

DESIGN OF ROBUST RECURSIVE IDENTIFICATION ALGORITHMS FOR LARGE-SCALE STOCHASTIC SYSTEMS

UDC (681.5.037+681.5.15): 681.518.2

Vojislav Filipović, Vladimir Đorđević

University of Kragujevac, Faculty for Mechanical and Civil Engineering in Kraljevo,
Kraljevo, Republic of Serbia

Abstract. *The robust recursive algorithms, for identification of decentralized stochastic systems, are developed. It is supposed that stochastic disturbance belongs to a specified class of distributions which include the gross error model suitable for the description of outliers presence. Such an assumption introduces into the recursive algorithms a nonlinear transformation of prediction error. The given algorithms are robust with respect to uncertainty in the disturbance distribution. The individual subsystems are described with SISO (single-input single output) ARMAX model. Two algorithms are considered: the stochastic approximation and the least squares. Their comparison is based on simulations.*

Key words: *large-scale systems, outliers, Huber's theory, robust estimation, stochastic approximation, least squares*

1. INTRODUCTION

High performance requirement of complex industrial processes increases demands on control systems. The concept of driving a large system by a central computer has become unsuitable for either economic or reliability reasons. A number of large-scale systems founded in the real world are composed of a set of small, interconnected subsystem, such as power systems with strong interactions, water systems which are widely distributed in space, traffic systems with many external signals, large-space flexible structures, digital communication networks and economic systems. It is generally impossible to incorporate many feedback loops into the controller design and is too costly even if they can be implemented. These difficulties motivate the development of decentralized control theory, in which each subsystem is controlled independently [1-5].

Received February 18, 2015

Corresponding author: Vladimir Đorđević

Faculty for Mechanical and Civil Engineering in Kraljevo, Dositejeva 19, 36000 Kragujevac, Republic of Serbia

E-mail: djordjevic.v@mfkv.kg.ac.rs

Within the control theory of complex systems, there are two important approaches:

- i) the first approach is based on the principle of decentralization which allows the decentralization of multivariable systems into a collection of SISO (single-input single-output) or collection of MIMO (multi-input multi-output) systems with lower order than original MIMO system
- ii) the second approach for the control of complex system are a hierarchical control systems[6-7]. This theory is based on decomposition-coordination principle.

The hierarchical identification, inspired by the hierarchical control is presented in [8-10]. Hierarchical identification uses subsystem decomposition in identification, and is, also, called bootstrap identification. In this area a system is decomposed into several subsystems with smaller dimension and fewer variables, and then the parameter vector is identified. The difficulties arise is that there exist common unknown quantities among subsystems. These problems are resolved for deterministic and stochastic systems (with Gaussian noise) in [8-10].

The hierarchical identification has important application in parameter identification of multirate systems [11-12].

The papers [13-14] describe a switching scheme for MIMO systems so that switching in different control agents is carried out in a decentralized way.

In [15] a decentralized adaptive synchronization problem is considered. The topic is based on capability and limitation of the feedback mechanism, traditional adaptive control and complex systems (especially complex networks). In the study of complex adaptive systems, theory plays an important role which mainly focuses on agent-based modelling and simulations. The topic of the [16] covers a kind of common and important phenomenon in nature (for example, chaos synchronization has been found to be useful in secure communication). In recent years, several synchronization-related topics (coordination, rendezvous, consensus, formation, etc.) have become active in the research community.

Two important attributes of large-scale systems are:

- i) they often represent complex, real-life systems
- ii) they have hierarchical (multilevel) and decentralized information structures.

Recent advances of soft computing (fuzzy logic, neural network and evolutionary algorithms) have led us to new approaches for control of complex systems. Connections between two approaches are presented in [17].

In the recursive identification of MIMO systems (hierarchical identification) [8-10] the key assumption about the stochastic disturbance is that disturbance has a Gaussian distribution. In this paper, we supposed that disturbance has non-Gaussian distribution. Justification of such approach was confirmed in practice [18-20]. To be more specific, the real stochastic disturbance rarely, if ever, has Gaussian distribution. Practical investigations show that, in the population of observations, rare large observations (outliers) are present so that their distribution is non-Gaussian. Because, estimation algorithms, based on Gaussian disturbance model, have been found to be inefficient in the above case. Considerable effort has been oriented towards the design of robust estimation algorithm possessing a low sensitivity to distribution changes. The fundamental contribution has been given by Huber [21-22]. The field is known as robust statistics and min-max robust estimation. Further development of this idea and application to different types of problems has led to many achievements. The applications in system identification are given in [23-25] and in the field of adaptive control in [26-27].

A problem with stochastic description of the disturbance is that the probability distribution of a disturbance must be known a priori. Such assumption is unrealistic. In the robust identification [23-25] (or adaptive control [26-27]) it is supposed that a priori known only a class of distributions to which belongs the disturbance.

In this paper, we will consider robust identification of MIMO systems described with ARMAX model. The MIMO system is decomposed into a collection of SISO subsystems. It is supposed that stochastic disturbance has a non-Gaussian distribution. Using prediction errors concept two algorithms are derived: stochastic approximation algorithm and least squares algorithm.

Applications of decentralization in the field of adaptive control, for the Gaussian noise, are considered in [28] where the reference decentralized indirect adaptive control is considered. Reference [29] describes the case when stochastic disturbance has a non-Gaussian distribution and interconnections of subsystems are nonlinear functions. For estimation of unknown parameters of the local controllers the stochastic approximation type algorithms are used. Using martingale theory, it is shown that the closed-loop system is globally stable and that the overall mean-square tracking error is bounded. In reference [30] the decentralized stochastic minimum variance controller, for the MIMO subsystems, is considered.

In this paper analysis of recursive algorithms for identification of large scale systems is based on simulations. The expected conclusion is that the convergence speed of stochastic approximation (SA) algorithm is significantly lower than the convergence speed of least squares. In the future we will consider SA with extended memory. Such intervention significantly increases the convergence speed of the SA algorithm while preserving the simplicity of the algorithm.

2. PROBLEM FORMULATION

The overall system consists of N subsystems (see Fig. 1).

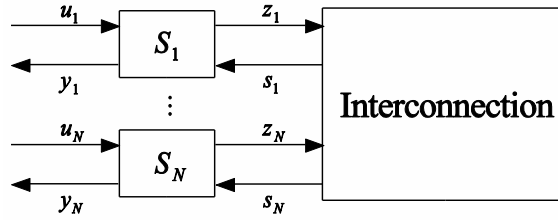


Fig. 1 Interconnected systems

In Fig. 1 u_i , s_i , y_i , $i = (1, 2, \dots, N)$ are control inputs, interconnection inputs, control outputs and interconnection outputs, respectively.

Let us consider the large-scale system S , which consists of N interconnected stochastic systems S_1, S_2, \dots, S_N . Each interconnected system S_i , $1 \leq i \leq N$, can be described in the next form.

$$S_i : A_i(q^{-1})y_i(k) = B_i(q^{-1})u_i(k) + C_i(q^{-1})e_i(k) + \sum_{\substack{j=1 \\ j \neq i}}^N B_{ij}(q^{-1})u_j(k) + \sum_{\substack{j=1 \\ j \neq i}}^N A_{ij}(q^{-1})y_j(k), \quad i \in \mathcal{N} \quad (1)$$

where

$$\begin{aligned} A_i(q^{-1}) &= 1 + a_{i,1}q^{-1} + \dots + a_{i,nA_i}q^{-nA_i} \\ B_i(q^{-1}) &= b_{i,1}q^{-1} + \dots + b_{i,nB_i}q^{-nB_i} \\ C_i(q^{-1}) &= 1 + c_{i,1}q^{-1} + \dots + c_{i,nC_i}q^{-nC_i} \\ B_{ij}(q^{-1}) &= b_{ij,1}q^{-1} + \dots + b_{ij,nB_{ij}}q^{-nB_{ij}} \\ A_{ij}(q^{-1}) &= a_{ij,0} + a_{ij,1}q^{-1} + \dots + a_{ij,nA_{ij}}q^{-nA_{ij}} \\ a_{if,0} &= \begin{cases} 1, & f = i \\ 0, & f \neq i \end{cases} \end{aligned} \quad (2)$$

and where nA_i is the degree of polynomial $A_i(q^{-1})$, and nB_i , nC_i , nB_{ij} and nA_{ij} are degrees of corresponding polynomials. The block scheme of model (1) is given on the Fig. 2.

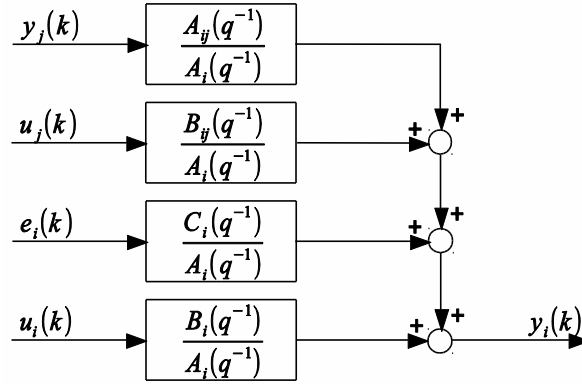


Fig. 2 Structure of subsystem

For the system presented on Fig. 2 ..., $\{y_i(k)\}$ and $\{e_i(k)\}$ are input, output and stochastic disturbance of the i -th subsystem S_i , $(i,j) \in \mathcal{N}$, $j \neq i$.

Remark 1. In references [31-32] more general model than (1) is considered. The form of model is

$$\begin{aligned} S_i : A_i(q^{-1})y_i(k) &= B_i(q^{-1})u_i(k) + C_i(q^{-1})e_i(k) + D_i(q^{-1})v_i(k) \\ P_i(q^{-1})x_i(k) &= Q_i(q^{-1})y_i(k) \quad , \quad i \in \mathcal{N} \end{aligned} \quad (3)$$

Sequences $\{v_i(k)\}$ and $\{x_i(k)\}$ are the interaction inputs and outputs of S_i from and to S_j

$$v_i = f_i(k, \mathbf{x}), \quad i \in \mathcal{N} \quad (4)$$

where $\mathbf{x}^T = [x_1, \dots, x_N]^T$.

The function $f_i(k, \mathbf{x})$ may contain the non-linearities of the i -th subsystem and non-linear interactions with other subsystems. Polynomials $A_i(q^{-1})$, $B_i(q^{-1})$ and $C_i(q^{-1})$ are the same as in relation (1) and polynomials $D_i(q^{-1})$, $P_i(q^{-1})$ and $Q_i(q^{-1})$ have a form

$$\begin{aligned} D_i(q^{-1}) &= 1 + d_{i,1}q^{-1} + \dots + d_{i,n}D_iq^{-nD_i} \\ P_i(q^{-1}) &= 1 + p_{i,1}q^{-1} + \dots + p_{i,n}P_iq^{-nP_i} \\ Q_i(q^{-1}) &= 1 + q_{i,1}q^{-1} + \dots + q_{i,n}Q_iq^{-nQ_i} \end{aligned} \quad (5)$$

The overall system S , which is composed of subsystems S_i interconnected as in (4), can be described as

$$S : \mathbf{A}(q^{-1})\mathbf{y}(k) = \mathbf{B}(q^{-1})\mathbf{u}(k) + \mathbf{C}(q^{-1})\mathbf{e}(k) + \mathbf{D}(q^{-1})\mathbf{v}(k) \quad (6)$$

where $\mathbf{y}(k) = [y_1, \dots, y_N]^T$, $\mathbf{u}(k) = [u_1, \dots, u_N]^T$ and $\mathbf{e}(k) = [e_1, \dots, e_N]^T$ are output, input and disturbance. The corresponding matrix polynomials are given by the following relations

$$\begin{aligned} \mathbf{A}(q^{-1}) &= \text{diag}\{A_1(q^{-1}), \dots, A_N(q^{-1})\} \\ \mathbf{B}(q^{-1}) &= \text{diag}\{B_1(q^{-1}), \dots, B_N(q^{-1})\} \\ \mathbf{C}(q^{-1}) &= \text{diag}\{C_1(q^{-1}), \dots, C_N(q^{-1})\} \\ \mathbf{D}(q^{-1}) &= \text{diag}\{D_1(q^{-1}), \dots, D_N(q^{-1})\} \\ \mathbf{P}(q^{-1}) &= \text{diag}\{P_1(q^{-1}), \dots, P_N(q^{-1})\} \\ \mathbf{Q}(q^{-1}) &= \text{diag}\{Q_1(q^{-1}), \dots, Q_N(q^{-1})\} \end{aligned} \quad (7)$$

The vector $\mathbf{v}(k) \in R^N$, $\mathbf{v}(k) = [v_1, \dots, v_N]^T$ and $\mathbf{x}(k) \in R^N$ are interconnected as

$$\mathbf{v}(k) = \mathbf{f}(k, \mathbf{x}(k)) \quad (8)$$

The nonlinear function $\mathbf{f} : R \times R^N \rightarrow R^N$ is continuous and bounded in both arguments. Also, there exist non-negative, but unknown numbers such that ξ_{ij}

$$|\mathbf{f}_i(k, \mathbf{x}(k))| \leq \sum_{j=1}^N \xi_{ij} |x_j(k)| \quad (9)$$

The composite system (1) can be described in the form of ARMAX model

$$\mathbf{A}(q^{-1})\mathbf{y}(k) = \mathbf{B}(q^{-1})\mathbf{u}(k) + \mathbf{C}(q^{-1})\mathbf{e}(k) \quad (10)$$

where polynomial matrices $\mathbf{A}(q^{-1})$, $\mathbf{B}(q^{-1})$, and $\mathbf{C}(q^{-1})$ have a form

$$\mathbf{A}(q^{-1}) = \begin{bmatrix} A_{11}(q^{-1}) & -A_{12}(q^{-1}) & \dots & -A_{1N}(q^{-1}) \\ -A_{21}(q^{-1}) & A_{22}(q^{-1}) & \dots & -A_{2N}(q^{-1}) \\ -A_{N1}(q^{-1}) & \dots & -A_{NN-1}(q^{-1}) & A_{NN}(q^{-1}) \end{bmatrix} \quad (11)$$

$$\mathbf{B}(q^{-1}) = \begin{bmatrix} B_{11}(q^{-1}) & B_{12}(q^{-1}) & \cdots & B_{1N}(q^{-1}) \\ B_{21}(q^{-1}) & B_{22}(q^{-1}) & \cdots & B_{2N}(q^{-1}) \\ B_{N1}(q^{-1}) & \cdots & B_{NN-1}(q^{-1}) & B_{NN}(q^{-1}) \end{bmatrix} \quad (12)$$

$$\mathbf{C}(q^{-1}) = \begin{bmatrix} C_1(q^{-1}) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & C_N(q^{-1}) \end{bmatrix} \quad (13)$$

The i -th interconnected system output variable in the equation (1) can be expressed as

$$\begin{aligned} y_i(k) = & -a_{i,1}y_i(k-1) - \dots - a_i nA_i y_i(k-nA_i) + \\ & + b_{i,1}u_i(k-1) + \dots + b_i nB_i u_i(k-nB_i) + \\ & + e_i(k) + c_{i,1}e_i(k-1) + \dots + c_i nC_i e_i(k-nC_i) + \\ & + b_{ij,1}(k)u_j(k-1) + \dots + b_{ij} nB_{ij} u_j(k-nB_{ij}) + \\ & + a_{ij,1}(k)y_j(k-1) + \dots + a_{ij} nA_{ij} y_j(k-nA_{ij}) \end{aligned} \quad (14)$$

with $j=1,2,\dots,N$ and $j \neq i$.

The output $y_i(k)$ of the interconnected stochastic system S_i can be expressed in the following compact form

$$y_i(k) = \boldsymbol{\theta}_i^T \boldsymbol{\varphi}_i^o(k) + e_i(k) \quad (15)$$

where

$$\boldsymbol{\theta}_i^T = [a_{i,1}, \dots, a_{i,nA_i}, b_{i,1}, \dots, b_{i,nB_i}, c_{i,1}, \dots, c_{i,nC_i}, b_{ij,1}, \dots, b_{ij,nB_{ij}}, a_{if,1}, \dots, a_{if,nA_{if}}] \quad (16)$$

$$\begin{aligned} \boldsymbol{\varphi}_i^o(k) = & [-y_i(k-1), \dots, -y_i(k-nA_i), u_i(k-1), \dots, u_i(k-nB_i), e_i(k-1), \dots, \\ & e_i(k-nC_i), u_i(k-1), \dots, u_i(k-nB_{ij}), y_j(k-1), \dots, y_i(k-nA_{ij})]^T \end{aligned} \quad (17)$$

In relation (10) values $e(k-i)$, $i=1,2,\dots, nB_{ij}$ are not measurable and because the following vector is introduced

$$\begin{aligned} \boldsymbol{\varphi}_i(k) = & [-y_i(k-1), \dots, -y_i(k-nA_i), u_i(k-1), \dots, u_i(k-nB_i), \varepsilon_i(k-1), \dots, \\ & \varepsilon_i(k-nC_i), u_i(k-1), \dots, u_i(k-nB_{ij}), y_j(k-1), \dots, y_i(k-nA_{ij})]^T \end{aligned} \quad (18)$$

where

$$\varepsilon_i(k) = y_i(k) - \hat{\boldsymbol{\theta}}_i^T(k-1)\boldsymbol{\varphi}_i(k) \quad (19)$$

represents the prediction error, and is a base for the formulation of the generator (functional) for a recursive algorithm.

3. ROBUST RECURSIVE ALGORITHM

Dominant approach in the literature devoted to the identification of stochastic systems is based on assumption that the probability distribution of stochastic disturbance is exactly known (usually it is assumed that it have Gaussian distribution). In this way, an absolutely optimal recursive algorithm can be obtained [24]. Analysis of real data, however, is shown that the population measurements also contain outliers (measurements inconsistent with the majority of the population measurements) [18-20]. The presence of outliers significantly degrades performances of linear recursive algorithms. A mathematical theory (robust statistics [21-22]) was developed and provides a framework for the efficient solution of the problem described above. For stochastic disturbance it is assumed that a priori is known the class of distributions to which disturbance belongs. In this paper, is considered the class of approximately normal distributions

$$P_\varepsilon = \{P : P = (1-\varepsilon)N(0, \sigma_1^2) + \varepsilon N(0, \sigma_2^2)\}, \quad \sigma_1^2 \ll \sigma_2^2 \quad (20)$$

where $\varepsilon \in [0,1)$ is the degree of contamination and $N(0, \sigma^2)$ denotes the Gaussian distribution with zero mean and variance σ^2 . The class of distributions (9) is suitable to describe the presence of outliers in the systems output measures.

For the design of robust algorithms the generalised maximum likelihood method is used. According to the philosophy of robust statistics, there is, for a given class of distributions, the least favourable distribution. For the i -th subsystem and the class of distributions (20) the least favourable probability distribution is

$$p_i^*(e_i(k)) = \begin{cases} \frac{1-\varepsilon}{\sqrt{2\pi}\sigma_{i1}} \exp\left\{-\frac{(e_i(k))^2}{2\sigma_{i1}^2}\right\}, & |e_i(k)| \leq k_\varepsilon \\ \frac{1-\varepsilon}{\sqrt{2\pi}\sigma_{i1}} \exp\left\{-\frac{k_\varepsilon}{\sigma_{i1}^2} \left(-|e_i(k)| - \frac{k_\varepsilon}{2}\right)\right\}, & |e_i(k)| > k_\varepsilon \end{cases} \quad (21)$$

The connection between parameters ε and k_ε is [21]

$$\frac{2\Phi_N(k_\varepsilon)}{k_\varepsilon} - 2\Phi_N(-k_\varepsilon) = \frac{\varepsilon}{1-\varepsilon}, \quad \Phi_N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy \quad (22)$$

$$\Phi_i(\cdot) = -\log p_i^*(\cdot) \quad (23)$$

By using relations (19) and (23), empiric functional for i -th subsystem is obtained

$$J_i^k(\boldsymbol{\theta}) = \frac{1}{k} \sum_{j=1}^k \Phi_i(\varepsilon_i(j)) \quad (24)$$

The recursive minimization of such a criterion can be done by using the approximate Newton Raphson type method [24-25]

$$\begin{aligned} \hat{\boldsymbol{\theta}}_i(k) &= \hat{\boldsymbol{\theta}}_i(k-1) - [k \nabla_{\boldsymbol{\theta}}^2 J_i^k(\hat{\boldsymbol{\theta}}_i(k-1))]^{-1} [k \nabla_{\boldsymbol{\theta}} J_i^k(\hat{\boldsymbol{\theta}}_i(k-1))] \\ \hat{\boldsymbol{\theta}}_i(0) &= \boldsymbol{\theta}_{i0} \end{aligned} \quad (25)$$

Moreover, with large k and by virtue of the approximation truth of the optimality conditions, yielding $\nabla_{\theta} J_{i-1}^k(\hat{\boldsymbol{\theta}}_i) \approx 0$ and we have

$$k \nabla_{\theta} J_i^k(\hat{\boldsymbol{\theta}}_i(k-1)) = \boldsymbol{\varphi}_i(k) \psi_i(\varepsilon_i(k)) \quad (26)$$

$$k \nabla_{\theta} J_i^k(\hat{\boldsymbol{\theta}}_i(k-1)) = \alpha_i \sum_{j=1}^k \boldsymbol{\varphi}_i(j) \boldsymbol{\varphi}_i^T(j), \quad \psi_i(\cdot) = (\Phi_i^*)'(\cdot) \quad (27)$$

where $\psi_i(\cdot) = (\Phi_i^*)'(\cdot)$ and $\alpha_i = E\{\psi_i'(\varepsilon_i(k))\}$. By introducing

$$\mathbf{P}_i(k) = [k \nabla_{\theta}^2 J_i^k(\hat{\boldsymbol{\theta}}_i(k-1))]^{-1} \quad (28)$$

and applying the matrix inversion lemma, we have

$$\hat{\boldsymbol{\theta}}_i(k) = \hat{\boldsymbol{\theta}}_i(k-1) + \mathbf{P}_i(k) \boldsymbol{\varphi}_i(k) \psi_i(\varepsilon_i(k)) \quad (29)$$

$$\mathbf{P}_i(k) = \mathbf{P}_i(k-1) - \frac{\mathbf{P}_i(k-1) \boldsymbol{\varphi}_i(k) \boldsymbol{\varphi}_i^T(k) \mathbf{P}_i(k-1)}{\alpha_i^{-1} + \boldsymbol{\varphi}_i^T(k) \mathbf{P}_i(k-1) \boldsymbol{\varphi}_i(k)} \quad (30)$$

$$\hat{\boldsymbol{\theta}}_i(0) = \boldsymbol{\theta}_{i0}, \quad \mathbf{P}_i(0) = \gamma_i \mathbf{I}, \quad \gamma_i \in (0, \infty) \quad (31)$$

The algorithm (29)-(31) represents a robust version of the conventional extended least squares.

Remark 2. In [24] it is shown that

$$\alpha_i = E\{\psi_i'(x)\} = I_i(p_i^*(e_i(k))) \quad (32)$$

where $I_i(p_i^*)$ is a Fisher information for the least favorable pdf. When the gross error model has a form

$$P_i = (1 - \varepsilon_i) N(0, \sigma_1^2) + \varepsilon_i N(0, \sigma_2^2), \quad \sigma_2^2 \gg \sigma_1^2 \quad (33)$$

where ε_i is the degree of contamination and $N(0, \sigma^2)$ is normal distribution with mean zero and variance σ^2 , the Fisher information is

$$I_i(p_i^*) = \frac{1}{(1 - \varepsilon_i) \sigma_1^2 + \varepsilon_i \sigma_2^2} \quad (34)$$

Remark 3. For the class of distribution (21) the nonlinear transformation of prediction error in relation (29) $\psi_i(\cdot)$ is Huber's function

$$\psi_i(x) = \max\{-k_{\varepsilon_i}, \min(k_{\varepsilon_i}, x)\} \quad (35)$$

and its derivative

$$\psi_i'(x) = \begin{cases} 1 & , \quad |x| < k_{\varepsilon_i} \\ 0 & , \quad \text{otherwise} \end{cases} \quad (36)$$

The graphical representation of function (35) and (36) is given on the next figure

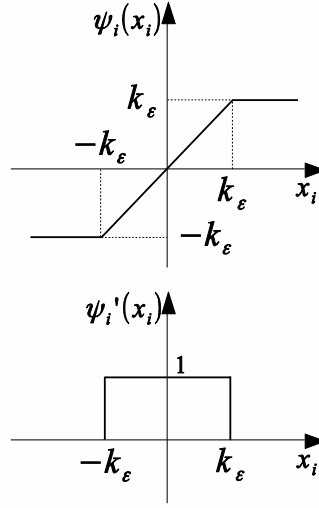


Fig. 3 Nonlinear function of prediction errors

Introducing value $r_i(k) = \text{tr}\mathbf{P}_i(k)$, from the least squares (29)-(31) the stochastic approximation algorithm is obtained.

$$\hat{\boldsymbol{\theta}}_i(k) = \hat{\boldsymbol{\theta}}_i(k-1) + \frac{1}{r_i(k)} \boldsymbol{\varphi}_i(k) \psi_i(\varepsilon_i(k)), \quad \hat{\boldsymbol{\theta}}_i(0) = \mathbf{0} \quad (37)$$

$$r_i(k) = r_i(k-1) + \alpha_i \boldsymbol{\varphi}_i^T(k) \boldsymbol{\varphi}_i(k), \quad r_i(0) = 1 \quad (38)$$

4. SIMULATIONS

In simulations is used following decentralised ARMAX model

$$\begin{aligned} y_1(k) &= -0.85y_1(k-1) + 0.28u_1(k-1) + e_1(k) + 0.14e_1(k-1) + 0.18u_2(k-1) + 0.26y_2(k-1) \\ y_2(k) &= -0.84y_2(k-1) + 0.36u_2(k-1) + e_2(k) + 0.2e_2(k-1) + 0.16u_1(k-1) - 0.25y_1(k-1) \end{aligned} \quad (39)$$

where $e_1(k)$ and $e_2(k)$ are stochastic disturbances

$$\begin{aligned} e_1(k) &\sim (1-\varepsilon)\mathcal{N}(0, 0.25) + \varepsilon\mathcal{N}(0, 25) \\ e_2(k) &\sim (1-\varepsilon)\mathcal{N}(0, 0.25) + \varepsilon\mathcal{N}(0, 12.25) \end{aligned} \quad (40)$$

In simulations, $k_\varepsilon = 3$, while the contamination ε has values 0.05, 0.1, 0.15, and 0.2. The error of estimation is calculated as follows

$$\delta = \ln \frac{\|\hat{\boldsymbol{\theta}}(k) - \boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|} \quad (41)$$

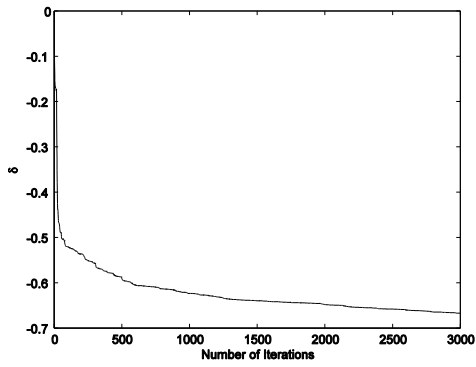


Fig. 4 Error of SA algorithm for $\varepsilon = 0.05$

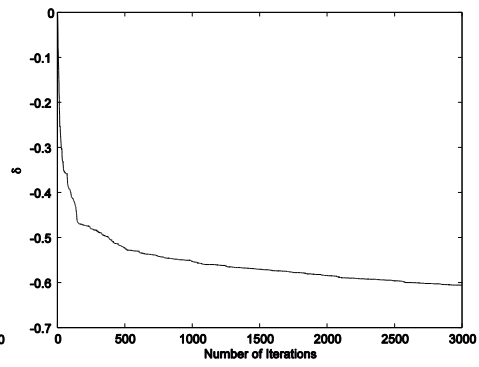


Fig. 5 Error of SA algorithm for $\varepsilon = 0.1$

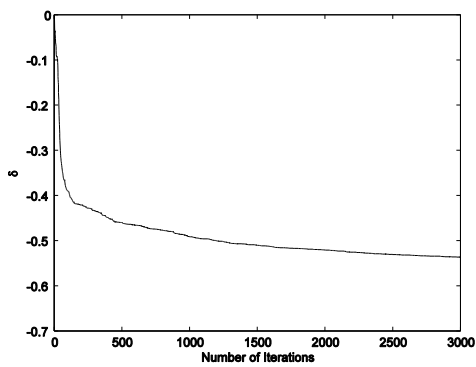


Fig. 6 Error of SA algorithm for $\varepsilon = 0.15$

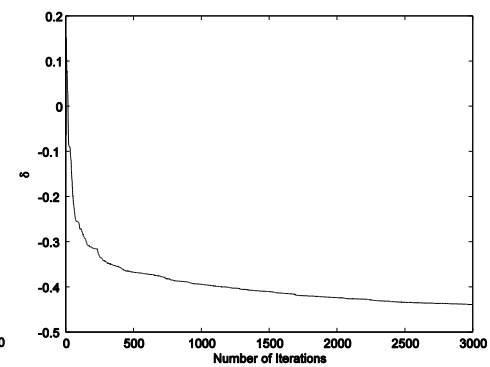


Fig. 7 Error of SA algorithm for $\varepsilon = 0.2$

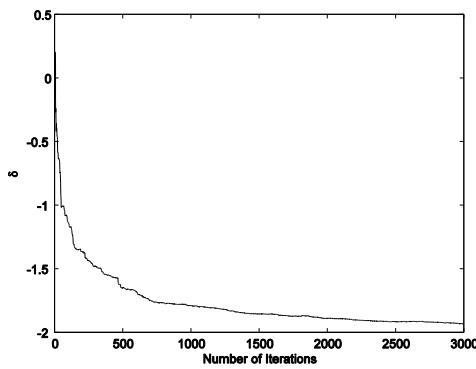


Fig. 8 Error of LS algorithm for $\varepsilon = 0.05$

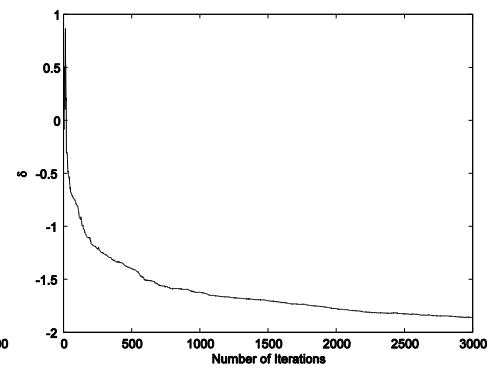


Fig. 9 Error of LS algorithm for $\varepsilon = 0.1$

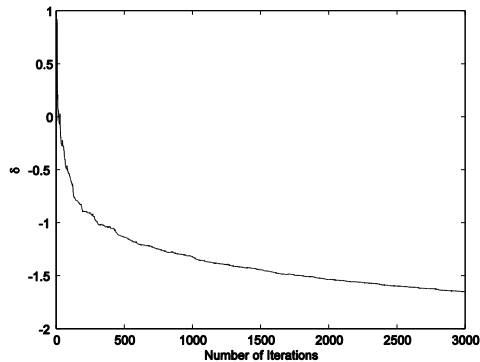


Fig. 10 Error of LS algorithm for $\varepsilon = 0.15$

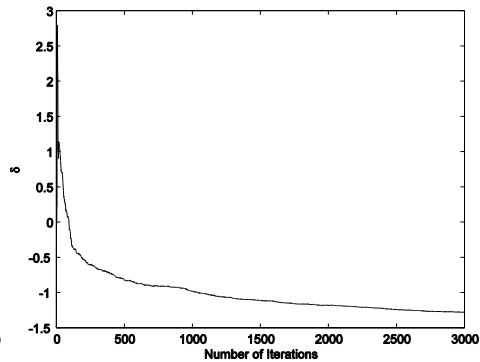


Fig. 11 Error of LS algorithm for $\varepsilon = 0.2$

The Fig. 4 - Fig. 7 presents the behaviour of robust stochastic approximation algorithm for different degrees of contamination. It follows that the estimation error increases with the increase of degree of contamination.

Similar simulations, Fig. 8 - Fig. 11, are presented for robust least squares. It is possible to conclude that robust least squares algorithm is more superior than stochastic approximation algorithm.

CONCLUSION

In the paper has been considered robust identification of decentralized systems. Two algorithms have been proposed: the algorithm of stochastic approximation and the extended least squares algorithms. Simulations show that the stochastic approximation algorithm has smaller convergence speed in comparison with the least squares algorithm. Further investigations will be devoted for increasing speed of convergence of stochastic approximation. The candidates for that are recursive algorithms with extended memory.

Acknowledgement: *The paper is a part of the research done within projects TR33026 and TR33027 at the University of Kragujevac, Faculty of Mechanical and Civil Engineering in Kraljevo, and was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.*

REFERENCES

- [1] D. Siljak, *Large-Scale Dynamic Systems. Stability and Structure*, North-Holland, Amsterdam, 1978.
- [2] D. Siljak, *Decentralized Control of Complex Systems*, Academic Press, New York.
- [3] A.I. Zecevic, *Control of Complex Systems*, Springer, Berlin, 2010.
- [4] J. Lunze, *Feedback Control of Large-Scale Systems*, Prentice-Hall, New Jersey, 1992.
- [5] M. S. Mahmoud, *Distributed Control and Filtering for Industrial Systems*, The Institute of Engineering and Technology, London, 2013.
- [6] M. D. Mesarovic, M. Macko, F. Takahara, *Theory of Hierarchical Multilevel Systems*, Academic Press, New York, 1970.
- [7] M. G. Singh, *Dynamical Hierarchical Control*, North-Holland, Amsterdam, 1980.

- [8] F. Ding, T. Chen, "Hierarchical Gradient Based Identification of Multivariable Discrete-Time Systems", *Automatica* vol. 41, no. 1, pp. 315–325, 2005. [Online]. Available: <http://dx.doi.org/10.1016/j.automatica.2004.10.010>
- [9] F. Ding, T. Chen, "Hierarchical Least Squares Identification Methods for Multivariable Systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 397–402, 2005. [Online]. Available: <http://dx.doi.org/10.1109/TAC.2005.843856>
- [10] Y. Liu, F. Ding, Y. Shi, "An Efficient Hierarchical Identification Method for General Dual-Rate Sampled-Data Systems," *Automatica*, vol. 50, no. 4, pp. 962–970, 2014. [Online]. Available: <http://dx.doi.org/10.1016/j.automatica.2013.12.025>
- [11] L. Qui, T. Chen, " H^2 -Optimal Design of Multirate Sampled-Data Systems," *IEEE Transactions on Automatic Control*, vol. 39, no. 12, pp. 2506–2511, 1994. [Online]. Available: <http://dx.doi.org/10.1109/9.362836>
- [12] L. Qui, T. Chen, "Multirate Sampled-Data Systems: All H^2 -Suboptimal Controllers and the Minimum Entropy Controllers," *IEEE Transactions on Automatic Control*, vol. 44, no. 3, pp. 537–550, 1999. [Online]. Available: <http://dx.doi.org/10.1109/9.751347>
- [13] A. G. Aghdam, E. J. Davison, "Pseudo-Decentralized Switching Control," *Automatica*, vol. 39, no. 2, pp. 317–327, 2003. [Online]. Available: [http://dx.doi.org/10.1016/S0005-1098\(02\)00208-X](http://dx.doi.org/10.1016/S0005-1098(02)00208-X)
- [14] A. G. Aghdam, E. J. Davison, "Decentralized Switching Control for Hierarchical Systems," *Automatica*, vol. 43, no. 3, pp. 1092–1100, 2007. [Online]. Available: <http://dx.doi.org/10.1016/j.automatica.2006.12.005>
- [15] H. Ma, "Decentralized Adaptive Synchronization of a Stochastic Discrete-Time Multiagent Dynamic Model," *SIAM Journal of Control and Optimization*, vol. 48, no. 2, pp. 559–580, 2009.
- [16] J. H. Holland, *Hidden Order: How Adaption Build Complexity*, Addison – Wesley, New York, 1996
- [17] M. Jamshidi, *Large-Scale Systems. Modeling, Control and Fuzzy Logic*, Prentice-Hall, New Jersey, 1997.
- [18] V. Barnett and T. Lewis, *Outliers in Statistical Data*, 3rd Ed., Wiley-Blackwell, New York, 1994.
- [19] K. Pearson, *Mining Imperfect Data*, SIAM, Philadelphia, 2005.
- [20] K. Pearson, *Exploring Data in Engineering, the Science and Medicine*, Oxford University Press, Oxford, 2011.
- [21] P. Huber, "Robust Estimation of a Location Parameter," *Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964. [Online]. Available: <http://dx.doi.org/10.1214/aoms/1177703732>
- [22] P. Huber, E. Ronchetti, *Robust Statistics*, Wiley, London, 2009.
- [23] B. Poljak, Y. Z. Tsytkin, "Adaptive Estimation Algorithms: Convergence, Optimality, Stability", *Automation and Remote Control*, vol. 43, no.3, pp. 71–81, 1979.
- [24] Y. Tsytkin, *On the Foundation of Information and Identification (in Russian)*, Nauka, Moscow, 1984.
- [25] V. Filipovic, B. Kovacevic, "On Robust AML Identification Algorithm," *Automatica*, vol. 30, no. 11, pp. 1775–1778, 1994. [Online]. Available: [http://dx.doi.org/10.1016/0005-1098\(94\)90081-7](http://dx.doi.org/10.1016/0005-1098(94)90081-7)
- [26] V. Filipovic, B. Kovacevic, "On Robustified Adaptive Minimum Variance Controller," *International Journal of Control*, vol. 65, no. 1, pp. 117–129, 1996. [Online]. Available: <http://dx.doi.org/10.1080/00207179608921689>
- [27] V. Filipovic, "Stochastic Multivariable Self-Tuning Tracker for Non-Gaussian Systems," *International Journal of Applied Mathematics and Computer Science*, vol. 15, no. 3, pp. 351–357, 2005.
- [28] S. Kamoun, "Design of Optimal Self-Tuning Regulators for Large-Scale Stochastic Systems," *International Journal of Information and Systems Science*, vol. 6, no. 2, pp. 201–219, 2010.
- [29] V. Filipovic, "Decentralized Stochastic Adaptive Control of Dynamic non-Gaussian Systems," in *Proceedings of the European Control Conference*, Porto, Portugal, pp. 2016–2021, 2001.
- [30] V. Filipovic, "Decentralized Stochastic Minimum Variance Controller," in *Proceedings of 10th IEEE Mediterranean Conference on Control and Automation (MED'02)*, July 9–13, Lisabon, Portugal, 2002.
- [31] D. T. Gavel, D. D. Siljak, "Decentralized Adaptive Control: Structural Conditions for Stability," *IEEE Transactions on Automatic Control*, vol. 34, no. 4, pp. 413–426, 1989. [Online]. Available: <http://dx.doi.org/10.1109/9.28016>
- [32] M. S. Radenkovic, "Decentralized Stochastic Adaptive Control of Interconnected Systems," *International Journal of Adaptive Control and Signal Processing*, vol. 4, no. 1, pp. 1–16, 1990. [Online]. Available: <http://dx.doi.org/10.1002/acs.4480040102>