

This paper addresses the task of identifying the parameters of a linear object in the presence of non-Gaussian interference. The identification algorithm is a gradient procedure for minimizing the combined functional. The combined functional, in turn, consists of the fourth-degree functional and a modular functional, whose weights are set using a mixing parameter. Such a combination of functionals makes it possible to obtain estimates that demonstrate robust properties. We have determined the conditions for the convergence of the applied procedure in the mean and root-mean-square measurements in the presence of non-Gaussian interference. In addition, expressions have been obtained to determine the optimal values of the algorithm's parameters, which ensure its maximum convergence rate. Based on the estimates obtained, the asymptotic and non-asymptotic values of errors in estimating the parameters and identification errors. Because the resulting expressions contain a series of unknown parameters (the values of signal and interference variances), their practical application requires that the estimates of these parameters should be used.

We have investigated the issue of stability of the steady identification process and determined the conditions for this stability. It has been shown that determining these conditions necessitates solving the third-degree equations, whose coefficients depend on the specificity of the problem to be solved. The resulting ratios are rather cumbersome but their simplification allows for a qualitative analysis of stability issues. It should be noted that all the estimates reported in this work depend on the choice of a mixing parameter, the task of determining which remains to be explored.

The estimates obtained in this paper allow the researcher to pre-evaluate the capabilities of the identification algorithm and the effectiveness of its use in solving practical problems

Keywords: combined functional, gradient algorithm, weighing parameter, asymptotic assessment, identification accuracy

STUDYING THE PROPERTIES OF A ROBUST ALGORITHM FOR IDENTIFYING LINEAR OBJECTS, WHICH MINIMIZES A COMBINED FUNCTIONAL

O. Rudenko

Doctor of Technical Sciences, Professor, Head of Department*

O. Bezsonov

Doctor of Technical Sciences, Professor*
E-mail: oleksandr.bezsonov@nure.ua

O. Lebediev

PhD, Associate Professor**

V. Lebediev

Postgraduate Student**

K. Oliinyk

Postgraduate Student

Department of Informatics***

*Department of Computer Intelligent Technologies and Systems***

Department of Electronic Computers*

***Kharkiv National University of Radio Electronics
Nauky ave., 14, Kharkiv, Ukraine, 61166

Received date 27.07.2020

Accepted date 17.08.2020

Published date 31.08.2020

Copyright © 2020, O. Rudenko, O. Bezsonov, O. Lebediev, V. Lebediev, K. Oliinyk

This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0>)

1. Introduction

Underlying many of the tasks related to processing information (processing and filtering of complex signals, identifying and managing objects, predicting time sequences, classification, etc.) is the task of building a model of the following form:

$$y(k) = \theta^{*T} x(k) + \xi(k), \tag{1}$$

where $y(k)$ is the observed output signal;

$$x(k) = (x_1(k), x_2(k), \dots, x_N(k))^T$$

is the vector of input signals $N \times 1$; $\theta^* = (\theta_1^*, \theta_2^*, \dots, \theta_N^*)^T$ is the vector of the desired parameters $N \times 1$; $\xi(k)$ is the interference that implies minimizing some of the predefined quality functional (identification criteria). A quadratic functional, the most widely used in practice, leads to various identification

algorithms, making it possible to obtain the estimates of the desired vector θ^* at the normal interference distributions, that is, $\xi(k) \sim N(0, \sigma_\xi^2)$.

Based on this assumption, the LSM-solution is asymptotically optimal with minimal variance in the class of non-displaced grades. However, this assumption does not generally hold in real-world conditions as almost always the *a priori* information about distributions is typically inaccessible, or interference is clogged with non-Gaussian noise. This results in some measurements being significantly removed from the core of the data, thereby forming so-called "tails". The instability of the LSM evaluation in the presence of such interference was the basis for the development of an alternative, robust assessment in statistics, which was aimed at eliminating the effects of interference [1–9].

If one has information about the interference ξ belonging to a certain class of distributions, the task is simplified. In this case, it is possible to obtain the maximum plausibility

(M-assessment) assessment by minimizing the optimal criterion, which is the reverse logarithm of the interference distribution function. If such information is not available, a non-quadratic criterion must be applied to assess the vector of θ^* parameters. This ensures that the estimate obtained is robust. One of these criteria is a modular one, whose minimization leads to a symbolic algorithm.

2. Literature review and problem statement

Theoretical research into the properties of the symbolic evaluation algorithm was reported for the first time in [1]. The practical application of this criterion in the task of identifying an object in the presence of pulse interference was considered in [2–6]. In particular, the effectiveness of the affinity projection symbolic algorithm was studied in [2, 3]; the affinity projection symbolic algorithm with a variable gain factor was used in [4]. It should be noted that the symbolic algorithms, while ensuring the robustness of the resulting assessment, have a low convergence rate. Therefore, in order to speed up the evaluation process, a normalized symbolic identification algorithm was proposed and examined in [5]. A simple-to-implement algorithm that uses a root mean square error and the estimated interference power to correct the length of the step is studied in [6].

The positive properties of the modular criterion are used in the so-called combined criteria, the most common of which is the combined functionals proposed in [7, 8]. They include a quadratic functional that provides optimal ratings for the Gauss distribution, as well as modular, producing an estimate that is more robust to the distributions with heavy “tails” (emissions). It should be noted, however, that the effectiveness of the robust estimates obtained depends significantly on the many parameters used in these criteria. The cited works provide some recommendations for choosing these parameters. In most cases, however, they are selected based on the experience of the researcher [9]. The task of robust neural network training based on the functionals [6, 7] by Huber and Hempel is considered in [10–12]; some practical recommendations on the choice of the functionals’ parameters are devised. A more general issue of robust assessment in the presence of interference with asymmetrical distributions was investigated in [13]. However, the task of choosing the functionals’ parameters remains to be resolved.

Works [14–18] developed a simpler approach to building combined functionals, consisting of both quadric and modular, and without the specified flaw.

Such a criterion was for the first time proposed in [14]. In [14–18], this criterion was used to solve the problem of identification in the presence of pulse interference. The stability of the normalized algorithm was studied in [15]; the applied identification problem was solved in [16]. An adaptive combination of normalized filters was proposed in [17]; the convergence of the identification algorithm was studied in [18], where the task of selecting the optimal parameters’ values of the algorithm was addressed.

The minimum fourth-degree criterion was proposed in [19], the properties of which were studied in [19–24]. Thus, the stability of the normalized algorithm in the presence of the non-Gaussian input signals was considered in [20]; the process of the algorithm normalization was described in [21]; papers [22, 23] considered the global stability of the appropriate algorithms; the problem of stochastic

analysis of the stability of the adaptive algorithm was tackled in [24]. The task of increasing the convergence rate of a given algorithm by using the optimal setting step parameter was studied in [25, 26]. Paper [27], in order to ensure the robustness and stability of the algorithm, proposed using a variable step parameter that takes into consideration the energy of the error (in the terms of the least squares). Study [28] proposed a modification of the algorithm of the method of the least fourth degree based on a quasi-Newtonian procedure. Finally, work [29] addressed the implementation of a given algorithm using quantum computations.

A combined assessment criterion to speed up the identification process, which uses the combination of the quadratic criterion and the fourth-degree criterion, is proposed in [30]. In [31], a given approach was used to speed up the identification process in the presence of pulse interference. The properties of the adaptive algorithm to minimize this combined criterion were studied in [32].

A combined criterion consisting of the fourth-degree and modular criteria was proposed in [33]; the specificity of its work was considered.

As revealed by an analysis of the above studies into the issue of the robust identification of control objects, the application of the combined criterion is quite effective. In addition, such an approach is much easier than when using traditional criteria. However, available papers do not include the results of studying the features of the robust algorithms for evaluating a model’s parameters built by using the combined criterion.

All this allows us to argue that it is appropriate to conduct a study on the analysis of the properties of the robust identification algorithm, which minimizes a combined functional, allowing for the combination of the LSM and LMM benefits.

3. The aim and objectives of the study

The aim of this work is to investigate issues related to the convergence and stability of the gradient algorithms that identify the parameters of a linear object in the presence of non-Gaussian noise.

To accomplish the aim, the following tasks have been set:

- to investigate the convergence of the robust identification procedure (to obtain analytical estimates of the convergence in the mean and the mean square of the gradient algorithm for minimizing the combined functional);
- to determine the most achievable (asymptomatic) values of errors in estimating the parameters and identification errors in the conditions under consideration;
- to define conditions for the stability of the steady identification process;
- to simulate the process of identification of a stationary linear object in the presence of non-Gaussian noise.

4. Studying the convergence of the robust identification procedure

The following combined functional is effective enough to ensure the robust properties of the estimates received.

$$F[e(k)] = \frac{1}{4} \lambda e^4(k) + (1 - \lambda) |e(k)|, \quad (2)$$

where

$$\begin{aligned} e(k) &= y(k) - \hat{y}(k) + \xi(k) = \\ &= y(k) - \hat{\theta}^T(k-1)x(k) + \xi(k), \end{aligned}$$

$\hat{y}(k)$ – a model's output signal;

$$\hat{\theta}(k-1) = (\hat{\theta}_1(k-1), \hat{\theta}_2(k-1), \dots, \hat{\theta}_N(k-1))^T$$

– a vector of the rated parameters of $N \times 1$; $\lambda \in [0, 1]$ – a mixing parameter.

When using criterion (2), the gradient minimization procedure takes the following form

$$\theta(k) = \theta(k-1) + \gamma(k) [\lambda e^3(k) + (1-\lambda) \text{sign} e(k)] x(k), \quad (3)$$

where $\gamma(k)$ is some parameter that affects the speed of the algorithm's convergence.

A given procedure combines the properties of LSM with the properties of LMM as, at $\lambda=1$, we have, from (3) the LSM algorithm, and, at $\lambda=0$, the LMM algorithm; that makes it possible to eliminate the non-Gaussian interference. By varying the λ parameter, one can change the properties of the algorithm.

Introduce an evaluation error

$$\tilde{\theta}(k) = \theta^*(k) - \hat{\theta}(k), \quad (4)$$

which makes it possible to write down an expression for $e(k)$ in the following form:

$$e(k) = \tilde{\theta}^T(k-1)x(k) + \xi(k) = e_a(k) + \xi(k), \quad (5)$$

where

$$e_a(k) = \tilde{\theta}^T(k-1)x(k)$$

is the *a priori* identification error.

Since it is assumed that $\xi(k) \sim N(0, \sigma_\xi^2)$, we have

$$M\{e^2(k)\} = \sigma_\xi^2 + \sigma_x^2 M\{\|\tilde{\theta}(k)\|^2\}, \quad (6)$$

where $M\{\bullet\}$ is the symbol of mathematical expectation; $\|\bullet\|$ is the Euclid norm.

Record algorithm (3) relative to the identification errors $\tilde{\theta}(i)$

$$\begin{aligned} \tilde{\theta}(k) &= \tilde{\theta}(k-1) - \gamma [\lambda e^3(k) + (1-\lambda) \text{sign} e(k)] x(k) = \\ &= \tilde{\theta}(k-1) - \gamma \left[\lambda (\tilde{\theta}(k-1)x(k) + \xi(k))^3 + \right. \\ &\quad \left. + (1-\lambda) \text{sign} (\tilde{\theta}(k-1)x(k) + \xi(k)) \right] x(k). \quad (7) \end{aligned}$$

Considering (5), rewrite expression (7) in the following form:

$$\begin{aligned} \tilde{\theta}(k) &= \tilde{\theta}(k-1) - \gamma \left[\lambda \left(e_a^3(k) + 3e_a^2(k)\xi(k) + \right. \right. \\ &\quad \left. \left. + 3e_a(k)\xi^2(k) + \xi^3(k) \right) + \right. \\ &\quad \left. + (1-\lambda) \text{sign}(e_a(k) + \xi(k)) \right] x(k). \quad (8) \end{aligned}$$

Consider the convergence of procedure (3) in the absence of interference, that is, $\xi(k)=0$. In this case, we shall use the approach applied in [25, 26].

Introduce the Lyapunov function $V(k) = \|\tilde{\theta}(k)\|^2$ and consider its increment

$$\Delta V(k) = \|\tilde{\theta}(k)\|^2 - \|\tilde{\theta}(k-1)\|^2. \quad (9)$$

After multiplying (7) on the left by $\tilde{\theta}^T(k)$ and considering that, in the case in question,

$$e(k) = e_a(k) = \tilde{\theta}^T(k-1)x(k),$$

after simple transformations, we obtain the following condition of convergence of the algorithm ($\Delta V(k) < 0$):

$$\left(\lambda e_a^4(k) + \right. \\ \left. + 2(1-\lambda)e_a(k) \right) > \gamma \left(\lambda e_a^3(k) + \right. \\ \left. + (1-\lambda)e_a(k) \right)^2 \|x(k)\|^2. \quad (10)$$

Thus, the algorithm convergence condition (3) is met if the γ parameter satisfies the following inequality

$$0 < \gamma < \frac{2e_a(k)}{(\lambda e_a^3(k) + (1-\lambda) \text{sign} e_a(k)) \|x(k)\|^2}. \quad (11)$$

An expression for the optimal value of the γ parameter, which provides the maximum convergence rate, is determined from the equation obtained by the differentiation of (8) for γ and equating the derivative to zero. Thus

$$\gamma^{\text{opt}} = \frac{e(k)}{(\lambda e^3(k) + (1-\lambda) \text{sign} e(k)) \|x(k)\|^2}. \quad (12)$$

Examine the statistical properties of evaluation procedure (3) in the presence of measurement interference, that is,

$$y^*(k) = \theta^{*T} x(k) + \xi(k), \quad \xi(k) \sim N(0, \sigma^2).$$

Suppose the interference is not correlated with usable signals. Having written down (3) relative to the errors of the assessment, we have (7).

Consider the mathematical expectation $M\{\tilde{\theta}(k)\}$. Given (5), after averaging both sides of (7), we obtain

$$\begin{aligned} M\{\tilde{\theta}(k)\} &= \\ &= M \left\{ \begin{aligned} &\tilde{\theta}(k-1) - \gamma \lambda (\tilde{\theta}^T(k-1)x(k) + \xi(k))^3 x(k) - \\ &- \gamma (1-\lambda) \text{sign}(\tilde{\theta}^T(k-1)x(k)) x(k) \end{aligned} \right\}. \quad (13) \end{aligned}$$

It is easy to see

$$\begin{aligned} M\{\tilde{\theta}^T(k-1)x(k)x(k)\} &= M\{x(k)x^T(k)\tilde{\theta}(k-1)\} = \\ &= M\{x(k)x^T(k)\} M\{\tilde{\theta}(k-1)\} = R_{xx} M\{\tilde{\theta}(k-1)\}; \end{aligned}$$

$$\begin{aligned} M\{(\tilde{\theta}^T(k-1)x(k))^3 x(k)\} &= \\ &= M \left\{ \begin{aligned} &(\tilde{\theta}^T(k-1)x(k))^2 \times \\ &\times x(k)x^T(k)\tilde{\theta}(k-1) \end{aligned} \right\} = \\ &= M\{e_a^2(k)\} M\{x(k)x^T(k)\} M\{\tilde{\theta}(k-1)\} = \\ &= \sigma_{e_a}^2 R_{xx} M\{\tilde{\theta}(k-1)\}, \quad (14) \end{aligned}$$

where $\sigma_{e_a}^2$ is the root mean square value of the $e_a(k)$ error; R_{xx} – the correlation matrix of the input signal.

Let us take a closer look at the expression

$$M\{x(k)\text{sign}(\tilde{\theta}^T(k-1)x(k))x(k)\}.$$

For the case when the signal is

$$x(k) \sim N(0, \sigma_x^2),$$

we obtain

$$\begin{aligned} &M\{x(k)\text{sign}(\tilde{\theta}^T(k-1)x(k))\} = \\ &= M\{M\{x(k)\text{sign}(\tilde{\theta}^T(k-1)x(k))\}\} = \\ &= M\left\{\sqrt{\frac{2}{\pi}} \frac{1}{\sigma_x} M\{x(k)(\theta^T(k-1)x(k))\}\right\} = \\ &= M\left\{\sqrt{\frac{2}{\pi\sigma_x^2}} M\{x(k)x^T(k)\theta(k-1)\}\right\} = \\ &= \sqrt{\frac{2}{\pi\sigma_x^2}} M\{x(k)x^T(k)\theta(k-1)\} = \\ &= \sqrt{\frac{2}{\pi\sigma_x^2}} R_{xx} M\{\theta(k-1)\}. \end{aligned} \tag{15}$$

Expression (15) is derived from a Price's theorem, whereby for two random Gaussian quantities x and y with zero mathematical expectations, the following form is true

$$M\{x \text{sign } y\} = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_y} M\{xy\},$$

where σ_y is the root mean square value of y .

Taking into consideration the properties of interference

$$M\{\xi(k)x(k)\} = 0, \quad M\{\xi(k)\} = M\{\xi^3(k)\} = 0$$

and expressions (8), (9), we have

$$M\{\tilde{\theta}(k)\} = \left\{ \begin{aligned} &I - \gamma\lambda\sigma_x^2 R_{xx} - 3\gamma\lambda \times \\ &\times \sigma_\xi^2 R_{xx} - \gamma(1-\lambda)\beta R_{xx} \end{aligned} \right\} M\{\tilde{\theta}(k-1)\}, \tag{16}$$

hence, it follows that procedure (3) will converge on average if the γ parameter satisfies the following inequality

$$0 < \gamma < \frac{2}{\left(\lambda(\sigma_x^2 + 3\sigma_\xi^2) + (1-\lambda)\beta\right) tr R_{xx}}. \tag{17}$$

Here, $\beta = \sqrt{\frac{2}{\pi\sigma_x^2}}$; $tr R_{xx}$ is the trace of the R_{xx} matrix.

$$0 < \gamma < \frac{2}{\left(\lambda(\sigma_x^2 + 3\sigma_\xi^2) + (1-\lambda)\beta\right) N\sigma_x^2}. \tag{18}$$

To study the convergence of the algorithm in root mean square, let us consider the Lyapunov function $M\{\|\tilde{\theta}(k)\|^2\}$.

Multiplying both sides of (7) on the left by $\tilde{\theta}^T(k)$, considering (5), we obtain

$$\begin{aligned} \|\tilde{\theta}(k)\|^2 &= \|\tilde{\theta}(k-1)\|^2 - \\ &- 2\gamma \left[\lambda \left(e_a^4(k) + 3e_a^3(k)\xi(k) + 3e_a^2(k)\xi^2(k) + \right. \right. \\ &\left. \left. + e_a(k)\xi^3(k) + (1-\lambda)e_a \text{sign} e_a(k) \right) \right] + \\ &+ \gamma^2 \left\{ \lambda^2 \left[e_a^6(k) + 6e_a^5(k)\xi(k) + 15e_a^4(k)\xi^2(k) + \right. \right. \\ &\left. \left. + 20e_a^3(k)\xi^3(k) + 15e_a^2(k)\xi^4(k) + 6e_a\xi^5(k) + \xi^6(k) \right] \|x(k)\|^2 + \right. \\ &\left. + 2\lambda(1-\lambda) \left[\left(e_a^3(k) + 3e_a^2(k)\xi(k) + 3e_a(k)\xi^2(k) + \xi^3(k) \right) \text{sign} e_a(k) \right] \|x(k)\|^2 \right\} + \\ &+ \gamma^2(1-\lambda)^2 \|x(k)\|^2. \end{aligned} \tag{19}$$

Then

$$\begin{aligned} M\{\|\tilde{\theta}(k)\|^2\} &= M\{\|\tilde{\theta}(k-1)\|^2\} - \\ &- 2\gamma M \left\{ \left[\lambda \left(e_a^4(k) + 3e_a^3(k)\xi(k) + 3e_a^2(k)\xi^2(k) + \right. \right. \right. \\ &\left. \left. + e_a(k)\xi^3(k) + (1-\lambda)e_a \text{sign} e_a(k) \right) \right] \right\} + \\ &+ \gamma^2 M \left\{ \left\{ \lambda^2 \left[e_a^6(k) + 6e_a^5(k)\xi(k) + 15e_a^4(k)\xi^2(k) + \right. \right. \right. \\ &\left. \left. + 20e_a^3(k)\xi^3(k) + 15e_a^2(k)\xi^4(k) + 6e_a\xi^5(k) + \xi^6(k) \right] \|x(k)\|^2 + \right. \\ &\left. + 2\lambda(1-\lambda) \left[\left(e_a^3(k) + 3e_a^2(k)\xi(k) + 3e_a(k)\xi^2(k) + \xi^3(k) \right) \text{sign} e_a(k) \right] \|x(k)\|^2 \right\} + \\ &+ \gamma^2(1-\lambda)^2 M\{\|x(k)\|^2\}. \end{aligned} \tag{20}$$

It is easy to obtain formulae to calculate the expressions included in (20). For example, for $M\{e_a \text{sign } e_a\}$, we have

$$\begin{aligned} M\{e_a \text{sign } e_a\} &= M\{M\{e_a \text{sign } e_a \mid \theta(k-1)\}\} = \\ &= M\left\{ \tilde{\theta}^T(k-1) \sqrt{\frac{2}{\pi\sigma_e}} M\{x(k)x^T(k)\tilde{\theta}(k-1) \mid \tilde{\theta}(k-1)\} \right\} \approx \\ &\approx M\{\tilde{\theta}^T(k-1)\beta\tilde{\theta}(k-1)\} R_{xx} = \\ &= \beta tr R_{xx} M\{\|\tilde{\theta}(k-1)\|^2\} = \beta\sigma_x^2 M\{\|\tilde{\theta}(k-1)\|^2\}. \end{aligned}$$

The formulae below are derived similarly:

$$M\{e_a^2(k)\} = \sigma_x^2 M\{\|\theta(k-1)\|^2\};$$

$$M\{e_a^2(k)\|x(k)\|^2\} = (N+2)\sigma_x^4 M\{\|\theta(k-1)\|^2\};$$

$$M\{e_a^4(k)\} \approx 3\sigma_x^4 M\{\|\theta(k-1)\|^2\};$$

$$M\{e_a^4(k)\|x(k)\|^2\} \approx (3N+12)\sigma_x^6 \left(M\{\|\theta(k-1)\|^2\} \right)^2;$$

$$M\{e_a^6(k)\|x(k)\|^2\} \approx (15N+90)\sigma_x^8 \left(M\{\|\theta(k-1)\|^2\} \right)^3;$$

$$\begin{aligned} M\{e_a^3(k)\text{sign} e_a(k)\|x(k)\|^2\} &\approx (3N+12)\beta\sigma_x^6 \times \\ &\times \left(M\{\|\theta(k-1)\|^2\} \right)^2; \end{aligned}$$

$$M\{e_a(k)\text{sign} e_a(k)\|x(k)\|^2\} \approx (N+2)\beta\sigma_x^4 M\{\|\theta(k-1)\|^2\};$$

$$M\{e_a^2(k)\} M\{\xi^4(k)\} \approx \sigma_x^2 M\{\xi^4(k)\} M\{\|\tilde{\theta}(k-1)\|^2\};$$

$$M\{\|x(k)\|^2 \xi^6\} = M\{\|x(k)\|^2\} M\{\xi^6\} = N\sigma_x^2 M\{\xi^6\};$$

$$M\{\|x(k)\|^2\} = N\sigma_x^2.$$

Here,

$$\beta = \sqrt{\frac{2}{\pi\sigma_{e_a}^2}}.$$

Substituting these expressions in (20) and taking into consideration the statistical properties of the signals and interference, in particular,

$$M(\xi(k)) = M(\xi^3(k)) = M(\xi^5(k)) = 0,$$

we obtain

$$\begin{aligned} M\{\|\tilde{\theta}(k)\|^2\} = & \\ & \left(1 - 6\gamma\lambda\sigma_x^2\sigma_\xi^2 + 15(N+2)\gamma^2\lambda^2 \times \right. \\ & \left. \times \sigma_x^4 M\{\xi^4(k)\} - 2\gamma(1-\lambda)\beta\sigma_x^2\right) M\{\|\tilde{\theta}(k)\|^2\} + \\ & + \left(15(3N+12)\gamma^2\lambda^2\sigma_x^6\sigma_\xi^2 + \right. \\ & \left. + 6\gamma^2\lambda(1-\lambda)\sigma_x^4\beta - 6\gamma\lambda\sigma_x^4\right) \left(M\{\|\tilde{\theta}(k-1)\|^2\}\right)^2 + \\ & + (15N+90)\gamma^2\lambda^2\sigma_x^8 \left(M\{\|\tilde{\theta}(k-1)\|^2\}\right)^3 + \\ & + \gamma^2\sigma_x^2 N \left(\lambda^2 M\{\xi^6(k)\} + (1-\lambda)^2\right). \end{aligned} \quad (21)$$

If the algorithm converges, the value of $M\{\|\tilde{\theta}(k)\|^2\}$ will be small. Therefore, to analyze the steady state, expression (21) can be simplified by neglecting the quantities $\left(M\{\|\tilde{\theta}(k-1)\|^2\}\right)^2$ and $\left(M\{\|\tilde{\theta}(k-1)\|^2\}\right)^3$ and by confining ourselves to the consideration of the quantity

$$\begin{aligned} M\{\|\tilde{\theta}(k)\|^2\} = & \left(\begin{array}{l} 1 - 6\gamma\sigma_x^2 M\{\xi^2(k)\} + \\ + 15\gamma^2\lambda^2(N+2) \times \\ \times \sigma_x^4 M\{\xi^4(k)\} + \\ + 6\gamma^2\lambda(1-\lambda)\beta(N+2) \times \\ \times \sigma_x^4 M\{\xi^2(k)\} + \\ + 2\gamma(1-\lambda)\beta\sigma_x^2 \end{array} \right) M\{\|\tilde{\theta}(k)\|^2\} + \\ & + \gamma^2\sigma_x^2 N \left(\lambda^2 M\{\xi^6(k)\} + (1-\lambda)^2\right). \end{aligned} \quad (22)$$

It follows from (22) that procedure (3) will converge in the root mean square (the increment of the Lyapunov function will be negative) when the following condition is met

$$\left| \begin{array}{l} 1 - 6\gamma\sigma_x^2 M\{\xi^2(k)\} + 15\gamma^2\lambda^2(N+2)\sigma_x^4 M\{\xi^4(k)\} + \\ + 6\gamma^2\lambda(1-\lambda)\beta(N+2)\sigma_x^4 M\{\xi^2(k)\} + 2\gamma(1-\lambda)\beta\sigma_x^2 \end{array} \right| < 1,$$

that is, if the γ parameter satisfies the following inequality

$$0 < \gamma < \frac{2(3M\{\xi^2(k)\} + (1-\lambda)\beta)}{3\lambda(N+2)[2(1-\lambda)\beta\sigma_x^2 M\{\xi^2(k)\} + 5\lambda\sigma_x^2 M\{\xi^4(k)\}]}. \quad (23)$$

The optimal value of this parameter, which ensures the maximum convergence rate of the algorithm, which is obtained by solving the following equation

$$\frac{\partial M\{\|\tilde{\theta}(k)\|^2\}}{\partial \gamma} = 0,$$

takes the following form:

$$\begin{aligned} \gamma^{opt} = & \\ = & \frac{3M\{\xi^2(k)\} + (1-\lambda)\beta}{3\lambda(N+2)[2(1-\lambda)\beta\sigma_x^2 M\{\xi^2(k)\} + 5\lambda\sigma_x^2 M\{\xi^4(k)\}]}. \end{aligned} \quad (24)$$

As can be seen from (24), the γ^{opt} magnitude depends on the dimensionality of the object N under study, the statistical properties of the signals and interference, the magnitude $\sigma_{e_a}^2$, and the mixing parameter λ . Usually, the dimensionality of an object is known, the parameter λ can be chosen by the researcher. The statistical characteristics of the signals and interference (especially the moments of the fourth order) are often unknown. Therefore, a given formula makes it possible to determine the effect of other parameters on the properties of the algorithm.

5. Determining the asymptomatic values of assessment and identification errors

Ratio (23) can be used to obtain an expression for an asymptotic assessment error

$$M\{\|\tilde{\theta}(\infty)\|^2\} = \frac{\gamma N [\lambda^2 M\{\xi^6\} + (1-\lambda)^2]}{15\gamma(N+2)M\{\xi^4\} - 6\lambda\sigma_\xi^2 - 2(1-\lambda)\beta}. \quad (25)$$

It follows from (25) that ensuring

$$\lim_{k \rightarrow \infty} M\{\|\tilde{\theta}(\infty)\|^2\} = 0$$

requires that the parameter γ should be chosen as variable and, with the growth of k , to strive to zero, that is, to meet the Dvoretzkiy conditions [35].

Fitting (24) to (6) produces an expression for the asymptomatic identification error

$$\begin{aligned} M\{e^2(\infty)\} = & \\ = & \sigma_\xi^2 + \frac{\sigma_x^2 \gamma N [\lambda^2 M\{\xi^6\} + (1-\lambda)^2]}{15\gamma(N+2)M\{\xi^4\} - 6\lambda\sigma_\xi^2 - 2(1-\lambda)\beta}. \end{aligned} \quad (26)$$

As it follows from (26), the magnitude of the asymptomatic error of identification depends on the dimensionality of the examined object N , the statistical properties of the signals and interference, the magnitude of the γ parameter and the mixing parameter λ . If the dimensionality of an object is known, and the parameters γ and λ can be chosen by the researcher, the statistical characteristics of the signals and noise (especially the moments of the fourth and sixth orders) are often unknown. Therefore, a given formula is rather of theoretical interest as it characterizes the limits of the algorithm.

6. Determining the stability of the steady evaluation process

When studying the stability of an evaluation process, we shall take the approach proposed in [24].

Write down expression (22) in the following form

$$y(k+1) = (1-a)y(k) - by^2(k) + cy^3(k) + d, \tag{27}$$

where

$$y(k) = M \left\{ \left\| \tilde{\theta}(k) \right\|^2 \right\},$$

$$a = A_1\gamma - A_2\gamma^2,$$

$$b = B_1\gamma - B_2\gamma^2,$$

$$c = C\gamma^2,$$

$$d = D\gamma^2,$$

$$A_1 = 6\lambda\sigma_x^2 M \left\{ \xi^2(k) \right\} - 2(1-\lambda);$$

$$A_2 = 15\lambda^2(N+2)\sigma_x^4 M \left\{ \xi^2(k) \right\} + 6\lambda(1-\lambda)(N+2)\beta\sigma_x^4 M \left\{ \xi^2(k) \right\};$$

$$B_1 = 6\lambda\sigma_x^4;$$

$$B_2 = 15\lambda^2(3N+12)\sigma_x^6 M \left\{ \xi^2(k) \right\} + 2\lambda(1-\lambda)(3N+12)\beta\sigma_x^6;$$

$$C = \lambda^2(15+90)\sigma_x^8;$$

$$D = \lambda M \left\{ \xi^6(k) \right\} + (1-\lambda)\sigma_x^2.$$

Nonlinear differential equation (27) describes the dynamics of algorithm (3). Obviously, the convergence depends on the magnitude of the initial error

$$y(0) = M \left\{ \tilde{\theta}(0) \right\}.$$

To study the conditions of stability, one needs to find the equilibrium points.

Consider the steady state of the evaluation process. Assuming

$$y(k+1) = y(k) = y(\infty),$$

record (27) in the following form:

$$y(\infty) = (1-a)y(\infty) - by^2(\infty) + cy^3(\infty) + d, \tag{28}$$

Since $c > 0$, (28) can be rerecorded in the following form

$$y^3(\infty) - \frac{b}{c}y^2(\infty) - \frac{a}{c}y(\infty) + \frac{d}{c} = 0. \tag{29}$$

A given equation has three roots that define the equilibrium points and take the following form [36]

$$y_1(\infty) = (s_1 + s_2) + \frac{b}{3c};$$

$$y_2(\infty) = -\frac{1}{2}(s_1 + s_2) + \frac{b}{3c} + \frac{j\sqrt{3}}{2}(s_1 - s_2);$$

$$y_3(\infty) = -\frac{1}{2}(s_1 + s_2) + \frac{b}{3c} - \frac{j\sqrt{3}}{2}(s_1 - s_2);$$

$$s_1 = \left(r + \sqrt{q^3 + r^2} \right)^{\frac{1}{3}};$$

$$s_2 = \left(r - \sqrt{q^3 + r^2} \right)^{\frac{1}{3}};$$

$$q = -\frac{a}{3c} - \frac{b^2}{9c^2};$$

$$r = \frac{1}{6} \left(\frac{ab}{c^2} - \frac{3d}{c} \right) - \frac{b^3}{27c^3}.$$

Depending on the q and r values, three cases are possible

$$q^3 + r^2 < 0; \tag{30}$$

$$q^3 + r^2 = 0; \tag{31}$$

$$q^3 + r^2 > 0. \tag{32}$$

In the first case, as shown in paper [24], equation (29) has either three negative physical roots or one negative and two positive physical roots. Negative roots are of no interest as $y(k)$ is a square, that is, it must be non-negative. To study the positive roots, work [24] investigated the behavior of curve (27) $y(k+1)$ and defined the condition for a stable equilibrium point.

$$\min \left\{ \frac{dy(k+1)}{dy(k)} \right\} = \frac{3c(1-a) - b^2}{3c} > -1. \tag{33}$$

In particular, the convergence will be monotonous if

$$\min \left\{ \frac{dy(k+1)}{dy(k)} \right\} = \frac{3c(1-a) - b^2}{3c} \geq 0.$$

And, since $c > 0$, this condition is met at

$$3c(1-a) \geq b^2. \tag{34}$$

To make sure under which parameters of the algorithm (34) holds, we shall substitute in this inequality the expressions for a , b , and c . After simple transformations, we obtain

$$\begin{aligned} & 12\lambda - 60\gamma\sigma_x^2 M \left\{ \xi^2 \right\} (3N+12) - 8\gamma(1-\lambda)\sigma_x^2\beta + \\ & + 75\gamma^2\sigma_x^4 M \left\{ \xi^4 \right\} (3N+12)^2 + \\ & + 20\sigma_x^4\beta M \left\{ \xi^2 \right\} (1-\lambda)(3N+12)^2 + \\ & + 4\sigma_x^4\beta^2(1-\lambda)(3N+12)^2 \leq \\ & \leq ((15N+90) - 6\gamma\lambda\sigma_x^2 M \left\{ \xi^2 \right\} (15N+90) + \\ & + 15\lambda^2\gamma^2\sigma_x^4 M \left\{ \xi^4 \right\} (N+2)(15N+90) + \\ & + 6\lambda(1-\lambda)\gamma^2\sigma_x^4 M \left\{ \xi^2 \right\} (N+2)(15N+90). \end{aligned} \tag{35}$$

(35) shows that meeting the condition (34) depends on the dimensionality of the problem N , the magnitudes of γ and λ , as well as the statistical properties of the signals and interference $\sigma_x^2, \sigma_x^4, M\{\xi^2\}, M\{\xi^4\}$. Since N is defined by the problem being solved, the degree of robustness of the solution is determined by λ , the only freely chosen parameter is γ . Inequality (35) can be used to obtain the conditions that this parameter must satisfy to ensure the stability of the estimation process.

To this end, rewrite inequality (35) in the following form:

$$P\gamma^2 + R\gamma + Q \geq 0, \quad (36)$$

where

$$\begin{aligned} P &= 15\lambda^2\sigma_x^4M\{\xi^4\}(N+2)(15N+90) + \\ &+ 6\lambda(1-\lambda)\sigma_x^4M\{\xi^2\}(N+2)(15N+90) - \\ &- 75\sigma_x^4M\{\xi^4\}(3N+12)^2; \\ R &= 60\sigma_x^2M\{\xi^2\}(3N+12) + 8(1-\lambda)\sigma_x^2\beta; \\ Q &= 15N+90 - 12\lambda - 20\sigma_x^4\beta M\{\xi^2\}(1-\lambda)(3N+12)^2 - \\ &- 4\sigma_x^4\beta^2(1-\lambda)(3N+12)^2. \end{aligned}$$

By solving inequality (36), we obtain a condition for choosing γ

$$0 \leq \gamma \leq \frac{-R \pm \sqrt{R^2 - 4PQ}}{2P}. \quad (37)$$

Substituting in (37) the expressions for P , R and Q , one can obtain a rather cumbersome analytical expression for choosing γ . Even for a simple case corresponding to $\lambda=1$, studied in [24], the analysis is very complex and requires significant simplifications. A qualitative analysis of inequality (37) shows that the value of the parameter depends on all the quantities included in (36).

In the second case, that is, when meeting (31), equation (29) has two valid positive and equal roots and one negative, which again is of no interest. This case defines the resistance limit for choosing the γ step length parameter. For this case, ($q^3+r^2=0$), we obtain

$$\begin{aligned} q^3 + r^2 &= \left\{ \frac{a}{3c} - \frac{b^2}{9c^2} \right\}^3 + \left\{ \frac{1}{6} \left(\frac{ab}{c^2} - \frac{3d}{c} \right) + \frac{b^3}{27c^3} \right\}^2 = \\ &= -\frac{1}{729}(3ac+b^2)^3 + \frac{1}{729c^6} \left(\frac{9}{2}abc - \frac{27}{2}dc^2 + b^3 \right)^2 = 0. \end{aligned}$$

Hence, we obtain

$$4(3ac+b^2)^3 = (9abc - 27dc^2 + 2b^3)^2. \quad (38)$$

By recording (38) in the form of a polynomial from γ and by substituting the expressions for a , b , c , and d as the functions from A_1, A_2, B_1, B_2, C , and D , we obtain

$$P_4\gamma^4 + P_3\gamma^3 + P_2\gamma^2 + P_1\gamma + P_0 = 0, \quad (39)$$

where

$$P_4 = -4A_2^3C + A_2^2B_2^2 + 18A_2B_2 - 4B_2^3D - 27C^2D^2;$$

$$\begin{aligned} P_3 &= 12A_1A_2^2C - 2(A_1A_2B_2^2 + A_2^2B_1B_2) - \\ &- 18(A_1B_2 + A_2B_1)CD + 12B_1B_2^2D; \end{aligned}$$

$$\begin{aligned} P_2 &= -12A_1^2A_2C + A_1^2B_2^2 + A_2^2B_1^2 + \\ &+ 4A_1A_2B_1B_2 + 18A_1B_1CD - 12B_1^2B_2D; \end{aligned}$$

$$P_1 = 4A_1^3C - 2(A_1A_2B_1^2 + A_1^2B_1B_2) + 4B_1^3D; \quad P_0 = A_1^2B_1^2.$$

The lowest positive root γ_0 (39) corresponds to the resistance limit. At the same time, as noted in [24], the existence of interference, even when the parameters $\hat{\theta}(0)$ are initiated based on the values close to θ^* , leads to the instability of the evaluation process.

Finally, for the third case, when (32) is met, equation (29) may have three physical negative roots, which is of no interest, or one physical negative root, but two complex ones. This case corresponds to the lack of stability in the learning process.

7. Modeling the identification process

We have considered the problem of identifying a stationary linear object, which is described by equation (1), with the following parameters

$$\theta^* = (-110; -25; -19; -40; -68; 16; 32; 90; 120; 240)^T.$$

The chosen input signal $x(k)$ was the sequences of the normally distributed quantities $x(k) \sim N(0; 1)$. When testing the robustness of the algorithms, independent noise was added to the output signal of the object, with the Laplace distribution ($\alpha=1.0$) and clogging Gaussian noise with $\sigma=48$. The histogram of such interference is shown in Fig. 1.

The results of the simulation at different values of the λ parameter are shown in Fig. 2, 3. Fig. 2, *a, c* shows the diagrams of setting the model parameters when selecting $\lambda=1, \lambda=0.5$, and $\lambda=0$, respectively; Fig. 3, *a-c* – the identification errors.

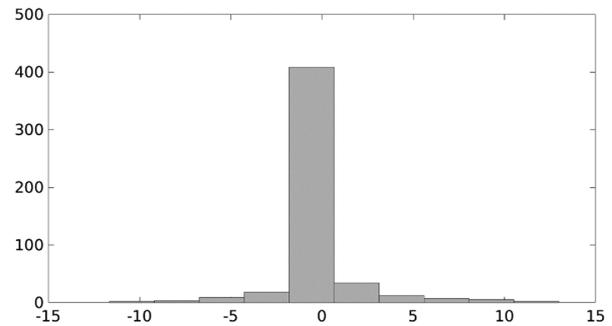


Fig. 1. Interference histogram

As the modeling results show, when using only the fourth-degree criterion (Fig. 3, *a*), it becomes impossible to evaluate the parameters of the model in the presence of mixed interference. When using only a modular criterion, an estimate is possible, but the convergence of the algorithm stretches over time (Fig. 3, *b*). The application of the mixed criterion (Fig. 3, *c*) is optimal.

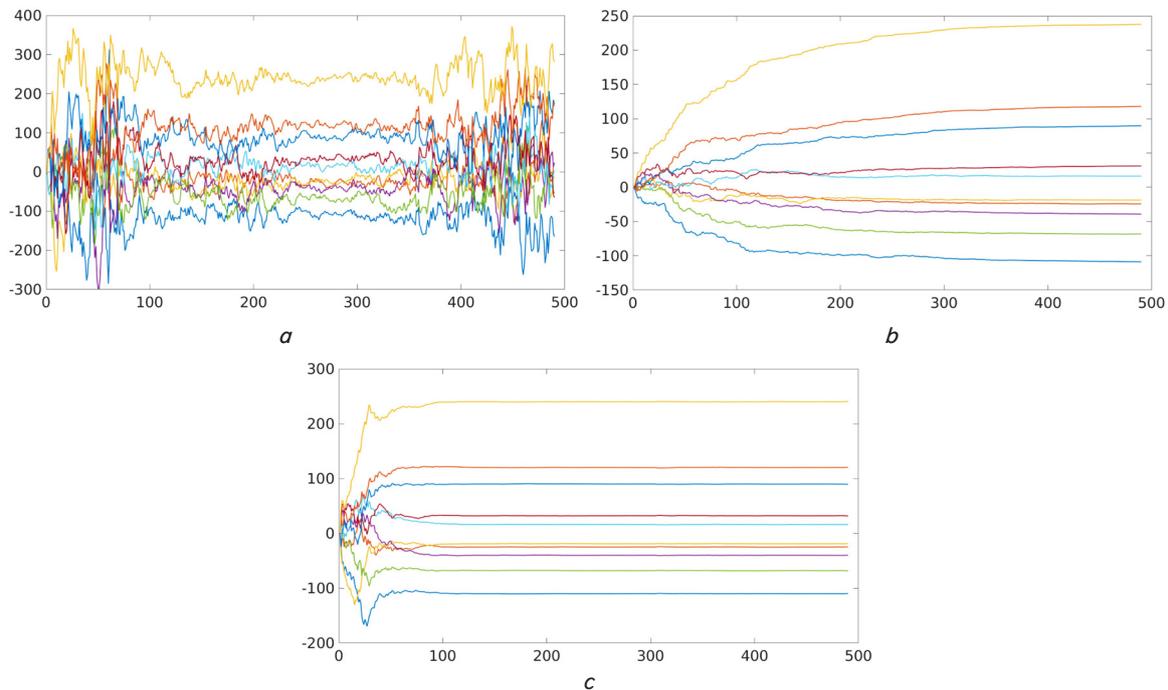


Fig. 2. Diagram of setting the model's parameters: *a* – at $\lambda=1$; *b* – $\lambda=0.5$; *c* – $\lambda=0$

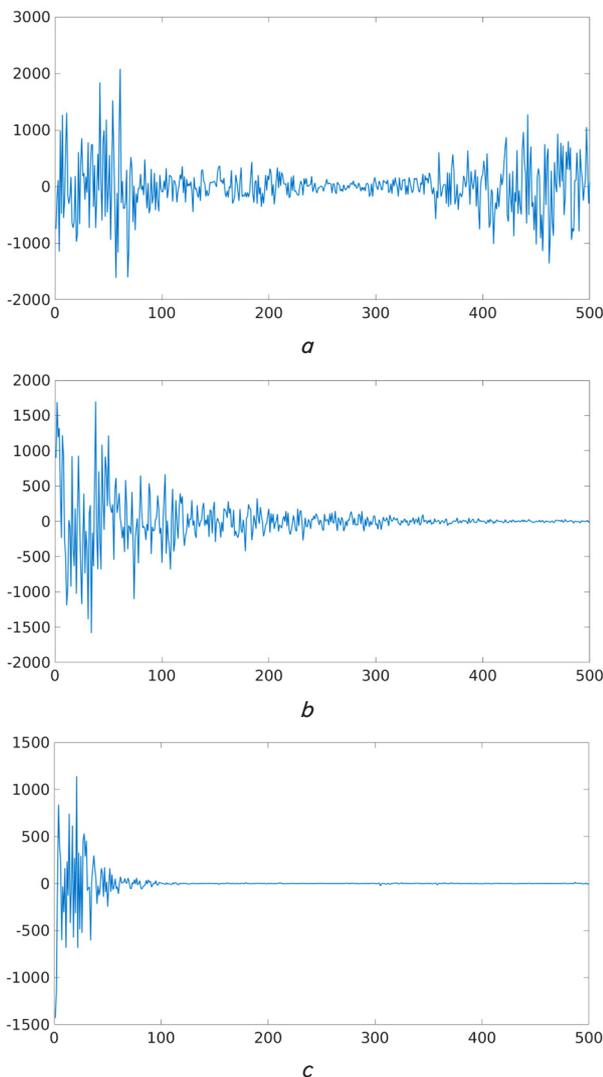


Fig. 3. Identification error: *a* – at $\lambda=1$; *b* – $\lambda=0.5$; *c* – $\lambda=0$

8. Discussion of results of studying the convergence of a gradient algorithm for the identification of a linear object

The issues of the convergence of the gradient algorithms to identify the parameters of a linear object in the presence of non-Gaussian noise have been investigated. We have determined the conditions of the convergence of algorithms in the absence of interference (11), on average (18), and root mean square (23), if the noise is present.

The expressions have been derived to determine the optimal values of the algorithm's parameters, ensuring its maximum convergence rate in the absence and presence of interference – formulae (12) and (24), respectively.

Based on the estimates obtained, the most achievable (asymptotic) values of identification errors (25) and parameters assessment errors (26) under the conditions under consideration have been determined. Even though the derived expressions contain a series of unknown parameters (the signals σ_x^2 and interference σ_ξ^2 variances), they allow for qualitative analysis. These formulae are more of theoretical interest as they characterize the limits of the algorithm.

A non-linear differential equation (27), describing the dynamics of the algorithm, was used to study the stability of the steady identification process. It has been shown that meeting stability condition (34) depends on the dimensionality of the problem N , the quantities γ and λ , as well as the statistical properties of the signals and interference. Since N is defined by the problem being solved, λ characterizes the degree of robustness of the solution derived, the only freely chosen parameter is γ . The condition for the choice of γ (37), which is rather cumbersome, has been obtained. Its simplification, however, allows for a qualitative analysis of stability issues.

In modeling the identification process, the non-Gaussian noise was added to the object's output signal (with the Laplace distribution and the clogging Gaussian noise). If one uses the fourth-degree criterion only, it becomes impossible to evaluate the parameters of the model in the presence of

mixed interference. The results of the simulation show that with only the fourth-degree criterion, studied in [19–28], the evaluation of model parameters in the presence of mixed interference becomes ineffective. Using only a modular criterion makes it possible to obtain an assessment of the parameters but the identification process is delayed. In this case, the best option is to apply the combined criterion considered in the present work.

The limitations of this study include the need for information on the statistical properties of the usable signals and interference. In addition, the effectiveness of solving the identification problem depends significantly on the choice of the mixing parameter λ that determines the robustness of the assessment. At present, there are no general recommendations for the choice of λ . Since the algorithms considered are designed to solve the problem of real-time identification, it seems appropriate to further develop effective procedures for assessing the statistical characteristics of signals and interference and the rules for selecting the weighting parameter. This is even more important because a given approach can be applied to the identification of the dynamic objects, represented, for example, by a pseudolinear regression model. It is therefore appropriate to conduct research on the development of recommendations for the choice of the λ parameter for different types of distributions and their possible combinations.

The estimates obtained in this paper allow the researcher to pre-evaluate the capabilities of a given algorithm and the effectiveness of its application when solving practical problems.

In conclusion, the study reported in this paper is a continuation and the advancement of earlier research, the results of which are described in [18, 33].

9. Conclusions

1. The convergence of the identification algorithm, which minimizes a combined functional, consisting of the fourth-degree and modular functionals, and ensuring the robustness of estimates has been investigated. The conditions for the convergence of algorithms have been determined, in the absence of interference, as well as on average and root mean square if it is present. Expressions have been obtained

to determine the optimal values of the algorithm's parameters, which ensure its maximum convergence rate in the absence and presence of interference.

2. Based on the estimates obtained, the most achievable (asymptomatic) values of identification errors and errors in the assessment of parameters in the conditions under consideration have been determined. Although the expressions derived contain a series of unknown parameters (the signals σ_x^2 and interference σ_ξ^2 variance), they allow for qualitative analysis. The resulting formulae are more of theoretical interest as they characterize the limits of the algorithm.

3. The stability of the steady identification process using a non-linear differential equation describing the dynamics of the algorithm has been investigated. It has been shown that meeting stability condition (34) depends on the dimensionality of the N problem, the quantities γ and λ , as well as the statistical properties of the signals and interference. Since N is defined by the problem being solved, λ characterizes the degree of robustness of the solution, the only freely chosen parameter is γ . Its simplification, however, makes it possible to perform a qualitative analysis of stability issues.

4. The identification process was simulated when there was a non-Gaussian noise in the object's output signal (with the Laplace distribution and the clogging Gaussian noise). The results of the simulation have confirmed the effectiveness of the developed approach. However, this approach requires information on the statistical properties of the usable signals and interference. In addition, the effectiveness of solving an identification problem depends significantly on the choice of the mixing parameter γ that determines the robustness of the assessment.

The estimates obtained in this work allow the researcher to pre-evaluate the limits of the algorithm and the effectiveness of its application when solving practical tasks.

Acknowledgments

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

References

1. Bedel'baeva, A. A. (1978). Relay estimation algorithm. *Avtomat. i Telemekh.*, 39 (1), 87–95.
2. Shao, T., Zheng, Y. R., Benesty, J. (2010). An Affine Projection Sign Algorithm Robust Against Impulsive Interferences. *IEEE Signal Processing Letters*, 17 (4), 327–330. doi: <https://doi.org/10.1109/lsp.2010.2040203>
3. Shin, J., Yoo, J., Park, P. (2012). Variable step-size affine projection sign algorithm. *Electronics Letters*, 48 (9), 483. doi: <https://doi.org/10.1049/el.2012.0751>
4. Lu, L., Zhao, H., Li, K., Chen, B. (2015). A Novel Normalized Sign Algorithm for System Identification Under Impulsive Noise Interference. *Circuits, Systems, and Signal Processing*, 35(9), 3244–3265. doi: <https://doi.org/10.1007/s00034-015-0195-1>
5. Huang, H.-C., Lee, J. (2012). A New Variable Step-Size NLMS Algorithm and Its Performance Analysis. *IEEE Transactions on Signal Processing*, 60 (4), 2055–2060. doi: <https://doi.org/10.1109/tsp.2011.2181505>
6. Casco-Sánchez, F. M., Medina-Ramírez, R. C., López-Guerrero, M. (2011). A New Variable Step-Size NLMS Algorithm and its Performance Evaluation in Echo Cancelling Applications. *Journal of Applied Research and Technology*, 9 (03). doi: <https://doi.org/10.22201/icat.16656423.2011.9.03.425>
7. Huber, P. J. (1977). Robust methods of estimation of regression coefficients. *Series Statistics*, 8 (1), 41–53. doi: <https://doi.org/10.1080/02331887708801356>
8. Hampel, F. R. (1974). The Influence Curve and its Role in Robust Estimation. *Journal of the American Statistical Association*, 69 (346), 383–393. doi: <https://doi.org/10.1080/01621459.1974.10482962>

9. Adamczyk, T. (2017). Application of the Huber and Hampel M-estimation in real estate value modeling. *Geomatics and Environmental Engineering*, 11 (1), 15. doi: <https://doi.org/10.7494/geom.2017.11.1.15>
10. Rudenko, O. G., Bezsonov, O. O. (2011). Robust training of radial basis networks. *Cybernetics and Systems Analysis*, 47 (6), 863–870. doi: <https://doi.org/10.1007/s10559-011-9365-8>
11. Rudenko, O. G., Bezsonov, O. O. (2014). Robust Neuroevolutionary Identification of Nonlinear Nonstationary Objects. *Cybernetics and Systems Analysis*, 50 (1), 17–30. doi: <https://doi.org/10.1007/s10559-014-9589-5>
12. Rudenko, O. G., Bezsonov, O. O., Rudenko, S. O. (2013). Robust identification of nonlinear objects with the help of an evolving radial basis network. *Cybernetics and Systems Analysis*, 49 (2), 173–182. doi: <https://doi.org/10.1007/s10559-013-9497-0>
13. Rudenko, O., Bezsonov, O. (2011). Function Approximation Using Robust Radial Basis Function Networks. *Journal of Intelligent Learning Systems and Applications*, 03 (01), 17–25. doi: <https://doi.org/10.4236/jilsa.2011.31003>
14. Chambers, J. A., Tanrikulu, O., Constantinides, A. G. (1994). Least mean mixed-norm adaptive filtering. *Electronics Letters*, 30 (19), 1574–1575. doi: <https://doi.org/10.1049/el:19941060>
15. Rakesh, P., Kumar, T. K., Albu, F. (2019). Modified Least-Mean Mixed-Norm Algorithms For Adaptive Sparse System Identification Under Impulsive Noise Environment. 2019 42nd International Conference on Telecommunications and Signal Processing (TSP). doi: <https://doi.org/10.1109/tsp.2019.8768813>
16. Papoulis, E. V., Stathaki, T. (2004). A Normalized Robust Mixed-Norm Adaptive Algorithm for System Identification. *IEEE Signal Processing Letters*, 11 (1), 56–59. doi: <https://doi.org/10.1109/lsp.2003.819353>
17. Arenas-García, J., Figueiras-Vidal, A. R. (2005). Adaptive combination of normalised filters for robust system identification. *Electronics Letters*, 41 (15), 874. doi: <https://doi.org/10.1049/el:20051936>
18. Rudenko, O., Bezsonov, O., Lebediev, O., Serdiuk, N. (2019). Robust identification of non-stationary objects with nongaussian interference. *Eastern-European Journal of Enterprise Technologies*, 5 (4 (101)), 44–52. doi: <https://doi.org/10.15587/1729-4061.2019.181256>
19. Walach, E., Widrow, B. (1984). The least mean fourth (LMF) adaptive algorithm and its family. *IEEE Transactions on Information Theory*, 30 (2), 275–283. doi: <https://doi.org/10.1109/tit.1984.1056886>
20. Bershada, N. J., Bermudez, J. C. M. (2011). Mean-square stability of the Normalized Least-Mean Fourth algorithm for white Gaussian inputs. *Digital Signal Processing*, 21 (6), 694–700. doi: <https://doi.org/10.1016/j.dsp.2011.06.002>
21. Eweda, E., Zerguine, A. (2011). New insights into the normalization of the least mean fourth algorithm. *Signal, Image and Video Processing*, 7 (2), 255–262. doi: <https://doi.org/10.1007/s11760-011-0231-y>
22. Eweda, E. (2012). Global Stabilization of the Least Mean Fourth Algorithm. *IEEE Transactions on Signal Processing*, 60 (3), 1473–1477. doi: <https://doi.org/10.1109/tsp.2011.2177976>
23. Eweda, E., Bershada, N. J. (2012). Stochastic Analysis of a Stable Normalized Least Mean Fourth Algorithm for Adaptive Noise Canceling With a White Gaussian Reference. *IEEE Transactions on Signal Processing*, 60 (12), 6235–6244. doi: <https://doi.org/10.1109/tsp.2012.2215607>
24. Hubscher, P. I., Bermudez, J. C. M., Nascimento, V. H. (2007). A Mean-Square Stability Analysis of the Least Mean Fourth Adaptive Algorithm. *IEEE Transactions on Signal Processing*, 55 (8), 4018–4028. doi: <https://doi.org/10.1109/tsp.2007.894423>
25. Zhang, S., Zhang, J. (2015). Fast stable normalised least-mean fourth algorithm. *Electronics Letters*, 51 (16), 1276–1277. doi: <https://doi.org/10.1049/el.2015.0421>
26. Guan, S., Meng, C., Biswal, B. (2019). Optimal step-size of least mean absolute fourth algorithm in low SNR. arXiv. Available at: <https://arxiv.org/ftp/arxiv/papers/1908/1908.08165.pdf>
27. Asad, S. M., Moinuddin, M., Zerguine, A., Chambers, J. (2019). A robust and stable variable step-size design for the least-mean fourth algorithm using quotient form. *Signal Processing*, 162, 196–210. doi: <https://doi.org/10.1016/j.sigpro.2019.04.021>
28. Bin Mansoor, U., Mayyala, Q., Moinuddin, M., Zerguine, A. (2017). Quasi-Newton least-mean fourth adaptive algorithm. 2017 25th European Signal Processing Conference (EUSIPCO). doi: <https://doi.org/10.23919/eusipco.2017.8081689>
29. Sadiq, A., Usman, M., Khan, S., Naseem, I., Moinuddin, M., Al-Saggaf, U. M. (2019). q-LMF: Quantum Calculus-Based Least Mean Fourth Algorithm. *Fourth International Congress on Information and Communication Technology*, 303–311. doi: https://doi.org/10.1007/978-981-15-0637-6_25
30. Zerguine, A., Cowan, C. F. N., Bettayeb, M. (1996). LMS-LMF adaptive scheme for echo cancellation. *Electronics Letters*, 32 (19), 1776. doi: <https://doi.org/10.1049/el:19961202>
31. Zerguine, A., Aboulnasr, T. (2000). Convergence analysis of the variable weight mixed-norm LMS-LMF adaptive algorithm. *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers (Cat. No.00CH37154)*. doi: <https://doi.org/10.1109/acssc.2000.910959>
32. Zerguine, A. (2012). A variable-parameter normalized mixed-norm (VPNMN) adaptive algorithm. *EURASIP Journal on Advances in Signal Processing*, 2012 (1). doi: <https://doi.org/10.1186/1687-6180-2012-55>
33. Rudenko, O., Bezsonov, O., Serdiuk, N., Oliinyk, K., Romanyk, O. (2020). Workable identification of objects based on minimization of combined functional. *Information Processing Systems*, 1 (160), 80–88. doi: <https://doi.org/10.30748/soi.2020.160.10>
34. Price, R. (1958). A useful theorem for nonlinear devices having Gaussian inputs. *IEEE Transactions on Information Theory*, 4 (2), 69–72. doi: <https://doi.org/10.1109/tit.1958.1057444>
35. Gladyshev, E. G. (1965). On Stochastic Approximation. *Theory of Probability & Its Applications*, 10 (2), 275–278. doi: <https://doi.org/10.1137/1110031>
36. Spiegel, M. R., Lipschutz, S., Liu, J. (2008). *Mathematical Handbook of Formulas and Tables*. McGraw Hill.