_{t=1}^{n} E\left\{\|c_t - c_t^*\|_P^2\right\} \geq tr\left\{PI^{-1}(q)\right\},\ 0 \leq P = P^T.$$

The next section states such type of inequalities for adaptive asymptotically optimal control strategies $\{u_n^*\} \in \mathbf{U}_p^*(\mathbf{\Theta})$ for the class of ARX models.

3 General Information Inequality

This section presents two main theorems concerning the estimation of lower bounds for the performance index (2.3). The second theorem is a particular case of the first one for the special choice of the free vector-parameter h.

Theorem 1 *Let us assume that*

1. ξ_n $(n = 1, 2\ldots)$ *are independent, identically distributed random variables with*

$$E\{\xi_n\} = 0;\ E\{\xi_n\xi_n^T\} = D > 0 \qquad \forall n \in \mathbf{N}$$

and with a piecewise continuous differentiable density $q(x)$ *and finite non singular and bounded Fisher information matrix* $I(q)$, *i.e.*

$$0 < I(q) = \int \frac{\nabla q(x)\nabla^T q(x)}{q(x)} dx < \infty, \int \frac{\|\nabla q(x)\|^2}{q(x)} x^2 dx < \infty.$$

2. *A dynamic system described by equation (2.1) is controllable by any almost surely asymptotically optimal* L_4*-realizable strategy*

$$\{u_n\} \in \mathbf{U}_4^*(\mathbf{\Theta})$$

where Θ *is an open and convex set,* $\Theta \subset \mathbf{R}^{m\times(N_a m+N_b K)}$ *and* $\theta \in \Theta$.

Then for any vector

$$h \in \mathbf{R}^{m+K},\ h \neq 0, h = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}, h_1 \in \mathbf{R}^m, h_2 \in \mathbf{R}^K$$

such that

$$0 < G := h_1^T I^{-1}(q) h_1 N_a m + \pi^T(h_2) I^{-1}(q) \pi(h_2) N_b \min(m, K)$$

the following inequality holds:

$$\liminf_{n \to \infty} \sup_{\theta \in \Theta} \frac{1}{\ln(n)} \sum_{t=1}^{n} E\left\{ < P^{1/2} \left[\begin{array}{c} y_n - y_n^* \\ u_n - u_n^* \end{array} \right], h >^2 \right\} \geq \frac{H^2}{G} \qquad (3.1)$$

where

$$0 \leq P = P^T \in \mathbf{R}^{(m+K)\times(m+K)}$$

is some given matrix, and

$$H := h^T, I^{-1}(q) \left(h_1 N_a Tr\left\{ D^{1/2} D_{y^*}^{-1/2} \right\} + \pi(h_2) N_b Tr\left\{ D^{1/2} U D_{u^*}^{-1/2} \Phi \right\} \right)$$

$$, := P^{1/2} \left[\begin{array}{c} B_0 \\ I_{K\times K} \end{array} \right] \left[B_0^T Q B_0 + R \right]^{-1} B_0^T Q$$

$$U := \{\delta_{i,j}\} \qquad i = 1, \ldots, m, \; j = 1, \ldots, K, \qquad \delta_{i,j} \text{ - Kronecker symbol}$$

$$\Phi := \frac{1}{2\pi i} \oint_{\|q\|=1} \Lambda(q) \left[A(q) - B(q)\Lambda(q) \right]^{-1} q^{N_a - 1} dq$$

$$\pi^T \left(\begin{array}{cccc} a_1, & \ldots & , a_K \end{array} \right) := \left\{ \begin{array}{ll} \left(\begin{array}{cccccc} a_1, & \ldots & a_K, & 0, & \ldots & ,0 \end{array} \right)^T & K < m \\ \left(\begin{array}{ccc} a_1, & \ldots & , a_m \end{array} \right)^T & K \geq m \end{array} \right. .$$

Proof: We separate (for more clarity) this proof into several intermediate steps (statements).

Without lost of generality, we can consider $\Theta = V(\theta_0, r)$ defined by the following way:

$$V(\theta_0, r) \;:= \left\{ \theta \in \mathbf{R}^{m \times (N_a m + N_b K)} \mid \| \theta - \theta_0 \| < r \right\}$$

for some positive constant $r > 0$. Define also

$$\widehat{\vartheta}_n := \pi_{[a,b]} \left\{ \left\langle P^{1/2} \left[y_n^T \vdots u_n^T \right], h \right\rangle \right\}$$

where $\pi_{[a,b]}$ denote the projection operator on the real interval $[a, b]$

$$a := \langle, \theta_0 \; z_n, h \rangle - r \|, h\| \|z_n\|, \; b := \langle, \theta_0 \; z_n, h \rangle + r \|, h\| \|z_n\|$$

and

$$\varepsilon_n := \widehat{\vartheta}_n - \langle, \theta \; z_n, h \rangle .$$

Then we derive

$$\left|\left\langle P^{1/2}\left[y_n^T \vdots u_n^T\right], h\right\rangle - \langle, \theta\ z_n, h\rangle\right| =$$

$$= \left|\widehat{\vartheta}_n - \langle, \theta\ z_n, h\rangle\right| + \inf_{x\in[a,b]}\left|x - \left\langle P^{1/2}\left[y_n^T \vdots u_n^T\right], h\right\rangle\right| \geq$$

$$\geq \left|\widehat{\vartheta}_n - \langle, \theta\ z_n, h\rangle\right| = \varepsilon_n.$$

Hence

$$E\left\{\left\langle P^{1/2}\left[\begin{array}{c} y_n - y_n^* \\ u_n - u_n^* \end{array}\right], h\right\rangle^2\right\} \geq E\left\{\varepsilon_n^2\right\}. \tag{3.2}$$

Statement 1 Consider the function Δ_n

$$\Delta_n : \mathbf{R}^{m\times(N_a m + N_b K)} \to \mathbf{R}^{m\times(N_a m + N_b K)}$$

which is defined for any $\theta \in \mathbf{R}^{m\times(N_a m + N_b K)}$ as follows:

$$\Delta_n(\theta) = E\{\varepsilon_n W_n\}$$

$$W_n := \left[W_1^n, \ldots, W_N^n\right], \qquad W_i^n := \Xi_i h \zeta_i^{nT} C_i \tag{3.3}$$

where matrices $\Xi_i \in \mathbf{R}^{m\times(m+K)}$, $C_i \in \mathbf{R}^{m\times m_i}$ are defined as

$$\Xi_i := \left\{\begin{array}{ll} I^{-1\cdot}(q)\left[\begin{array}{ccc} I_{m\times m} & \vdots & 0 \end{array}\right] & i = 1, \ldots, N_a \\ I^{-1\cdot}(q)\left[\begin{array}{ccc} 0 & \vdots & U \end{array}\right] & i = N_a + 1, \ldots, N_a + N_b \end{array}\right.$$

$$C_i := \left\{\begin{array}{ll} D^{-1/2} D_{y^*}^{-1/2} & i = 1, \ldots, N_a \\ D^{-1/2} U D_{u^*}^{-1/2} & i = N_a + 1, \ldots, N_a + N_b \end{array}\right.$$

$$\zeta_i^n := y_{n-i} + \theta\ z_{n-i} - B_0 u_{n-i}$$

(which coincide with ξ_{n-i} if $n > i$).

Then for some positive constant k_1, the following inequality holds:

$$div \Delta_n(\Theta) + \sum_{i=1}^{N} h^T, \Xi_i h E\left\{\zeta_i^{nT} C_i\ z_i^n\right\}$$

$$\leq k_1 \sum_{i=1}^{N} E^{1/2}\left\{\varepsilon_n^2\right\} E^{1/2}\left\{\sum_{i=1}^{N}\left\|\ z_i^{n-i}\right\|^2\right\} +$$

$$+ \sum_{i=1}^{N} E^{1/2}\left\{\varepsilon_n^2\right\} E^{1/2}\left\{\left(Tr\left(W_i^{n\ T} G_i^n\right)\right)^2\right\}. \tag{3.4}$$

Proof of statement 1: Consider the joint density of $\xi_1, ..., \xi_n$

$$Q_n = Q_n(\theta, y_1, ..., y_n) := \prod_{t=1}^{n} q\left(\zeta_0^t\right).$$

Define

$$m_i := \begin{cases} m & i = 1, ..., N_a \\ K & i = N_a + 1, ..., N_a + N_b \end{cases}$$

and calculate $div\Delta_n(\theta)$

$$div\Delta_n(\theta) = \sum_{i=1}^{N} \sum_{j=1}^{m} \sum_{k=1}^{m_i} \frac{\partial}{\partial \theta^i_{j\ k}} \int_{\Omega} \varepsilon_n \left(W_i^n\right)_{jk} Q_n \prod_{t=1}^{n} dy_t$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{m} \sum_{k=1}^{m_i} E\{-(h^T,\ )_j \left(z_i^n\right)_k \left(W_i^n\right)_{jk} +$$

$$+\varepsilon_n((\Xi_i h)_j ([C_i]_{j\ k} \left(z_i^{n-i}\right)_k + \left(W_i^n\right)_{jk} \left(G_i^n\right)_{jk})\}$$

where

$$G_i^n := \sum_{t=1}^{n} \nabla ln(x_t)\ z_i^{t\ T} \tag{3.5}$$

$$x_t := y_t + \theta\ z_t - B_0 u_t$$

$$\nabla ln(x_t) := \nabla ln(x) \mid_{x=x_t}.$$

So, we obtain

$$div\Delta_n(\Theta) + \sum_{i=1}^{N} h^T,\ \Xi_i h E\left\{\zeta_{n-i}^T C_i^n\ z_i^n\right\} =$$

$$\sum_{i=1}^{N} E\left\{\varepsilon_n\left(\sum_{i=1}^{N} h^T \Xi_i C_i^T z_i^{n-i} + Tr\left(W_i^{n\ T} G_i^n\right)\right)\right\}.$$

Then (3.4) follows from this relation if we apply the Cauchy-Bouniakowsky inequality to the last term. □

Let $\{\alpha_n\}$ be any sequence of functions such that

$$\alpha_n : V(\theta_0, r) \to \mathbf{R}.$$

Define now the averaging operator $\overline{E}_{V,n}$ as follows:

$$\overline{E}_{V,n}\{\alpha_n\} := \frac{1}{n} \sum_{t=n_1}^{n} \frac{1}{vol(V(\theta_0, r))} \int_{V(\theta_0,r)} E\{\alpha_t(\theta)\}\, dv$$

$$n \geq n_1;\; n_1 := 1 + 2N.$$

Dividing (3.4) by n and then applying the operator $\overline{E}_{V,n}$ to both sides of this inequality, we obtain

$$\frac{1}{vol(V(\theta_0,r))} \int\limits_{V(\theta_0,r)} \frac{1}{n}\sum_{t=n_1}^{n} \frac{1}{t} div \Delta_t(\theta) dv +$$
$$+\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}\sum_{i=1}^{N} h^T,\, \Xi_i h E\left\{\zeta_i^{tT} C_i\ z_i^t\right\} \leq$$
$$\leq \overline{E}_{V,n}^{1/2}\left\{\varepsilon_n^2\right\}\left(k_1 \overline{E}_{V,n}^{1/2}\left\{\frac{1}{n^2}\sum_{i=1}^{2N} \|\ z_i^n\|^2\right\} +\right.$$
$$\left.+\overline{E}_{V,n}^{1/2}\left\{\frac{1}{n^2}\sum_{i=1}^{N} Tr\left(W_i^{n\ T} G_i^n\right)^2\right\}\right). \tag{3.6}$$

We will try now to investigate the behavior of the term

$$\overline{E}_{V,n}^{1/2}\left\{\frac{1}{n^2}\sum_{i=1}^{N} Tr\left(W_i^{n\ T} G_i^n\right)^2\right\}.$$

Statement 2: Let $\gamma > 0$ and

$$\Psi(j,s,i,t) := h^T \Xi_j^T \nabla \ln q(x_s) \nabla^T \ln q(x_t) \Xi_i h \in \mathbf{R}^1.$$

Then there exists a positive constant k_3 such that

$$\overline{E}_{V,n}^{1/2}\left\{\frac{1}{n^2}\sum_{i=1}^{N} Tr\left(W_i^{nT} G_i^n\right)^2\right\} \leq$$
$$\leq \frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t^2}[(1+\gamma)\sum_{i=1}^{N}\sum_{k=1}^{t-N-1} E\left\{\Psi(i,k,i,k)\right\} E_V\left\{z_i^{k\ T} C_i^T \zeta_i^t \zeta_i^{tT} C_i z_i^k\right\}$$
$$+ (1+\gamma^{-1}) k_3 \sup_{l\in\mathbf{N}} \sup E_V\left\{\|z_l\|^2\right\}]^{1/2} \tag{3.7}$$

where

$$\overline{E}_V\left\{X\right\} := \frac{1}{vol(V(\theta_0,r))}\int\limits_{V(\theta_0,r)} X(\theta) dv.$$

Proof of statement 2: From the definition of G_i^n it follows that

$$G_i^n = G_i^{n-N-1} + \sum_{t=n-N+1}^{n} \nabla ln(x) \mid_{x=x_t}\ z_i^{t\ T}.$$

Applying now the inequality

$$(a+b)^2 \le (1+\gamma)a^2 + (1+\gamma^{-1})b^2, \ \forall \gamma > 0, \tag{3.8}$$

we obtain

$$\left(\sum_{i=1}^{N} TrW_i^{n\ T} G_i^n\right)^2 \le (1+\gamma)\left(\sum_{i=1}^{N} TrW_i^{n\ T} G_i^{n-N-1}\right)^2 +$$

$$+ (1+\gamma^{-1})\left(\sum_{i=1}^{N}\sum_{t=n-N}^{n} TrW_i^{n\ T} \nabla ln(x)\ z_i^{t\ T}\right)^2. \tag{3.9}$$

Taking into account that

$$\left(\sum_{i=1}^{N} TrW_i^{n\ T} G_i^{n-N-1}\right)^2 = \left(\sum_{i=1}^{N}\sum_{t=1}^{n-N-1} TrW_i^{n\ T} \nabla ln(x_t)\ z_i^{t\ T}\right)^2 =$$

$$\sum_{i=1}^{N}\sum_{t=1}^{n-N-1}\sum_{j=1}^{N}\sum_{s=1}^{n-N-1} \{z_j^{t\ T} C_j^T \zeta_j^n h^T \Xi_j^T \nabla \ln q(x_s) \nabla^T \ln q(x_t) \Xi_i h \zeta_i^{nT} C_i z_i^t\}$$

and using the definition of $\Psi(i,t,i,t)$:

$$E\{\Psi(i,t,i,t)\} = h^T \Xi_i^T I(q) \Xi_i h,$$

we obtain

$$E\left\{\sum_{i=1}^{N} Tr\left(W_i^{nT} G_i^{n-N-1}\right)^2\right\} =$$
$$= \sum_{i=1}^{N}\sum_{t=1}^{n-N-1} E\{\Psi(i,t,i,t)\} E\left\{\ z_i^{t\ T} C_i^T \zeta_i^n \zeta_i^{nT} C_i z_i^t\ \right\}$$

and hence the following inequality is fulfilled:

$$E\left\{\left(\sum_{i=1}^{N}\sum_{t=n-N}^{n} TrW_i^{nT} \nabla ln(x_t)\ z_i^{t\ T}\right)^2\right\} \le k_2 \sum_{i=1}^{N}\sum_{t=n-N}^{n} E\left\{\|\ z_i^t\ \|^2\right\}$$

for some positive constant k_2. Then (3.7) immediately follows from these relations.

Statement 3. For each t define the variable v_t as follows

$$v_t := \frac{(1+\gamma)}{t^2}\sum_{i=1}^{N}\sum_{k=1}^{t-N-1} E\{\Psi(i,k,i,k)\}\, \overline{E}_V \left\{z_i^{k\ T} C_i^T \zeta_i^t \zeta_i^{tT} C_i z_i^k\right\}. \tag{3.10}$$

Then

a) there exist two positive constants k_4 and k_5 such that

$$\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}\sum_{i=1}^{N} h^T, \, \Xi_i h E\left\{\zeta_i^{tT} C_i \; z_i^t\right\} \leq$$

$$\leq \left(\sup_{\theta\in V(\theta_0,r)} \overline{E}_n\left\{\varepsilon_n^2\right\}\right)^{1/2}\left(\frac{1}{n}\sum_{t=n_1}^{n} v_t\right)^{1/2} +$$

$$+\left(\sup_{\theta\in V(0,r)} \overline{E}_{V,n}\left\{\varepsilon_n^2\right\}\right)^{1/2}$$

$$\left(\frac{k_4}{n^{1/2}} + (2+\gamma^{-1})k_5 \overline{E}_{V,n}\left\{\frac{1}{n^2}\sum_{i=1}^{2N}\| \; z_i^n\|^2\right\}\right) \tag{3.11}$$

where

$$\overline{E}_n\left\{\alpha_n\right\} := \frac{1}{n}\sum_{t=n_1}^{n} E\left\{\alpha_t(\theta)\right\}.$$

b) The following presentation holds:

$$\lim_{n\to\infty} tv_t = (1+\gamma)\left(h_1^T I^{-1}(q) h_1 N_a m + \pi^T\left(h_2\right) I^{-1}(q)\pi\left(h_2\right) N_b \min(m,K)\right) \tag{3.12}$$

Proof of statement 3.

a) Using the definitions (3.3), (3.5), (3.10) and applying the inequalities (3.8) and (3.9), we derive

$$\overline{E}_{V,n}^{1/2}\left\{\frac{1}{n^2}\sum_{i=1}^{N} Tr\left(W_i^{nT} G_i^n\right)^2\right\} \leq$$

$$\leq \left(\frac{1}{n}\sum_{t=n_1}^{n} v_t + (1+\gamma^{-1})k_3 \overline{E}_{V,n}\left\{\frac{1}{n^2}\sum_{i=1}^{2N}\| \; z_i^n\|^2\right\}\right)^{1/2}$$

and hence

$$\frac{1}{vol(V(\theta_0,r))}\int\limits_{V(\theta_0,r)} \frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t} div\Delta_n(\theta) dv +$$

$$\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}\sum_{i=1}^{N} h^T, \, \Xi_i h E\left\{\zeta_i^{tT} C_i \; z_i^t\right\}$$

$$\leq \overline{E}_{V,n}^{1/2}\left\{\varepsilon_n^2\right\} k_1 \overline{E}_{V,n}^{1/2}\left\{\frac{1}{n^2}\sum_{i=1}^{2N}\| \; z_i^n\|^2\right\} +$$

$$+\overline{E}_{V,n}^{1/2}\left\{\varepsilon_n^2\right\}\left(\frac{1}{n}\sum_{t=n_1}^{n}v_t+(1+\gamma^{-1})k_3\overline{E}_{V,n}\left\{\frac{1}{n^2}\sum_{i=1}^{2N}\|\ z_i^n\|^2\right\}\right)^{1/2}.$$

Applying the Gauss-Ostrogradskii theorem, which states that

$$\frac{1}{vol(V(\theta o,r))}\int_{V(\theta_0,r)}div\left(\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}\Delta_t(\theta)\right)dv=$$

$$=\int_{\partial V(\theta_0,r)}<\frac{\theta-\theta_0}{r},\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}\Delta_t(\theta)>dv\geq-\int_{\partial V(\theta_0,r)}\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}\|\ \Delta_t(\theta)\|\ dv,$$

and taking into account the inequalities

$$\|\Delta_t(\theta)\|^2\leq E\left\{\varepsilon_t^2\right\}E\left\{\|\ W_t\|^2\right\}\leq k_4E\left\{\varepsilon_t^2\right\}$$

which are true for some positive constant k_4, we obtain the estimation

$$\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}\|\ \Delta_t(\theta)\|\leq\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}k_4E^{1/2}\left\{\varepsilon_t^2\right\}\leq k_4n^{1/2}\left(\sup_{\theta\in V(\theta_0,r)}\overline{E}_n\left\{\varepsilon_n^2\right\}\right)^{1/2}$$

From these inequalities, taking into account Jensen's inequality, we finally derive (3.11).

b) Using the inequality from a) we can conclude that

$$\left(\frac{1}{n}\sum_{t=n_1}^{n}v_t\right)^{1/2}\sup_{\theta\in V(0,r)}\overline{E}_n^{1/2}\left\{\varepsilon_n^2\right\}+$$

$$+\left(\frac{k_4}{n^{1/2}}+(2+\gamma^{-1})k_5\overline{E}_{V,n}\left\{\frac{1}{n^2}\sum_{i=1}^{2N}\|\ z_{n-i}\|^2\right\}\right)\sup_{\theta\in V(0,r)}\overline{E}_n^{1/2}\left\{\varepsilon_n^2\right\}\geq$$

$$\geq\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}\sum_{i=1}^{N_a}h^T,\ \Xi_ihTr(E\left\{D^{1/2}D_{y*}^{-1/2}\right\})+$$

$$+\frac{1}{n}\sum_{t=n_1}^{n}\frac{1}{t}\sum_{i=N_a+1}^{N}h^T,\ \Xi_ihTr(D^{1/2}UD_{u*}^{-1/2}E\left\{u_{t-i+N_a}\zeta_{t-i}^T\right\}).$$

Also we have

$$tv_t:=\frac{(1+\gamma)}{t}\sum_{i=1}^{N_a}\sum_{k=1}^{t-N-1}h_1^TI^{-1}(q)h_1\overline{E}_V\left\{z_i^{k\ T}C_i^TDC_iz_i^k\right\}+$$

$$+\frac{(1+\gamma)}{t}\sum_{i=1}^{N}\sum_{k=1}^{t-N-1}E\left\{\Psi(i,k,i,k)\right\}\overline{E}_V\left\{z_i^{k\ T}C_i^T\zeta_i^t\zeta_i^{tT}C_iz_i^k\right\}.$$

Calculating the limit in the last expression, we obtain (3.12).

Now we return to the proof of the main theorem. If we take into account the following facts:

1) if

$$\lim_{n\to\infty}a_n=a,$$

then (Töeplitz's lemma)

$$\lim_{n\to\infty}\frac{1}{\ln n}\sum_{t=1}^{n}\frac{1}{t}a_t=a.$$

2) If

$$i=N_a+1,\ldots,N_a+N_b,$$

then

$$E\left\{\zeta_i^{tT}C_i\ z_i^t\right\}=E\left\{D^{-1/2}UD_{u^*}^{-1/2}\Phi D\right\}$$

where

$$\Phi=\frac{1}{2\pi i}\oint\limits_{\|\ q\|=1}\Lambda(q)\left[A(q)-B(q)\Lambda(q)\right]^{-1}q^{N_a-1}dq.$$

Taking into account (3.2), (3.4), (3.11), (3.12), we finally obtain (for any $r>0$)

$$\liminf_{n\ \to\ \infty}\sup_{\theta\in\Theta}\overline{E}_n\left\{\varepsilon_n^2\right\}\geq$$

$$\geq\frac{\left[h^T,\ I^{-1}(q)\left(h_1N_aTrE\left\{D^{1/2}D_{y^*}^{-1/2}\right\}+\pi\left(h_2\right)N_bTrE\left\{D^{1/2}UD_{u^*}^{-1/2}\Phi\right\}\right)\right]^2}{(1+\gamma)\left(h_1^TI^{-1}(q)h_1N_am+\pi^T\left(h_2\right)I^{-1}(q)\pi\left(h_2\right)N_b\min(m,K)\right)}$$

The main result follows from this inequality, if we consider that γ is any positive number. □

Theorem 2 *Under the assumptions of theorem 1 the following inequality holds:*

$$\liminf_{n\ \to\ \infty}\sup_{\theta\in\Theta}\frac{1}{\ln n}\sum_{t=1}^{n}E\left\{\left\|\begin{array}{c}y_n-y_n^*\\u_n-u_n^*\end{array}\right\|_{P^*}^2\right\}\geq\Xi_1+\Xi_2\qquad(3.13)$$

where

$$\Xi_1:=\frac{N_a}{m}\sum_{i=1}^{N_a}\left(\left[,\ I^{-1}(q)\left[I_{m\times m}\vdots0\right]\right]_{ii}\right)^2Tr^2\left\{D^{1/2}D_{y^*}^{-1/2}\right\}$$

$$\Xi_2 := \frac{N_b}{\min(m,K)} \sum_{i=N_a+1}^{N_a+\min(m,K)} \left(\left[, I^{-1}(q) \left[0 \vdots U \right] \right]_{ii} \right)^2 Tr^2 \left\{ D^{1/2} U D_{u^*}^{-1/2} \Phi \right\}$$

$$P^* := P^{1/2} \Upsilon P^{1/2}, \quad \Upsilon := \sum_{i=1}^{m+\min(m,K)} e_i e_i^T$$

and $\{e_i\}_{i=1}^{m+K}$ *represents the canonical base of* $\mathbf{R}^{m+K}$.

Proof: Applying sequentially the result of theorem 1 to the vector

$$h = \left(\left[I^{-1}(q) \right]_{ii} \right)^{1/2} e_i, \quad i = 1, ..., m + min(m, K)$$

$$h = \left(\left[I^{-1}(q) \right]_{i-m, i-m} \right)^{1/2} e_i, \quad i = m+1, ..., m + min(m, K)$$

and adding these inequalities, we obtain (3.13). □

Comment 4:

Note that the relation (3.13) depends on the *a priori constraint* on $\theta \in V(\theta_0, r)$ and hence, depends on the radius $r > 0$ of *"a priori knowledge ."* Results, given in [6], are not dependent on a priori information on the parameters of the system. They are stated for any concrete value θ_0 (which is fixed, but is not known). Making $r \to 0$ in the inequality (3.13), we have:

$$V(\theta_0, r) \to \theta_0,$$

i.e., we could be able to estimate the convergence rate of the procedure, suggested in [6], as a partial case of the main inequality.

Using the inequality (3.13) we can suggest the following definition of *"the convergence rate"* notion.

Definition 3 *For the class of adaptive control problems (under assumptions of the theorem 1) we can call the value*

$$C_r := \left[\liminf_{n \to \infty} \sup_{\theta \in \Theta} \frac{1}{\ln n} \sum_{t=1}^{n} E \left\{ \left\| \begin{array}{c} y_n - y_n^* \\ u_n - u_n^* \end{array} \right\|_P^2 \right\} \right]^{-1}$$

the **convergence rate** *of the corresponding adaptive asymptotically optimal* L_4 *-realizable control strategy* $\{u_n\} \in \mathbf{U}_4^*(\mathbf{\Theta})$.

Definition 4 *The quantity*

$$C_r^* = C_r^*(Q, R, P) := [\Xi]^{-1}$$

is called to be **the maximal possible convergence rate** *in the class of the problems under consideration. If there exists some adaptive strategy* $\{u_n^*\} \in \mathbf{U}_4^*(\mathbf{\Theta})$ *such that* $C_r^* = C_r$, *then the information inequality, written in the form*

$$C_r^* \geq C_r$$

is said to be **sharp**.

4 Two Partial Cases

Using the previous theorem we can derive the corresponding information inequalities for two very important (from the practical point of view) problems of the adaptive control: regulation and tracking problems.

4.1 Regulation problem for ARX models in $\mathbf{L}_4$, without delay in control action $(N_b = 0)$

Consider the system (2.1), then try to find the strategy $\{u_n^*\} \in \mathbf{U}_4^*(\mathbf{\Theta})$ such that

$$J(\{u_n\}) = \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E\left\{\|y_n\|^2\right\}$$

reaches its minimum value.

It can be shown that the optimal trajectory $\{y_n^*\}$ is precisely equal to $\{\xi_n\}$. Then the main theorem applied to this problem gives the lower bound for the convergence rate of the adaptive optimal control strategies without complete information on true values of parameters. Under assumptions of the main theorem we obtain [8]:

Corollary 1

$$\liminf_{n\,\to\,\infty} \sup_{\theta\in\Theta} \frac{1}{\ln n} \sum_{t=1}^{n} \left(E\left\{\|y_n\|^2\right\} - Tr(D)\right) \geq NmTr(I^{-1}(q)). \tag{4.1}$$

Proof: This inequality follows directly from (3.1), if we accept

$$P = \begin{bmatrix} I_{m\times m} & 0 \\ 0 & 0 \end{bmatrix},\ Q = I,\ B_0 = I,\ N_b = 0\ (i.e.\ N_a = N).$$

□

4.2 Tracking problem for ARX models in $\mathbf{L}_4$

Consider again the system (2.1). Find the strategy $\{u_n^*\} \in \mathbf{U}_4^*(\mathbf{\Theta})$, which minimizes the criterion:

$$J(\{u_n\}) := \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E\left\{\|y_n\|_Q^2 + \|u_n\|_R^2\right\}.$$

The optimal trajectories $\{y_n^*\}$ and $\{u_n^*\}$ satisfy (2.9) and (2.10) respectively. Considering the "tracking problem" for the optimal trajectories $\{y_n^*\}$ and $\{u_n^*\}$ and applying the basic information inequality (3.1), we obtain:

Corollary 2

$$\liminf_{n \to \infty} \sup_{\theta\in\Theta} \frac{1}{\ln n} \sum_{t=1}^{n} E\left\{ \|y_n - y_n^*\|_Q^2 + \|u_n - u_n^*\|_R^2 \right\} \geq \Xi_1 + \Xi_2 \qquad (4.2)$$

$$\Xi_1 := \frac{N_a}{m} Tr^2 \left\{ D^{1/2} D_{y^*}^{-1/2} \right\} \sum_{i=1}^{m} \left[Q^{1/2} B_0 \left[B_0^T Q B_0 + R \right]^{-1} B_0^T Q I^{-1}(q) \right]_{ii}^2$$

$$\Xi_2 := \frac{N_b Tr^2 \left\{ D^{1/2} D_{u^*}^{-1/2} \Phi \right\}}{\min(m, K)} \sum_{i=1}^{m} \left[R^{1/2} B_0 \left[B_0^T Q B_0 + R \right]^{-1} B_0^T Q I^{-1}(q) \right]_{ii}^2$$

Proof: This inequality follows directly from (3.1) if we accept

$$m = K, \ P = \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$$

and apply the formula 3.13. □

5 Examples

The next examples illustrate the approach suggested above for the estimation of the maximum possible adaptation rate for the different models.

1. *Regulation problem for an ARX model in L_4 (SISO-case).*

   Consider the system described by the following difference equation:

   $$y_n = 1.7 y_{n-1} - 0.7 y_{n-2} + u_n + 0.5 u_{n-1} + \xi_n$$

   where $\{\xi_n\}$ is a centered random sequence with independent and normal distributed values with variance

   $$D = E\{\xi_n^2\} = \sigma_\xi^2.$$

   The optimal control strategy $\{u_n^*\} \in \mathbf{U}_4^*(\boldsymbol{\Theta})$ which minimizes the criterion

   $$J(\{u_n\}) := \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E\left\{ \|y_n\|^2 \right\}$$

   can be calculated in accordance to expression (2.10) and is equal to:

   $$u_n^* = -0.5 u_{n-1}^* - 1.7 y_{n-1} + 0.7 y_{n-2};$$

   furthermore, $y_n^* = \xi_n$. Thus, applying the formula (4.1) for the Gaussian case, when

   $$D_{y^*} = D = I^{-1}(q) = \sigma_\xi^2$$

$$N_a = 2,\ K = m = 1,\ N_b = 1,\ B_0 = 1,\ Q = 1,\ R = 0,$$

we obtain the following inequality and

$$\liminf_{n\to\infty} \sup_{\theta\in\Theta} \frac{1}{\ln(n)} \sum_{t=1}^{n} \left(E\left\{ |y_n|^2 \right\} - \sigma_\xi^2 \right) \geq 2\sigma_\xi^2$$

which is true for any adaptive strategy $\{u_n\} \in \mathbf{U}_4^* (\mathbf{\Theta})$.

2. *Tracking problem for ARX models in* L_4 *(SISO-case).*
   Consider the system given by

$$y_n = -2y_{n-1} + y_{n-2} + u_n - 2.5u_{n-1} + \xi_n.$$

For

$$Q := 1, R := 1, P := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, D = 1$$

from lemma 1 it follows that:

$$\Lambda(z^{-1}) = \frac{2z^{-2} - 4z^{-1}}{5z^{-1} - 4}$$

$$A_c(z^{-1}) = \frac{5z^{-1} - 4}{2z^{-2} + z^{-1} - 4}$$

$$\Lambda(z^{-1})A_c(z^{-1}) = \frac{2z^{-2} - 4z^{-1}}{2z^{-2} + z^{-1} - 4}$$

$$u_n^* = 1.25u_{n-1} + y_{n-1} - 0.5y_{n-2}$$

$$D_{y^*} = 2.33.. \qquad D_{u^*} = 1.33.. \qquad \Phi = -0.25..$$

$$E\left\{ |y_n - y_n^*|_Q^2 + |u_n - u_n^*|_R^2 \right\} = E\left\{ |y_n - y_n^*|^2 + |u_n - u_n^*|^2 \right\},$$

$$J^* = 3.66\ldots$$

and applying inequality (4.2), we obtain the final result for $R = Q = 1$:

$$\liminf_{n\to\infty} \sup_{\theta\in\Theta} \frac{1}{\ln(n)} \sum_{t=1}^{n} E\left\{ \ |y_n - y_n^*|^2 + |u_n - u_n^*|^2 \right\} \geq 0.9I^{-1}(q).$$

3. *Tracking problem for ARX models in* L_2 *(MIMO-case):*
   Consider the system

$$A(z^{-1})y_n = B(z^{-1})u_n + \xi_n$$

with

$$A(z^{-1}) := \begin{bmatrix} 0.1 + z^{-1} & 0.06z^{-1} \\ -0.2z^{-1} & 0.3 + z^{-1} \end{bmatrix}$$

$$B(z^{-1}) := \begin{bmatrix} 1 + 2z^{-1} & 1 + 0.6z^{-1} \\ 0.5 + 0.4z^{-1} & 1 + 0.3z^{-1} \end{bmatrix}.$$

Let $\{\xi_n\}$ be some i.i.d. random process. Let us calculate the control strategy $\{u_n^*\} \in \mathbf{U}_4^* (\mathbf{\Theta})$, which minimizes

$$J(\{u_n\}) := \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E\left\{ \|y_n\|^2 + \|u_n\|^2 \right\}.$$

Applying now the result of corollary (4.2) for $R = Q = I$, we derive the following information inequality:

$$\liminf_{n \to \infty} \sup_{\theta\in\Theta} \frac{1}{\ln(n)} \sum_{t=1}^{n} E\ \left\{ \|y_n - y_n^*\|^2 + \|u_n - u_n^*\|^2 \right\} \geq$$

$$\geq 0.4149\alpha_{11}^2 + 0.3926\alpha_{11}\alpha_{12} + 0.0968\alpha_{12}^2 + 0.0034\alpha_{12}\alpha_{22} + 0.0008\alpha_{22}^2$$

where

$$I^{-1}(p) = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{12} & \alpha_{22} \end{bmatrix}$$

is the Fisher Information matrix of each random variable ξ_n. Notice that for the Gaussian processes $I^{-1}(p) = D$.

6 Concluding Remarks

- This paper suggests an information bound on the convergence rate for the class of adaptive strategies in the ARX model Control Problem.
- This bound gives **an objective estimation of the quality** of any available adaptive algorithm with respect to the optimal control strategy which uses full information about the parameters of a controlled system. If some considered algorithm reaches this bound it can be considered as "*an effective adaptive strategy*."
- If not, we can call such procedure "*not good*," because there exist some other algorithms, [7] and [8], which are better ("*more close*" to the optimal strategy) than the considered one. In [7] an algorithm has been described. It contains the nonlinear transformation φ in the identification part of the adaptive control strategy, in order to

reach the optimal converge rate. This identification algorithm for the system under consideration given by the equation

$$y_n = -\sum_{i=1}^{N_a} a_i y_{n-i} + u_n + \xi_n$$

has the following structure:

$$\widehat{\theta}_n = \pi_\Theta \left\{ \widehat{\theta}_{n-1} + , {}_n z_n \varphi(\varepsilon_n) \right\} \tag{6.1}$$

where

$$\theta^T := [a_1, ..., a_{Na}], \; z_n^T := [-y_{n-1}, ..., -y_{n-N_a}], \; \varepsilon_n := y_n - \widehat{\theta}_{n-1}^T z_n - u_n$$

and π_Θ is the projection operator to a given convex set Θ of a priori values of the unknown parameters.

- From the proof of the Information Inequality derived above (see (3.2), we can conclude that for any concrete adaptive strategy the adaptation rate depends on the applied identification algorithm:

  if the estimates, generated by this algorithm, are "asymptotically effective," then the corresponding adaptation process is "effective" too.

  Such asymptotically effective identification procedure can be given in the form (6.1) with the nonlinear transformation of a residual [10]:

$$\varphi(z) = -I^{-1}(q_\xi) \nabla \ln q_\xi(z).$$

So, only for Gaussian noise distribution can we use a linear identification algorithm with $\varphi(z) = z$.

References

[1] K. Aström. *Introduction to Stochastic Control Theory.* New York: Academic Press, 1970.

[2] V.N. Fomin, A.L. Fradkov and V.A. Yakubovich. *Adaptive Control of Dynamic Systems.* Moscow: Nauka, 1981 (in Russian).

[3] G.S. Goodwin and K.S. Sin. *Adaptive Filtering, Prediction and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1984.

[4] K. Aström, U. Borisson, L. Ljung, and K.B. Witternmark. Theory and applications of self-tuning regulators. *Automatica* **13** (1977), 457-476.

[5] G.K. Kelman, A.S. Poznyak and A.V. Chernitser. Adaptive locally optimal control. *Int. J. Systems Science,* **12(2)** (1981), 235-254.

[6] Graham C. Goodwin, Peter J. Ramadge, and Peter E. Caines. Discrete time stochastic adaptive control. *SIAM J. Control and Optimization,* **19(6)**, November (1981), 829-853.

[7] A.S. Nemirovskii and Ya.Z. Tsypkin. On optimal adaptive control algorithms, *Automat. Telemekh.*, **12** (1984), 64-67.

[8] A.V. Nazin. Information inequalities in adaptive control of a multidimensional linear plant, *Automation and Remote Control* **53(4)** April (1992), 570-577.

[9] A.S. Poznyak. New form of the strong large numbers law for dependent sequences and its applications to the analysis of recurrent identification algorithms. *Report of Russian Academy of Sciences, ser. Kibernetica i Teoriya Regulirovaniya,* **323(N4)** (1992), 645-648.

[10] A.S. Poznyak and S.N. Tikhonov. Strong consistency of the extended least square method with nonlinear error transformation, *Automation and Remote Control,* **51(8)** January (1991), 1105-1112.

[11] A.S. Poznyak and S.N. Tikhonov. Asymptotic normality and the rate of convergence of recursive stochastic process with a nonlinear residual transformation. *Automation and Remote Control,* **53(8)** January (1993), 1222-1230.

[12] M.J. Grimble, and M.A. Johnson. *Optimal Control and Stochastic Estimation, Theory and Applications,* Vol II. New York: John Wiley & Sons, 1987.

[13] Harald Crámer. *Mathematical Methods of Statistics.* Princeton: Princeton University Press, 1946.

CINVESTAV-IPN, Seccion de Control Automático. AP14-740. México 07300, México

CINVESTAV-IPN, Seccion de Control Automático. AP14-740. México 07300, México

Communicated by Clyde F. Martin