Control and Cybernetics

VOL. 21 (1992) No. 3/4

Design of multistep algorithms and local optimal input for dynamic system identification

by

A. Yastrebov

Department of Numerical Mathematics and Cybernetics
Nizhni Novgorod State University
603600, Nizhni Novgorod, Gagarin Ave. 23, Russia

M. Grzywaczewski

Department of Automatic Control
Technical University of Kielce
Al.1000–lecia Państwa Polskiego 7,
25–314 Kielce, Poland

A problem of on–line identification of dynamic systems is analyzed. For purposes of solution of this problem a certain class of multistep adaptive methods is proposed. The sufficient convergence conditions with probability one for this class are obtained and supplemented by results of computer analysis. The principle of the synthesis of multistep local optimal input design is given.

Keywords Multistep algorithms, estimation, local input design, dynamic system identification.

1. Introduction

In solving of problems of parameter identification of controlled discrete dynamics models the recursive adaptive algorithms realized through computer techniques are often used [6]. The accuracy of identification of unknown parameters can be increased by using the knowledge about values of estimation in former iterations [1, 2]. It is important in synthesis of either control or estimation procedure to use the data concerning former time steps. The recursive methods of input design or identification which implement the above principle are called multistep algorithms or M–algorithms [3]. The principles of synthesis of M–algorithms of estimation and input design for dynamic system identification are given here. Convergence of these multistep algorithms is investigated for the parameter estimation of linear dynamic system. The results of numerical investigations with IDDOL package of M–algorithms are described.

2. The recursive M–algorithms for identification

The idea of using former values of estimated or optimized quantity in building the actual estimation may be realized in two basic forms:

- the introductory filtration of residuals (or their squares) of the output signals of the model and the object with mean values obtained;
- the use of mean values in the number of points given a priori, the pseudo gradients for iterating over the identification quality functional.

Let us describe those two approaches to construction of M–algorithms in parameter identification of the following dynamic object and forecasting model:

$$y_n = X_n^T Q^* + \xi_n, \tag{1}$$

$$\hat{y}_n = \hat{X}_n^T \hat{Q}, \tag{2}$$

where $X_n = [X_n^1, X_n^2, \ldots, X_n^m]^T$, $Q^* = [Q_1, Q_2, \ldots, Q_m]^T$, ξ_n – statistical disturbance sequences, y_n – measurable output signals, $X_n^i (i = 1, \ldots, m)$ – partly measurable generalized inputs (for example, $X_n^i - y_{n-i}$ or u_{n-i} (measurable input) or ξ_{n-i}), $\hat{y}$ – output of the model, $\hat{X}_n = [\hat{X}_n^1, \hat{X}_n^2, \ldots, \hat{X}_n^m]^T$ –

generalized estimated inputs of model, $\hat{Q}$ is the estimation of Q^* which corresponds to minimization of the functional:

$$J(Q) \rightarrow \min_{Q} \tag{3}$$

The criterion $J(Q)$ of identification of Q^* leading to multistep methods can be presented in one of the following forms:
– filtration of the residuals

$$\begin{aligned} J(Q) &= E_N[(\hat{e}_n^2)/2], \\ \hat{e}_n &= H(z^{-1})e_n, \qquad e_n = y_n - \hat{y}_n \text{ (the residual)}, \end{aligned} \tag{4}$$

$$H(z^{-1}) = F(z^{-1})/D(z^{-1}) \quad \text{(the stable filter)},$$

$$D(z^{-1}) = 1 + d_1 z^{-1} + \ldots + d_{k1} z^{-k1},$$

$$F(z^{-1}) = f_0 + f_1 z^{-1} + \ldots + f_{k2} z^{-k2}, (z^{-k} x_n = x_{n-k},\ n = 1, 2, \ldots, N).$$

– filtration of the squares of residuals

$$\begin{aligned} J(Q) &= E_N[(\overline{e}_n)/2], \\ \overline{e}_n &= H(z^{-1})e_n^2, \end{aligned} \tag{5}$$

– mean value in a number of finite points given a priori the pseudo gradients of $J(Q)$ [1, 2, 3, 8]:

$$J(Q) = E_N[(e_n^2)/2], \tag{6a}$$

$$Z_n = [A_0 + \gamma_n^1 A_n] * Z_{n-1} + \gamma_n^2 P_n * S_n, \tag{6b}$$

$$S_n = \text{grad}_Q(e_n^2/2) \quad \text{(the gradient in the point } Q_{n-1}),$$

where A_0, A_n, P_n – known matrices, $\gamma_n^i > 0$ – the scalar sequence ($\gamma_n^i \rightarrow 0$ if $n \rightarrow \infty$), $\{f_i, d_j\}$ – are known coefficients,

$$E_N(\ ) = \begin{cases} E(\ ) - & \text{the expectation operator} \\ \Sigma(\ )/N - & \text{the empirical mean operation.} \end{cases}$$

The multistep pseudo gradient identification algorithms (M–algorithms) with (4)–(6) can be presented in the following form:

$$Q_n = Q_{n-1} - G * \Phi(z^{-1}) * Z_n, \tag{7a}$$

$$Z_n = [A + \gamma_n^1 B_n] * Z_{n-1} + \gamma_n^2 P_n * W_n, \tag{7b}$$

where – W_n – stochastic gradient of $J(Q)$ in (4)–(6), G, A, B_n – matrices of a given order, $\gamma_n^i > 0$ – numerical sequence, $\Phi(z^{-1}) = K(z^{-1})/L(z^{-1})$ – stable filter. In the M–algorithm (7) the form (7a) describes a model of algorithm estimation and (7b) is a certain filter, which produces directions for iteration $\{Q_n\}$.

The object (1) is here assumed to have the structure of

$$A(z^{-1})y_n = B(z^{-1})u_n + \xi_n, \tag{8}$$

where $A(z^{-1}), B(z^{-1})$ are polynomials in the backward shift operator z^{-1}: $A(z^{-1}) = 1 + a_1 z^{-1} + \ldots + a_m z^{-m}$, $B(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} + \ldots + b_l z^{-1}$, $(m \geq 1), y_{-n} = 0 \ (n \geq 0)$, $\{\xi_n\}$ is white noise, $Q^* = [a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_l]^T$ – is a vector of unknown parameters, $X_n^m = [-y_{n-1}, -y_{n-2}, \ldots, -y_{n-m}, u_{n-1}, u_{n-2}, \ldots, u_{n-1}]^T$ is a vector of measured data.

We shall state the following assumptions:

AP1: the system (7a) is asymptotically stable;

AP2: the polynomials $A(z^{-1}), B(z^{-1})$ have no common factors;

AP3: the coefficients in $B(z^{-1})$ are not equal to zero;

AP4: input sequence u_n (may be deterministic or stationary ergodic stochastic) is uniformly bounded and is persistently exciting of required ($\geq m+1$) order;

AP5: ξ_n is sequence of mutually independent random variables with mean zero (independent of $\{u_n\}$) with parameters $E(\xi_n)^2 \geq C_1, E(\xi_n)^4 \leq C_2$, where $C_i > 0$ $(i = 1, 2)$ are constants.

The convergence properties of the algorithm (7a)–(7b) for $W_n = S_n$ and $\gamma_n^1 = \gamma_n^2 = \gamma_n$ can now be summarized in the following results.

THEOREM 1 *Let the sequence $\{Q_n, Z_n\}$ be given by (7a)–(7b) with $G = I$, $A = 0$, $B_n \equiv 0, P_n \equiv I$, conditions AP1–AP5 hold and suppose also that:*

1) $L(z^{-1})$ have all zeros strictly outside of the unit circle in the complex z^{-1} plane and $K(1)/L(1) > 0$;

2) $\gamma_n \longrightarrow 0 \ (n \longrightarrow \infty)$, $\sum_{n=0}^{\infty} \gamma_n = \infty$, $\sum_{n=0}^{\infty} \gamma_n^2 < \infty$.

Then

$$\lim_{n \to \infty} Q_n = Q^* \ (w.p.1).$$

The proof is given in [8].

THEOREM 2 *Let the sequence Q_n, Z_n be given by (7a)–(7b) with $\Phi(z^{-1}) \equiv 1$ conditions AP1–AP5 hold and:*

1) A has all its eigenvalues inside the unit circle;

2) exist $\alpha_1 > 0$ such that for all $n, N_1 > 0$

$$\sum_{n=k}^{k+N1} [G - (I-A)^{-1}GA] P_n X_n X_n^T \geq \alpha_1 * \sum_{n=k}^{k+N1} X_n X_n^T.$$

3)

$$\|B_n\| \leq B < \infty, \quad \|P_n\| \leq P < \infty, \quad \gamma_n \to 0 \ (n \to \infty),$$

$$\sum_{n=0}^{\infty} \gamma_n = \infty, \quad \sum_{n=0}^{\infty} \gamma_n^2 < \infty.$$

Then

$$\lim_{n \to \infty} Q_n = Q^* \ (w.p.1).$$

The proof is given in [8].

The local optimal input design for dynamic system identification.

The approach to building multistep local optimal input design for linear parameter discrete systems is based on the following local optimum principle:

$$\delta\Psi(u_{n-1}, \ldots, u_{n-k}) = \Psi(I_n^D(Q_1^*, u_n)) - \Psi(I_{n-k}^D(Q_1^*, u^{n-k})) \longrightarrow \max_{u_{n-k}^{n-1} \in U} \quad (9)$$

where:

$$u_{n-k}^{n-1} = \{u_{n-k}, \ldots, u_{n-1}\}, \ u^n = \{u_0, u_1, \ldots, u_{n-1}\}, \ I_n^D(Q_1^*, u_n)$$

is a part of an informative Fisher matrix of the system shown which depends only on input signal or parameters of the deterministic part of the object (1), $\Psi(\ )$ scalar convex function (for example logdet ( ), tr( )), U – the set of time constraints for the input. A series of solutions of the problem (9) ($k = 1$) for linear or bilinear variants (one or multi dimensional types) of dynamic processes and function $\Psi(\ )$ can be found, for example, in [4, 5, 6, 7].

Consider a discrete–time linear system of type (1) modeled by the following difference equations (SISO model):

$$A(z^{-1})x_n = B(z^{-1})u_n, \quad (10)$$

$$C(z^{-1})\eta_n = D(z^{-1})\xi_n, \tag{11}$$

$$y_n = x_n + \eta_n, \tag{12}$$

where $C(z^{-1}) = 1 + c_1 z^{-1} + \ldots + c_{k1} z^{-k1}$, $D(\overline{z}^1) = 1 + d_1 \overline{z}^1 + \ldots + d_{k2} z^{-k2}$, $(k_1 \geq k_2)$, $\xi_n \in N(0, \sigma^2)$. The set of constraints U is given in the form:
1) $U^1 = \{\alpha_n \leq u_n \leq \beta_n, \ \ n = 0, 1, \ldots, N\}$;
2) $U^2 = \{\sum_{i=0}^{n} u_i^2 \leq \delta_n, n = 0, 1, \ldots, N\}$.

Then the solution of the problem (9) ($k = 1$) for model (10)–(12) and function $\Psi(\ ) = \text{logdet}(\ )$ (local optimal inputs signals – *LOS*) is given in the following theorem.

THEOREM 3 *Let the conditions AP2–AP3 hold and $A(z^{-1}, C(z^{-1})$ have all zeros strictly outside of the unit circle in the complex z^{-1} plane.*
Then LOS is given by

$$u^*_{n-1} = \left| \begin{array}{ll} gp_{n-1}, & F_{n-1} \geq 0 \\ g1_{n-1}, & F_{n-1} < 0, \end{array} \right. \tag{13}$$

1) for $U^1 - gp_{n-1} = \beta_{n-1}$, $gl_{n-1} = \alpha_{n-1}$;
2) for $U^2 - gp_{n-1} = (\delta_{n-1} - \sum_{i=1}^{n-1} u_i^2)$, $gl_{n-1} = -(\delta_{n-1} - \sum_{i=1}^{n-1} u_i^2)^{1/2}$,
where

$$\begin{aligned}
F_{n-1} &= (2Z_{n-1} * g_{n-1} + \alpha_{n-1} + \beta_{n-1})/(\beta_{n-1} - \alpha_{n-1}) - \lambda_{n-1} * h_{n-1}, \\
Z_{n-1} &= [u^F_{n-2}, \ldots, u^F_{n-1}, -x^F_{n-1}, \ldots, -x^F_{n-m}]^T, \\
g_{n-1} &= [P^{1,2}_{n-1}, \ldots, P^{1,m+1}_{n-1}]^T, \\
h_{n-1} &= P^{1,1}_{n-1}, \qquad \lambda_{n-1} = \sum_{i=1}^{m+k2} \phi_i u^F_{n-i-1} - \sum_{j=1}^{k2} c_j u_{n-j-1}, \\
F(z^{-1}) &= D(z^{-1})A(z^{-1}) = 1 + \phi_1 z^{-1} + \ldots + \phi_{m+k2} z^{-(m+k2)}, \\
x_n &= B(z^{-1})A^{-1}(z^{-1})u_n, \qquad u^F_n = D^{-1}(z^{-1})u_n, \\
x^F_n &= D^{-1}(z^{-1})x_n, \\
P_n &= P_{n-1} - P_{n-1} f_n f_n^T P_{n-1}/(\sigma + f_n^T P_{n-1} f_n), \\
f_n &= [u^F_{n-1}, \ldots, u^F_{n-1}, -x^F_{n-1}, \ldots, -x^F_{n-m}]^T.
\end{aligned}$$

The proof is given in [8].

3. Numerical example

In this section the example of the work package IDDOL [8] for the following process $ARMAX$ type of (1) are reported:

$$y_n + a_1 y_{n-1} + a_2 y_{n-2} = b_1 u_{n-1} + \xi_n + c_1 \xi_{n-1}, \tag{14}$$

- $Q^* = [a_1, a_2, b_1, c_1]^T = [-0.9, 0.2, 1.0, 0.5]^T, \ \ \xi_n \in N(0, \sigma)$;
- $\hat{X}_n = [-y_{n-1}, -y_{n-2}, u_{n-1}, \hat{\xi}_{n-1}]^T$;
- $\hat{\xi}_n = y_n - \hat{X}_n^T Q_{n-1}, \hat{y}_n = \hat{X}_n^T Q_{n-1}$;
- $U = \{-1 \leq u_n \leq 1, n = 1, 2, \ldots, N\}$;
- input signal u_n is LOS in (13) for (14),

where Q_n is the estimation of Q^* in time n with parameters of $M-$ algorithm (6b): $A = 0$, $B_n = 0$, $\gamma_n^1 = 0$, $\gamma_n^2 = 1$, $G = 1$ (unit matrix),

$$K(z^{-1}) = 1, \quad L(z^{-1}) = 1 + 0.5z^{-1}, \quad \sigma = 0.5,$$

$$P_n = P_{n-1} - P_{n-1}\hat{X}_n^T \hat{X}_n P_{n-1}/(1 + \hat{X}_n^T P_{n-1} \hat{X}_n),$$

$$P_0 = I, \quad Q_0 = 0 \quad (n = 1, 2, \ldots, N)$$

The following quantities are reported in Figs. 1–4:

- u_n, y_n, ξ_n and y_n together with ξ_n (Fig.1);
- $W_n = \|Q_n - Q^*\|/\|Q_0 - Q^*\|$, $D_n = \|Q_n\|$, $E_n^1 = [(y_n - \hat{y}_n)/y_n]^2$, $E_n^2 = (y_n - \hat{y})/y_n$ (Fig.2);
- $u_n, y_n, \hat{y}_n$ and y_n together with $\hat{y}_n$ (Fig.3);
- Q_N, Q_0, P_0, N, M–algorithm (Fig.4).

4. Conclusions

For on–line identification of dynamic systems a certain class of multistep adaptive algorithms (7) have been proposed. The sufficient convergence conditions for these methods (7) are obtained for ARX models of type (1). The recursive method (9) is proposed to obtain local optimal input signals (LOS) for dynamic system identification. A solution of the problem (9) for linear models of (1) with $\Psi(\cdot) = \log\det(\cdot)$ can be described. The results of numerical investigations with package IDDOL of $M-$ algorithms are described. A convergence comparison of multistep methods (7) with one–step algorithms is also provided and it is shown that proposed methods may be better than the one-step methods [4, 5, 8]. The simulation examination shows that LOS are quite effective for identification of systems of ARX and $ARMAX$ type of (1) [4, 5, 8].

References

[1] SCHILMAN S.V., YASTREBOV A.I. Convergence of a class of multistep stochastic adaptation algorithms. *Avtomatika i Telemekhanika* (in Russian), **8**, pp. 111–121, 1976

[2] SCHILMAN S.V., YASTREBOV A.I. Properties of a class of multistep gradient and pseudo gradient algorithms of adaptation and learning. *Avtomatika i Telemekhanika* (in Russian), **4**, pp. 95–103, 1978

[3] YASTREBOV A.I., STOCHNIOL A. Multistep adaptive algorithms for dynamic system identification, Abstracts of Papers, International Conference on Systems Science, Wrocław, Poland, pp. 89–90, 1989

[4] YASTREBOV A.I. On the choice of input signal for billinear MIMO dynamic system identification. Abstracts of Papers, International Conference on Stochastic Methods in Experimental Sciences, Szklarska Poręba, Poland, pp.89–91, 1989

[5] YASTREBOV A.I. Control for estimation of parameters of MIMO systems, (in Russian). In "Algebraic and stochastic methods in the applied mathematics", University of Gorky, Gorky, USSR, pp. 97–107, 1988

[6] GOODWIN G.C., PAYNE P. *Dynamic System Identification, Acad. Press,* New York, 1977

[7] KEVICZKY L. "Design of experiments" for the identification of linear systems. *Technometrics,* **17**, 3, pp. 303–315, 1975

[8] YASTREBOV A.I. Synthesis and analysis of multistep adaptive algorithms of parameter identification for static and dynamic objects for discrete data, (in Russian). Technical University of Kielce, Poland, 177 p., 1990

Projektowanie wielokrokowych algorytmów i lokalnie optymalnych sterowań w identyfikacji systemów dynamicznych

W pracy analizuje się zagadnienia identyfikacji systemów dynamicznych na bieżąco. Dla rozwiązania tego zagadnienia proponuje się pewną klasę wielokrokowych metod adaptacyjnych. Podaje się warunek zbieżności metody z prawdopodobieństwem 1. Rozważania są wsparte wynikami analizy komputerowej. Formułuje się zasadę syntezy wielokrokowej sterowań lokalnie optymalnych.

Проектирование многошаговых алгоритмов и локально оптимальных управлений в идентификации динамических систем

В работе анализируется вопрос текущей идентификации динамических систем. Для решения этого вопроса предлагается некоторый класс многошаговых адаптивных методов. Приводится условие сходимости метода с вероятностью 1. Рассмотрение дополняется результатами компьютерного анализа. Формулируется принцип синтеза многошаговых локально оптимальных управлений.

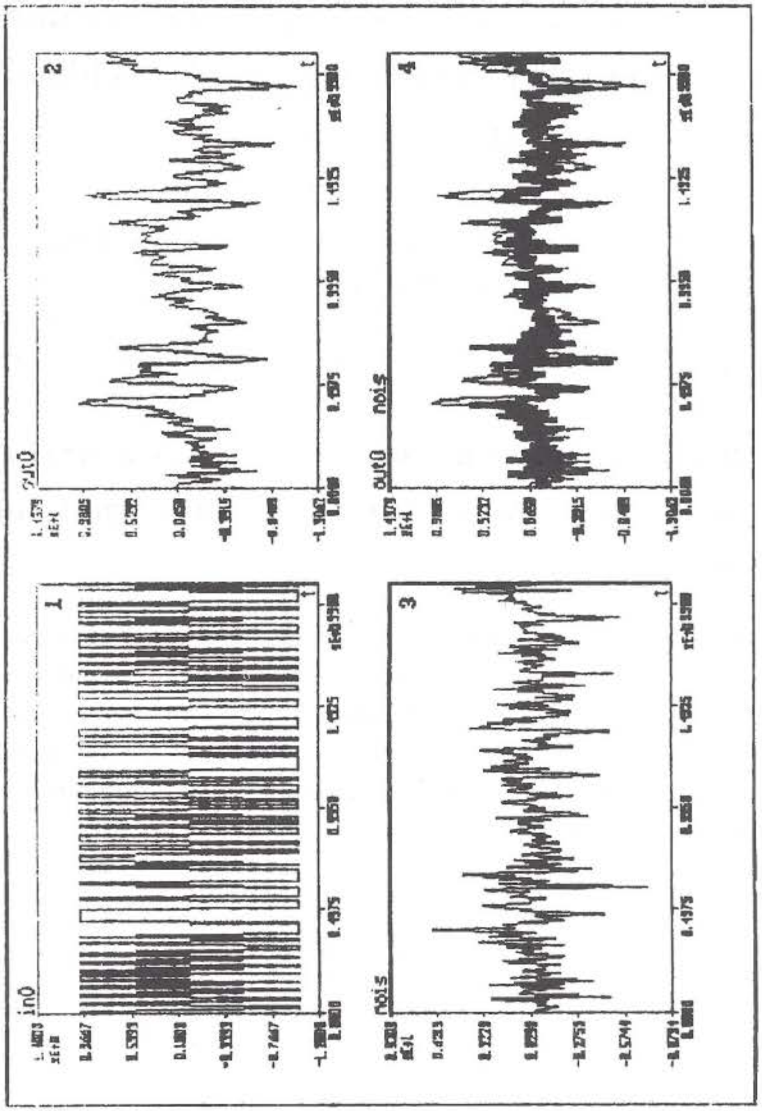

Fig. 1. 1 – Local optimal output signal (in0)
2 – Output signal in time (out0)
3 – Disturbantes in output ($\sigma = 0.05$) (nois)
4 – Comparison of useful signal and disturbances (out0, nois)

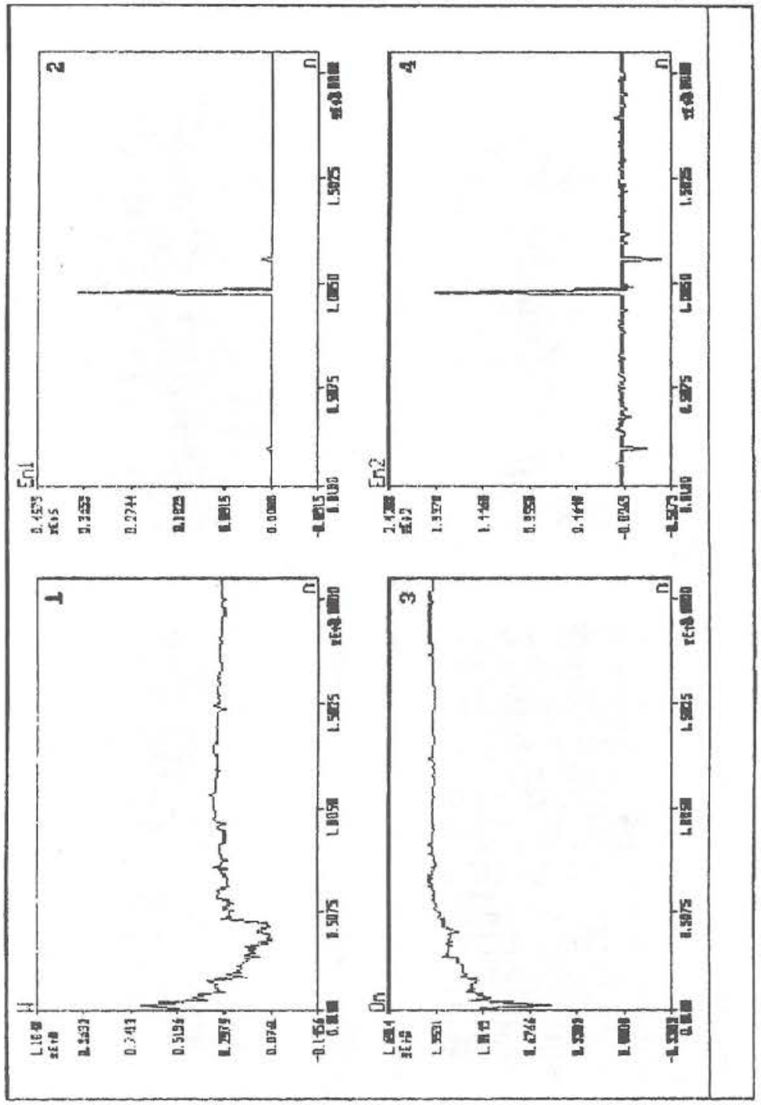

Fig. 2. 1 – Quality criterion of identification process

$$W = \|Q(n) - Q\| / \|Q(0) - Q\| \qquad \text{(W)}$$

2 – Estimation criterion of model quality

$$E_1(n) = (y(n) - y_m(n))^2/(y(n))^2 \qquad (\text{En}_1)$$

3 – Estimation criterion of identification process quality

$$D(n) = \|Q(n)\| \qquad (\text{Q}_n)$$

4 – Estimation criterion of model quality

$$E_2(n) = (y(n) - y_m(n))/(y(n)) \qquad (\text{En}_2)$$

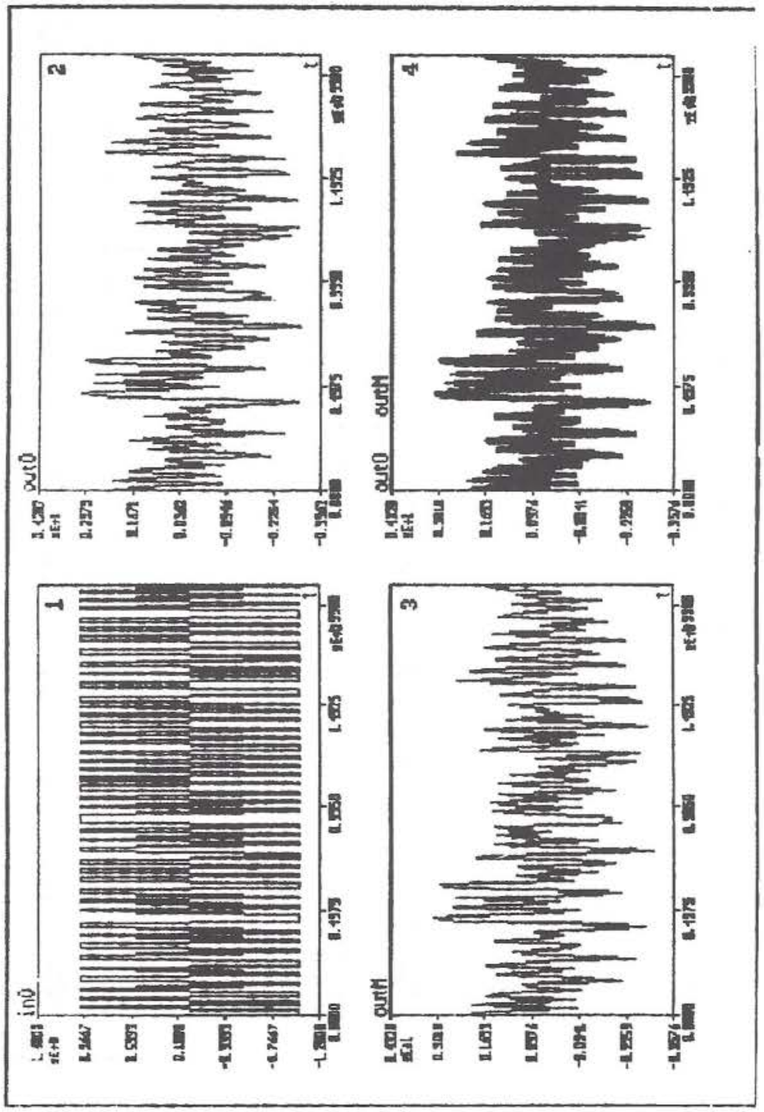

Fig. 3. 1 – Local optimal output signal (in0)
2 – Output signal in time (out0)
3 – Output signal in the model (outM)
4 – Comparison of output signals from the structure and the model (thick line) (out0, outM)

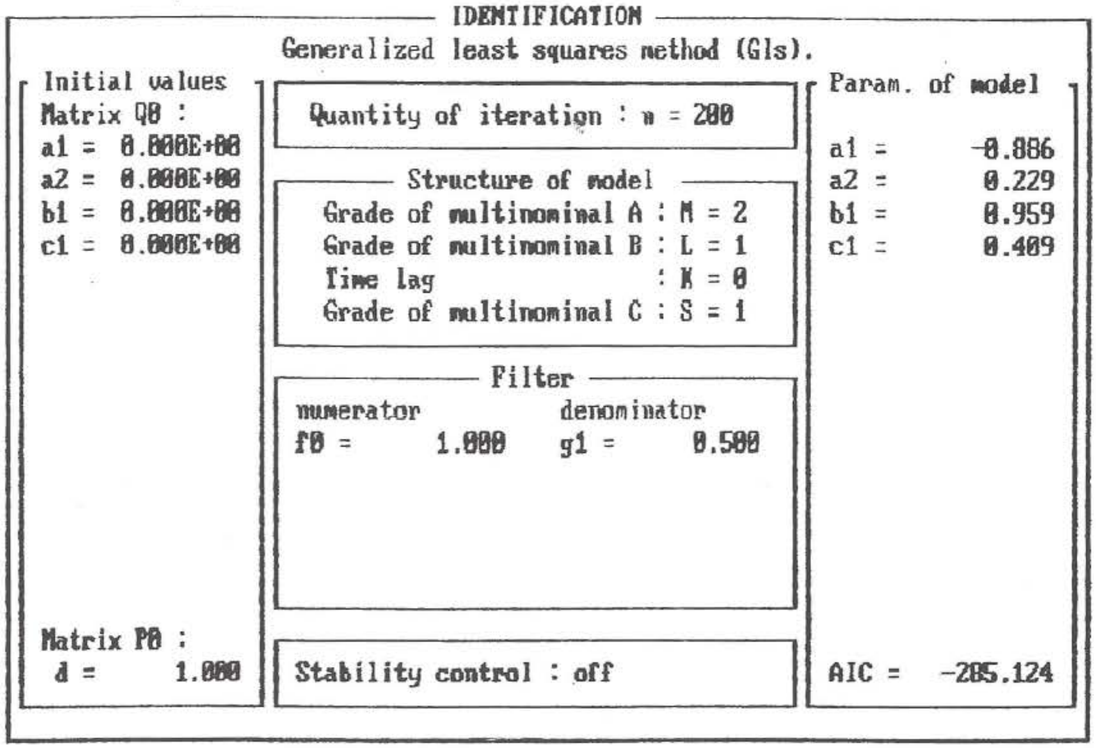

Fig. 4 – Q_N, Q_O, P_O, N, M – algorithm