On Recursive Calculation of M- and GM-Estimates by Direct Identification in LQG Control Systems

Nasko ATANASOV

Technical University, Studentska 1, 9010 Varna, Bulgaria
e-mail: nratanasov@yahoo.com

Rimantas PUPEIKIS

Institute of Mathematics and Informatics
Akademijos 4, LT-08663 Vilnius, Lithuania
e-mail: pupeikis@ktl.mii.lt

Abstract. In the previous papers (Pupeikis, 2000; Genov *et al.*, 2006), a direct approach for estimating the parameters of a discrete-time linear time-invariant (LTI) dynamic system, acting in a closed-loop in the case of additive correlated noise with contaminating outliers uniformly spread in it, is presented. It is assumed here that the parameters of the LQG (Linear Quadratic Gaussian Control) controller are known beforehand. The aim of the given paper is development of a parametric identification approach for a closed-loop system when the parameters of an LTI system as well as that of LQG controller are not known and ought to be estimated. The recursive techniques based on an the M- and GM- estimator algorithms are applied here in the calculation of the system as well as noise filter parameters. Afterwards, the recursive parameter estimates are used in each current iteration to determine unknown parameters of the LQG-controller, too. The results of numerical simulation by computer are discussed.

Keywords: LQG control systems, closed-loop, parametric identification, observations, outliers.

1. Introduction

The stochastic optimal control of a discrete-time LTI dynamic system is performed using the LQG approach (Isermann, 1984; Åström and Wittenmark, 1987; Glad and Ljung, 2000). In the case of various uncertainty reasons due to different sources, the solution to the problem of the worst LQG control system performance, based on a relative entropy constraint uncertainty description, is given in (Petersen, 2006). It has been emphasized here (Petersen, 2006) that in designing a robust control system, one ought to determine the type of uncertainties appearing in the system to be controlled. On the other hand, there are many types of uncertainties in system description models. One of the main ones of them is the uncertainty arising in the output disturbance description of a plant model to be used. It is frequently assumed that output of the system is affected by Gaussian disturbance. However, nonnormal noise, and particularly the presence of outliers, degrades the

performance of a system acting in a closed-loop. Therefore ordinary recursive techniques used for a parametric identification of LQG control systems, as a rule, are inefficient. In such a case, robust recursive techniques ought to be applied here according to (Huber, 1984) and (Pupeikis, 1991).

In what follows, we introduce the robust recursive generalized maximum likelihood (GM) and maximum likelihood (M) procedures for calculating robust estimates of the parameters of LTI dynamic systems, acting in a closed-loop in the case of correlated noise with outliers in it. Note that the class of GM-estimators (Denby and Martin, 1979) contains a class of maximum likelihood type estimators (M-estimators) (Huber, 1964). The class of GM-estimators is defined implicitly by the first order condition (Lucas, 1996)

$$\sum_{t=1}^{N} \mathbf{x}(t)\zeta\big\{\mathbf{x}(t), [y(t) - \mathbf{x}^T(t)\theta]/\sigma\big\} = 0. \tag{1}$$

Here $\mathbf{x}(t)$ is the set of regressors, σ denotes the scale of residuals $\mathbf{n}(t)$ of the linear regression model $y(t) = \mathbf{x}^T(t)\theta + \delta(t)$, $t = 1, \ldots, N$ where θ is a vector of unknown parameters. The function $\zeta\{\cdot, \cdot\}$ in (1) depends on both the set of regressors $\mathbf{x}(t)$ and the standardized residual $\delta(t)/\sigma$. The conditions that ought to be satisfied by $\zeta\{\cdot, \cdot\}$ in order that the GM-estimator have nice asymptotic properties are known in advance (Lucas, 1996). The ordinary least-squares estimator could be obtained as a special case of (1) by setting in it the function $\tau(\mathbf{x}(t), r) = r^2/2$ with $\partial\tau(\mathbf{x}(t), r)/\partial r = \zeta\{\mathbf{x}(t), r\}$, where r is a short form of the standardized residual. In such a case, the class of M-estimators is obtained by setting $\tau(\mathbf{x}(t), r) = \rho(r)$, with $d\rho(r)/dr = \psi(r)$. Various $\psi(\cdot)$ functions lead to various M-estimates. The class of M-estimates was proposed by Huber (1964) for the location parameter and was extended to the regression model (Huber, 1984; Novovičova, 1987).

In Section 2, the statement of the problem is presented. In Section 3, an ordinary direct approach is described for a parametric identification of the system transfer function. In Section 4, ordinary recursive parametric identification techniques are given for estimating the parameters of transfer functions of LTI systems in the case of various noise filters. We analyze the recursive parametric identification, based on M- and GM- estimators in the presence of outliers in output observations, in Section 5. Section 6 presents the simulation and parametric identification results. Section 7 contains conclusions.

2. The Statement of the Problem

Assume that a control system to be observed is causal, linear, and time-invariant with one output $\{y(k)\}$ and one input $\{u(k)\}$, expressed by the equation

$$y(k) = G_0(q^{-1}; \theta)u(k) + \underbrace{H_0(q^{-1}; \varphi)\xi(k)}_{v(k)}, \tag{2}$$

that consists of two parts (Fig. 1): a system model $G_0(q^{-1}; \theta)$ and a noise model $H_0(q^{-1}; \varphi)$. Here k is the current number of observations of a respective signal, θ, φ are
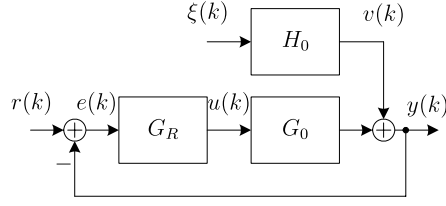
Fig. 1. A closed-loop system to be observed. Here $G_R \equiv G_R(q^{-1}; \alpha)$, $G_0 \equiv G_0(q^{-1}; \theta)$, and $H_0 \equiv H_0(q^{-1}; \varphi)$.

unknown parameter vectors to be estimated, q^{-1} is the backward time-shift operator such that $q^{-1}u(k) = u(k-1)$, $\{\xi(k)\}$ is used to generate unmeasurable noise $\{v(k)\}$ and it is assumed to be statistically independent and stationary with the following characteristics:

$$E\{\xi(k)\} = 0, \quad E\{\xi(k)\xi(k+\tau)\} = \sigma_\xi^2 \delta(\tau), \tag{3}$$

where $E\{\xi(k)\}$ is the mean value, σ_ξ^2 is the variance, $\delta(\tau)$ is the Kronecker delta function, and $H_0(q^{-1}, \varphi)$ is an inversely stable monic filter (Forsell and Ljung, 1999).

The input $\{u(k)\}$ is given by

$$u(k) = \underbrace{[r(k) - y(k)]}_{e(k)} G_R(q^{-1}; \alpha), \tag{4}$$

where the reference signal $\{r(k)\}$ is a quasi-stationary signal, independent of the stochastic disturbance $\{v(k)\}$, and the LQG controller $G_R(q, \alpha)$, which is designed by minimizing the quadratic performance function

$$\mathbf{J} = \lim_{N \to \infty} E\left\{ \frac{1}{N} \sum_{k=0}^{N-1} \left[y^2(k) + \varrho u^2(k) \right] \right\}, \tag{5}$$

is exponentially stable (Forsell and Ljung, 1999). Here α is the parameter vector of the controller $G_R(q^{-1}; \alpha)$ and ϱ is a weighting factor, the values of which are given in (Halwass, 1988).

The aim of the given paper is to estimate the parameter vector θ of the LTI system $G_0(q^{-1}; \theta)$, acting in the closed-loop (see Fig. 1) simultaneously with the current parameter vector α of the LQG controller $G_R(q^{-1}; \alpha)$, by observations $\{u(k), y(k)\}$ $\forall k = 1, 2, \ldots$, in the case of additive correlated noise $\{v(k)\}$, that contains large outliers and corrupts the output $\{y(k)\}$ of the system.

3. The Direct Approach

The direct approach ignores the feedback and identifies the system $G_0(q^{-1}; \theta)$ using measurements of the input $u(k)$ and the output $y(k)$ $\forall k = 1, 2, \ldots$ (Forsell and Ljung, 1999).

By the direct parameric identification method one has to estimate the prediction error value $\hat{\theta}_N$ of the vector of parameters θ by

$$\hat{\theta}_N = \arg \min_{\theta \in D_M} V_N(\theta, Z^N). \tag{6}$$

Here

$$V_N(\theta, Z^N) = \frac{1}{N} \sum_{k=1}^{N} e_F^T(k, \theta) \mathbf{\Lambda}^{-1} e_F^T(k, \theta), \tag{7}$$

with

$$e_F(k, \theta) = L(q, \theta)\epsilon(k, \theta), \tag{8}$$

$\mathbf{\Lambda}$ is a symmetric, positive definite weighting matrix and $L(q^{-1}; \theta)$ is a monic prefilter that can be used to enhance certain frequency regions (Forsell and Ljung, 1999). The prediction error is calculated by

$$\epsilon(k, \theta) = y(k) - \hat{y}(k, \hat{\theta}) = H^{-1}(q^{-1}; \hat{\varphi})\big[y(k) - G(q^{-1}; \hat{\theta})u(k)\big]. \tag{9}$$

Here the output $y(k)$ of the general model of the LTI system $G(q^{-1}; \theta)$ and noise filter $H(q^{-1}; \varphi)$, respectively, are in the form

$$y(k) = G(q^{-1}; \theta)u(k) + H(q^{-1}; \varphi)\xi(k), \tag{10}$$

where $G(q^{-1}; \theta)$ corresponds to the first part of the Eq. (2) and $H(q^{-1}; \varphi)$ to the second one. Then, the one-step-ahead predictor for the model structure (10) is

$$\hat{y}(k, \hat{\theta}) = H^{-1}(q^{-1}; \hat{\varphi})G(q^{-1}; \hat{\theta})u(k) + \big[1 - H^{-1}(q^{-1}; \hat{\varphi})\big]y(k). \tag{11}$$

Here $\hat{\varphi}$ is the estimate of the parameters vector φ. The parameter vector θ can be determined by the ordinary prediction error method, based on the recursive LS (RLS) of the form

$$\begin{aligned}
\hat{\theta}(k) &= \hat{\theta}(k-1) + \frac{\mathbf{\Gamma}(k-1)\mathbf{z}(k)}{1 + \mathbf{z}^T(k)\mathbf{\Gamma}(k-1)\mathbf{z}(k)}\hat{\varepsilon}(k), \\
\mathbf{\Gamma}(k) &= \mathbf{\Gamma}(k-1) - \frac{\mathbf{\Gamma}(k-1)\mathbf{z}(k)\mathbf{z}^T(k)\mathbf{\Gamma}(k-1)}{1 + \mathbf{z}^T(k)\mathbf{\Gamma}(k-1)\mathbf{z}(k)}
\end{aligned} \tag{12}$$

with the vector of observations $\mathbf{z}^T(k) = [-y(k{-}1), \ldots, -y(k{-}m), u(k{-}1), \ldots, u(k{-}m)]$, and some initial values of the vector $\hat{\theta}(0)$ and matrix $\mathbf{\Gamma}(0)$. Here

$$\hat{\theta}^T(k) = \big[\hat{\mathbf{a}}^T(k), \hat{\mathbf{b}}^T(k)\big] = \big[\hat{a}_1(k), \ldots, \hat{a}_m(k), \hat{b}_0(k), \hat{b}_1(k), \ldots, \hat{b}_m(k)\big], \tag{13}$$

is the current estimate of the vector $\theta^T = (\mathbf{a}^T, \mathbf{b}^T) = (a_1, \ldots, a_m, b_0, b_1, \ldots, b_m)$, and

$$\hat{\varepsilon}(k) = y(k) - \mathbf{z}^T(k)\hat{\theta}(k-1) \tag{14}$$

is the prediction error on the current k-th iteration, respectively, where $G_0(q^{-1}; \theta)$ is the system transfer function of the form

$$G_0(q^{-1}; \theta) = \frac{B(q^{-1}; \mathbf{b})}{A(q^{-1}; \mathbf{a})} = \frac{b_0 + b_1 q^{-1} + b_2 q^{-2} + \cdots + b_m q^{-m}}{1 + a_1 q^{-1} + \cdots + a_m q^{-m}}. \tag{15}$$

Here $\mathbf{b}^T = (b_0, b_1, \ldots, b_m)$, and $\mathbf{a}^T = (a_1, \ldots, a_m)$ are vectors of parameters to be estimated.

It is known that RLS is efficient only in the case where

$$H_0(q^{-1}; \varphi) = \frac{1}{1 + A(q^{-1}; \mathbf{a})} = \frac{1}{1 + a_1 q^{-1} + \cdots + a_m q^{-m}}, \quad \varphi \equiv \mathbf{a}. \tag{16}$$

It could be emphasized that, before the closed-loop direct parametric identification the respective identifiability conditions should be satisfied according to (Isermann, 1984).

4. Recursive Estimators for the Various Noise Filters

In practice, the assumption (16) is invalid as a rule, and the ordinary RLS is of little use. Therefore a multivariate set of recursive techniques is worked out. In the case where

$$H(q^{-1}; \varphi) = 1 \tag{17}$$

additive noise $\{v(k)\} \equiv \{\xi(k)\}$, and the recursive algorithm, based on the technique, is developed in (Steiglitz and McBride, 1965), may be used. Then, in formulas (12) and (14) the vector of observations $\mathbf{z}^T(k)$ and the equation error $\hat{\varepsilon}(k)$ must be replaced by

$$\mathbf{z}^{*T}(k) = \big[-y^*(k-1), \ldots, -y^*(k-m), u^*(k-1), \ldots, u^*(k-m) \big], \tag{18}$$

and

$$\hat{\varepsilon}^*(k) = y^*(k) - \mathbf{z}^{*T}(k)\hat{\theta}(k-1), \tag{19}$$

respectively. Here

$$y^*(k) = \frac{1}{\hat{A}_{k-1}(q^{-1}; \hat{\mathbf{a}})} y(k) = \frac{1}{1 + \hat{a}_{1(k-1)}q^{-1} + \cdots + \hat{a}_{m(k-1)}q^{-m}} y(k), \tag{20}$$

$$u^*(k) = \frac{1}{\hat{A}_{k-1}(q^{-1}; \hat{\mathbf{a}})} u(k) = \frac{1}{1 + \hat{a}_{1(k-1)}q^{-1} + \cdots + \hat{a}_{m(k-1)}q^{-m}} u(k), \tag{21}$$

are filtered observations after k samples; $\hat{\mathbf{a}}$ is the estimate of $\mathbf{a}$; $\hat{a}_{1(k-1)}, \ldots, \hat{a}_{m(k-1)}$ are estimates of $a_1, \ldots, a_m$, respectively, calculated in the recursive iteration $(k-1)$. The characteristics and convergence conditions of the above mentioned algorithm were investigated by Stoica and Söderström (1981).

Let us assume now that

$$H(q^{-1}; \varphi) = \frac{1}{[1 + G(q^{-1}; \mathbf{g})][1 + A(q^{-1}; \mathbf{a})]}. \tag{22}$$

Here

$$G(q^{-1}; \mathbf{g}) = g_1 q^{-1} + \cdots + g_{n\,g} q^{-n_g}, \quad \varphi \equiv (\mathbf{g}, \mathbf{a}) \tag{23}$$

with the known integer n_g. In such a case the generalized RLS (GRLS) algorithm, consisting of two RLS algorithms, is used for estimating the vector of unknown parameters $\theta^T = (\mathbf{a}^T, \mathbf{b}^T) = (a_1, \ldots, a_m, b_0, b_1, \ldots, b_m)$, and $\mathbf{g}^T = (g_1, \ldots, g_{n\,g})$ according to (Hastings–James and Sage, 1969). The first RLS algorithm calculates $\hat{\theta}$ by recursive formulas (12), where

$$e^*(k) = y^*(k) - \mathbf{z}^{*T}(k)\hat{\theta}(k-1), \tag{24}$$

$$\mathbf{z}^{*T}(k) = \big[-y^*(k-1), \ldots, -y^*(k-m), u^*(k-1), \ldots, u^*(k-m)\big], \tag{25}$$

$$y^*(k) = \big[1 + \hat{G}_{k-1}(q^{-1}; \hat{\mathbf{g}})\big] y(k)$$
$$= \big[1 + \hat{g}_{1(k-1)}q^{-1} + \cdots + \hat{g}_{n\,g(k-1)}q^{-n_g}\big] y(k), \tag{26}$$

$$u^*(k) = \big[1 + \hat{G}_{k-1}(q^{-1}; \hat{\mathbf{g}})\big] u(k)$$
$$= \big[1 + \hat{g}_{1(k-1)}q^{-1} + \cdots + \hat{g}_{n\,g(k-1)}q^{-n_g}\big] u(k), \tag{27}$$

$$\hat{G}_{k-1}(q^{-1}; \hat{\mathbf{g}}) = \hat{g}_{1(k-1)}q^{-1} + \cdots + \hat{g}_{n\,g(k-1)}q^{-n_g}. \tag{28}$$

The second algorithm calculates the vector of the estimates $\hat{\mathbf{g}}^T = (\hat{g}_{1\,k}, \ldots, \hat{g}_{n\,g\,k})$ using recursive equations of the form

$$\hat{\mathbf{g}}(k) = \hat{\mathbf{g}}(k-1) + \boldsymbol{\Lambda}(k)\mathbf{v}(k)\tilde{\varepsilon}(k),$$

$$\boldsymbol{\Lambda}(k) = \boldsymbol{\Lambda}(k-1) - \frac{\boldsymbol{\Lambda}(k-1)\mathbf{v}(k)\mathbf{v}^T(k)\boldsymbol{\Lambda}(k-1)}{1 + \mathbf{v}^T(k)\boldsymbol{\Lambda}(k-1)\mathbf{v}(k)}, \tag{29}$$

$$\tilde{\varepsilon}(k) = \big[1 + \hat{G}_{k-1}(q^{-1}; \hat{\mathbf{g}})\big]\hat{\varepsilon}(k), \tag{30}$$

where $\mathbf{v}^T(k) = (\hat{\varepsilon}(k), \ldots, \hat{\varepsilon}(k - n_q))$, $\hat{\varepsilon}(k)$ is of the form (14). The initial conditions for RGLS can be chosen according to (Clarce, 1967).

Further, suppose that

$$H(q^{-1}; \varphi) = \frac{1 + F(q^{-1}; \mathbf{f})}{1 + A(q^{-1}; \mathbf{a})}. \tag{31}$$

Here

$$F(q^{-1}; \mathbf{f}) = f_1 q^{-1} + \cdots + f_{n\,f} q^{-n_f} \tag{32}$$

with the known integer n_f, $\varphi \equiv (\mathbf{f}, \mathbf{a})$. Then the estimates of vectors of the parameters $\theta^T = (\mathbf{a}^T, \mathbf{b}^T)$, and $\mathbf{f}^T = (f_1, \ldots, f_{n_f})$ can be calculated using the maximum likelihood method (ML) or its on-line version consisting of two recursive algorithms, described in (Åström and Bohlin, 1965). The first recursive ML algorithm (RML) calculates the vector of the estimates $\hat{\theta}^T = (\hat{\mathbf{a}}^T, \hat{\mathbf{b}}^T)$ according to formulas (12), where $e^*(k)$ and $\mathbf{z}^{*T}(k)$ are of the form (24) and (25), respectively, with the exeption of

$$
\begin{aligned}
y^*(k) &= \left[1 + \hat{F}_{k-1}(q^{-1}; \hat{\mathbf{f}})\right]^{-1} y(k) \\
&= \left(1 + \hat{f}_{1(k-1)}q^{-1} + \cdots + \hat{f}_{n\,f(k-1)}q^{-n_f}\right)^{-1} y(k),
\end{aligned} \tag{33}
$$

$$
\begin{aligned}
u^*(k) &= \left[1 + \hat{F}_{k-1}(q^{-1}; \hat{\mathbf{f}})\right]^{-1} u(k) \\
&= \left(1 + \hat{f}_{1(k-1)}q^{-1} + \cdots + \hat{f}_{n\,f(k-1)}q^{-n_f}\right)^{-1} u(k).
\end{aligned} \tag{34}
$$

Here

$$
\hat{F}_{k-1}(q^{-1}; \hat{\mathbf{f}}) = \hat{f}_1 q^{-1} + \cdots + \hat{f}_{n\,f} q^{-n_f}. \tag{35}
$$

The second algorithm calculates the vector of the estimates $\hat{\mathbf{f}}^T = (\hat{f}_1, \ldots, \hat{f}_{n\,f})$ using recursive equations of the form

$$
\hat{\mathbf{f}}(k) = \hat{\mathbf{f}}(k-1) + \mathbf{\Pi}(k)\mathbf{v}(k)\hat{\varepsilon}(k), \tag{36}
$$

$$
\mathbf{\Pi}(k) = \mathbf{\Pi}(k-1) - \frac{\mathbf{\Pi}(k-1)\mathbf{v}(k)\mathbf{v}^T(k)\mathbf{\Pi}(k-1)}{1 + \mathbf{v}^T(k)\mathbf{\Pi}(k-1)\mathbf{v}(k)},
$$

$$
\varepsilon^*(k) = \left[1 + \hat{F}_{k-1}(q^{-1}; \hat{\mathbf{f}})\right]^{-1} \hat{\varepsilon}(k), \tag{37}
$$

where $\mathbf{v}^T(k) = (\varepsilon^*(k), \ldots, \varepsilon^*(k - n_q))$, and $\hat{\varepsilon}(k)$ is of the form (14). The initial conditions for RML can be chosen according to (Isermann, 1974; Eykhoff, 1975).

In a more general case (Van den Boom, 1981), i.e., where

$$
H_0(q^{-1}; \varphi) = \frac{(1 + F(q^{-1}; \mathbf{f}))}{(1 + G(q^{-1}; \mathbf{g}))(1 + A(q^{-1}; \mathbf{a}))}, \varphi \equiv (\mathbf{f}, \mathbf{g}, \mathbf{a}) \tag{38}
$$

the recursive algorithm of an extended least squares method (ERLS) can be used, which for $F(q^{-1}; \mathbf{f}) \equiv 0$ turned into GRLS anf for $G(q^{-1}; \mathbf{g}) \equiv 0$ – into RML. There also exist algorithms of stochastic approximation which are usually recursive procedures with a scalar step (Saridis, 1974). The convergence of the recursive techniques on the basis of a united approach is investigated in (Ljung, 1977). The simulation results of the above mentioned algorithms are given in (Isermann *et al.*, 1974).

5. Identification in the Presence of Outliers

Given the model (2) and measured data $\mathbf{Z}^N = \{u(1), \ldots, u(N), y(1), \ldots, y(N)\}$ and assuming that the white noise $\{\xi(k)\}, k = 1, 2, \ldots$ is really a sequence of independent

identically distributed variables with an ϵ-contaminated distribution of the form

$$p(\xi(k)) = (1 - \epsilon)N(0, \sigma_\mu^2) + \epsilon N(0, \sigma_\varsigma^2), \tag{39}$$

and the variance

$$\sigma_\xi^2 = (1 - \epsilon)\sigma_\mu^2 + \epsilon \sigma_\varsigma^2, \tag{40}$$

one can determine the prediction error estimate $\hat{\theta_N}$ of the parameter vector $\theta^T = (\mathbf{a}^T, \mathbf{b}^T) = (a_1, \ldots, a_m, b_0, b_1, \ldots, b_m)$ by minimizing

$$\hat{\theta}_N = \arg \min_{\theta \in D_M} \widetilde{V}_N(\theta, \mathbf{Z}^N) \tag{41}$$

with

$$\widetilde{V}_N(\theta, \mathbf{Z}^N) = \frac{1}{N} \sum_{k=1}^{N} \rho\big(e_F(k, \theta/s)\big), \tag{42}$$

or by solving the equation

$$\sum_{t=1}^{N} \mathbf{z}(t)\big\{\psi\big[y(t) - \mathbf{z}^T(t)\theta\big]\big\} = 0, \tag{43}$$

in the vector form. Here $p\{\xi(k)\}$ is the probability density distribution of the sequence $\{\xi(k)\}$, $k = 1, 2, \ldots$;

$$\xi(k) = (1 - \gamma_k)\mu_k + \gamma_k \varsigma_k \tag{44}$$

is the value of the sequence $\{\xi(k)\}$, $k = 1, 2, \ldots$ at a time moment k; γ is a random variable, taking values 0 or 1 with probabilities $p(\gamma_k = 0) = 1 - \epsilon$, $p(\gamma_k = 1) = \epsilon$; μ_k, ς_k are sequences of independent Gaussian variables with zero means and variances $\sigma_\mu^2, \sigma_\varsigma^2$, respectively; besides, $\sigma_\mu < \sigma_\varsigma$; $0 \leqslant \epsilon \leqslant 1$ is the unknown fraction of contamination; θ_N is the robust estimate of the parameter vector θ, established by processing N pairs of input-output samples; s is the scale of residual (examples of the scale are the standard deviation, the median, absolute deviation from the median, etc.,); $\rho(\cdot)$ is a real-valued function that is even and nondecreasing for positive residuals, and $\rho(0) = 0$, $\psi = \rho'$.

For the Huber M-estimator, the ρ-function is given by

$$\rho(x) = \begin{cases} x^2/2 & \text{if } \mid x \mid \leqslant c_H, \\ c_H \mid x \mid - c_H^2/2 & \text{if } \mid x \mid > c_H, \end{cases} \tag{45}$$

where c_H is a cutoff value. The mostly used function ψ is (Huber, 1964):

$$\psi(x) = \begin{cases} x & \text{if } \mid x \mid \leqslant c_H, \\ c_H \text{sign}(x) & \text{if } \mid x \mid > c_H, \end{cases} \tag{46}$$

with given $c_H > 0$. To get a better performance of $\hat{\theta}_N$ in the case of very long-tailed distributions, a function (42) satisfying $\psi(x) = 0$, if $\mid x \mid > c_H$, for some $c_H > 0$ could be selected. It is known (Novovičova, 1987) that, in both such cases, i.e., $\epsilon \neq 0$ and $H_0(q^{-1}; \varphi)$ of the form (16), the current M-estimates of an unknown vector of the parameters θ of LTI system (2) with $G(q, \theta)$ of the form (15) can be calculated using three techniques: the S-algorithm, the H-algorithm, and the W-one. All the three of them could be written in the general form:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + \frac{\mathbf{\Gamma}(k-1)\mathbf{z}(k)}{\lambda(k) + \mathbf{z}^T(k)\mathbf{\Gamma}(k-1)\mathbf{z}(k)}\beta(k), \tag{47}$$

$$\mathbf{\Gamma}(k) = \mathbf{\Gamma}(k-1) - \frac{\mathbf{\Gamma}(k-1)\mathbf{z}(k)\mathbf{z}^T(k)\mathbf{\Gamma}(k-1)}{\lambda(k) + \mathbf{z}^T(k)\mathbf{\Gamma}(k-1)\mathbf{z}(k)}.$$

Here

$$\beta(k) = \hat{s}\psi[\alpha(k)] \tag{48}$$

with

$$\alpha(k) = \hat{\varepsilon}(k)/\hat{s} \tag{49}$$

for S- and H-algorithms, and

$$\beta(k) = \hat{s}\hat{\varepsilon}(k) \tag{50}$$

for the W-algorithm;

$$\hat{\varepsilon}(k)/\hat{s} = \left\{y(k) - \mathbf{z}^T(k)\hat{\theta}(k-1)\right\}/\hat{s} \tag{51}$$

is the same for all the three algorithms, while

$$\lambda(k) = 1 \tag{52}$$

for the H-algorithm

$$\lambda(k) = \begin{cases} \left\{\hat{s}\psi[\alpha(k)]/\hat{\varepsilon}(k)\right\}^{-1} & \text{for } \hat{\varepsilon}(k) \neq 0, \\ 1 & \text{for } \hat{\varepsilon}(k) = 0, \end{cases} \tag{53}$$

for the W-algorithm,

$$\lambda(k) = \psi'\left[\alpha(k)\right]^{-1} \tag{54}$$

for the S-algorithm. Here $\hat{s}$ is the robust estimate of the scale s of residuals. In (Genov *et al.*, 2006) it has been proposed to use

$$\beta(k) = \hat{s}\phi_{z1}\psi\left[\alpha(k)/\phi_{z2}\right], \tag{55}$$

and

$$\lambda(k) = \begin{cases} \phi_{z1}\psi\big[\alpha(k)/\phi_{z2}\big] \big/ \big[\alpha(k)/\phi_{z2}\big] & \text{for } \alpha(k) \neq 0, \\ \phi_{z1} & \text{for } \alpha(k) = 0, \end{cases} \tag{56}$$

respectively, instead of (48) and (53). Here

$$\phi_{z1} = \phi_{z2} = 1 \tag{57}$$

for Hubers M-estimator;

$$\phi_{z1} = \phi_z[h(k)], \quad \phi_{z2} = 1 \tag{58}$$

for Mallow's, and

$$\phi_{z1} = \phi_{z2} = \phi_z[h(k)], \tag{59}$$

for Shweppe's GM-estimators, respectively, where

$$\phi_z[h(k)] = \sqrt{1 - h(k)} \tag{60}$$

with

$$h(k) = \mathbf{z}^T(k)\mathbf{\Gamma}(k)\mathbf{z}(k). \tag{61}$$

The S-algorithm represents a version of the algorithm proposed by Polyak and Tsypkin (1980) for an on-line robust identification of parameters of the linear dynamic model of the LTI system. The ordinary RLS (12) is modified by substituting the "winsorization" step of the residuals in the first equation and changing the second equation in Eqs. (12). The recursive H-algorithm is obtained only by inserting the "winsorization" step into the first equation of Eqs. (12). The W-algorithm is worked out by inserting different weights in respect to the function $\psi\{\cdot\}$ into the already existing ordinary RLS.

It is not difficult to show that, if $H_0(q^{-1};\varphi)$ is of the form (17), then in formulas (48), (51) the vector $\mathbf{z}^T(k)$ must be replaced by (18) and $\hat{\varepsilon}(k)$ of the form (14) by $\hat{\varepsilon}^*(k)$ of the form (19), where $y^*(k)$ and $u^*(k)$ are observations filtered according to (20) and (21), respectively. If $H_0(q^{-1};\varphi)$ is of the form (22), then the first RLS can be replaced by the generalized algorithm of the form (48)–(51) when calculating the vector of the estimates $\hat{\theta}(k)$ according to recursive formulas (12). Here the vector $\mathbf{z}^T(k)$ is replaced by (25) and $\hat{\varepsilon}(k)$ by $e^*(k)$ of the form (24). The observations $y^*(k)$ and $u^*(k)$ are filtered according to (26) and (27), respectively. The same preparations must be made when $H_0(q^{-1};\varphi)$ is of the form (31) and the ordinary RML is used. The difference is that the observations $y^*(k)$ and $u^*(k)$ in $\hat{\varepsilon}^*(k)$ and in the vector $\mathbf{z}^T(k)$ are filtered according to (33), (34), respectively. The matrix $\Gamma(k)$ for each case of $H_0(q^{-1};\varphi)$ is calculated by the second expression of (48) with $\lambda(k)$ that is different for each robust technique to be chosen.

Recursive expressions (29), and (37) are robustified by multiplying $\mathbf{\Lambda}(k)\mathbf{v}(k)\tilde{\varepsilon}(k)$ and $\mathbf{\Pi}(k)\mathbf{v}(k)\hat{\varepsilon}(k)$, respectively, by $\beta(k)$, in their first equations, and substituting the current values of $\lambda(k)$ instead of unities in the corresponding denominators in their second equations.

6. Simulation Example

A closed-loop system to be simulated is shown in Fig. 2 and described by a linear difference equation of the form (Åström, 1987; Halwass, 1988)

$$(1 + a_1 q^{-1})y(k) = (1 + b_1 q^{-1})u(k) + (1 + c_1 q^{-1})\xi(k), \tag{62}$$

while the controller design equation is

$$u(k) = e(k) + w_1 u(k-1) + w_2 u(k-2), \tag{63}$$

where

$$e(k) = r(k) - y(k). \tag{64}$$

Here $a_1 = -0.985, b_1 = 2$ and $c_1 = -0.7$. The coefficients of the LQG controller are found according to (Halwass, 1988) by the formulas:

$$w_1 = p + c_1 - a_1; \tag{65}$$

and

$$w_2 = \frac{(p - a_1)(c - a_1)}{b_1 - a_1}b, \tag{66}$$

assuming that time delay is equal to zero and $w_0 = 1$. Then such values of coefficients of the LQG controller described by Eqs. (65) and (66) are calculated beforehand: $w_1 = 0.1005, w_2 = -0.1016$. The output $y(k), k = 0, 1, 2, \ldots$ of the closed-loop system is observed without and under the additive noise $v(k), k = 0, 1, 2, \ldots$ containing outliers according to (39)–(40) (see Figs. 3, 4, 6, 8). We calculate estimates of the parameters a_1, b_1, c_1 of Eq. (62) by processing observations of $\{y(k)\}$ and $\{u(k)\}$ in each current iteration k using ordinary RML (see Figs. 5, 7, 9), the H-algorithm with version of M-estimator of Huber (Fig. 10), and the S-algorithm with version of Shweppe's GM-estimator (Fig. 11). It could be mentioned that both robust recursive procedures were applied only in the presence of outliers in the observations $\{y(k)\}$, while the ordinary RML was also used, in the opposite case. Afterwar ds, in each current iteration k the coefficients w_1, w_2 were determined using formulas (65)–(66) despite which recursive estimation technique of the parameters a_1, b_1, c_1 was used. The parameter estimation results are presented in Figs. 5, 7, 9–11. It follows that the accuracy of estimates of the
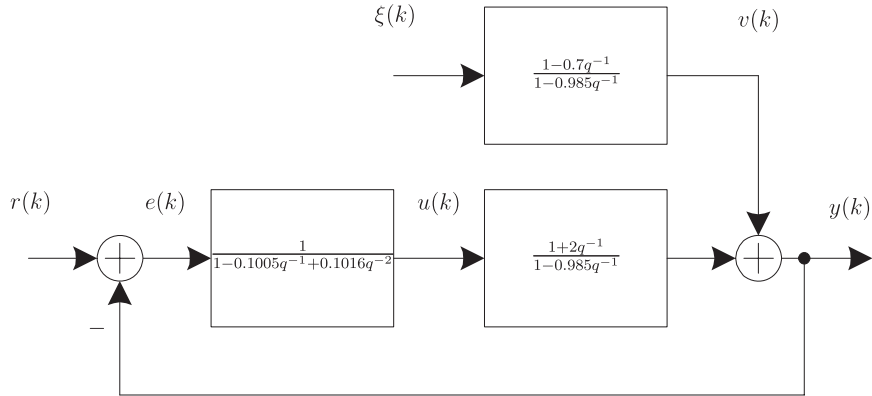
Fig. 2. A closed-loop system to be simulated. Signals: $r(k)$ is a reference signal, $u(k)$ is input, $y(k)$ is output, $v(k)$ is an additive correlated noise, $\xi(k)$ is a sequence of independent identically distributed variables with an ϵ-contaminated distribution of the form (39), and $e(k) = r(k) - u(k)$ is an error.
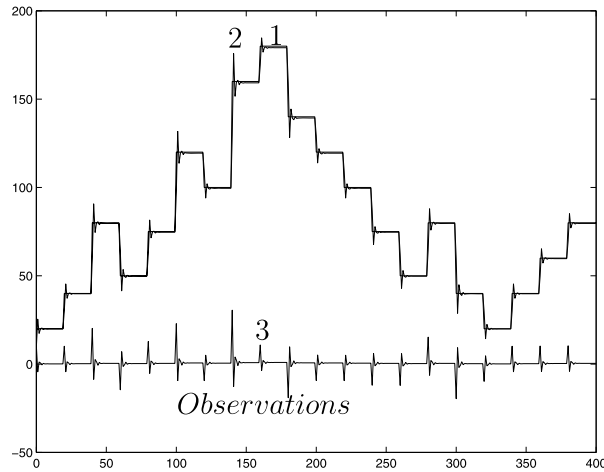

Fig. 3. Signals of the noiseless closed-loop system: 1) the reference signal $r(k)$, 2) output $y(k)$, 3) input $u(k)$.

parameters a_1, b_1, c_1 obtained by ordinary RML (see Figs. 5a, b, c, 7a, b, c, and 9a, b, c) decreases when the number of outliers in additive noise $\{v(k)\}$ is increasing. Thus, the ordinary RML turns out to be inefficient. In such a case, the accuracy of estimates of the coefficients w_1, w_2 of the LQG controller is decreasing as well. Therefore it is important to use here robust recursive techniques. The results of numerical simulation and estimation (Figs. 10, 11) allow us only to approximately consider the quality of estimates. Therefore, it is important to present also the averaged estimates of unknown parameters and quantitative indicators of the calculated estimates, i.e., the confidence intervals of the estimates with respect to the number of iterations.

10 experiments with different realizations of additive correlated noise $\{v(k)\}$ were carried out in order to investigate more precisely and to compare the accuracy of es-
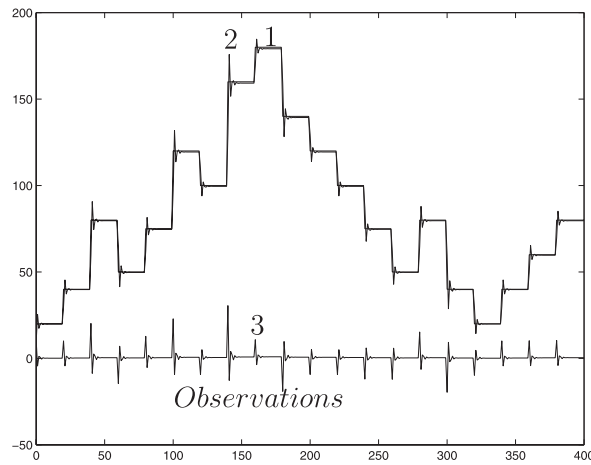
Fig. 4. Signals of a noisy closed-loop system in the absence of outliers in $v(k)$: 1) the reference signal $r(k)$, 2) noisy output $y(k)$, 3) input $u(k)$.
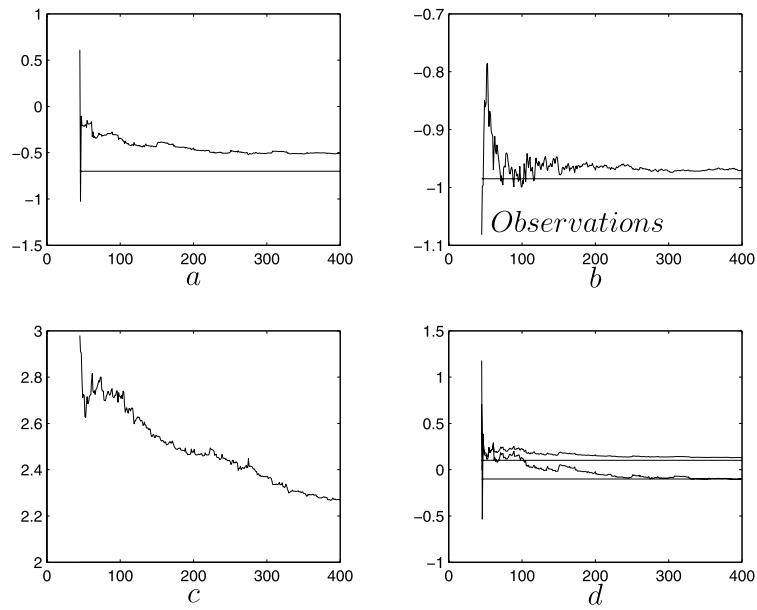

Fig. 5. Dependence of current estimates of the parameters of the closed-loop system on the number of recursive iterations when outliers in $v(k)$ are absent (see Fig. 4): x-axis – numbers of iterations, y-axis – meanings of the estimates. a is the estimate of the coefficient of the numerator of the noise filter transfer function, b is the estimate of the coefficient of the denominator of the system transfer function, c is the estimate of the coefficient of the numerator of the system transfer function, d are estimates of the coefficients of the denominator of the controller transfer function. Straight lines correspond to the exact values of estimated coefficients. The estimates are obtained using the ordinary RML.
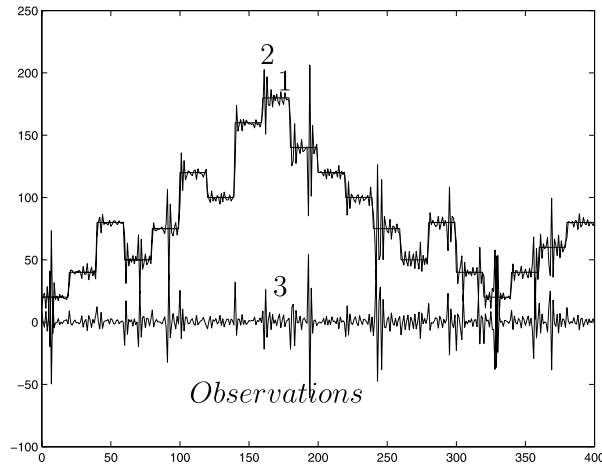
Fig. 6. Signals of a noisy closed-loop system in the presence of outliers in $v(k)$(ϵ=0.05 in (39)): 1) the reference signal $r(k)$, 2) the output $y(k)$ corrupted with an additive noise with outliers in it, 3) the input $u(k)$.
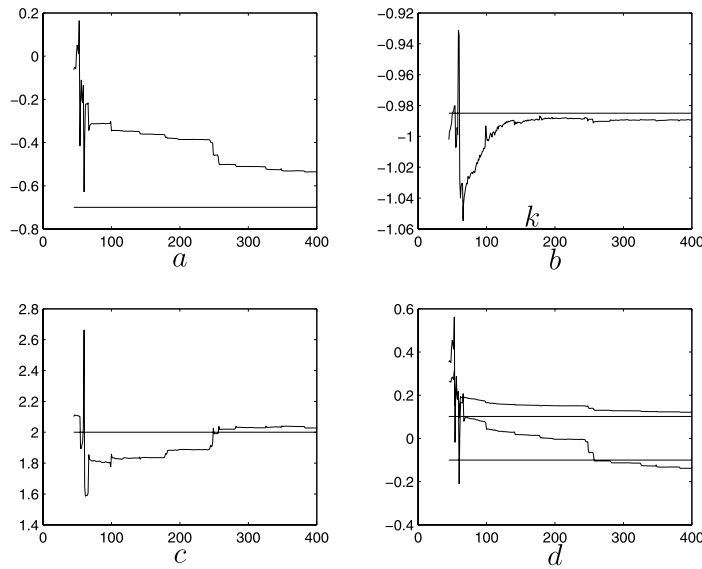

Fig. 7. Estimates of noisy closed-loop system parameters in the case of outliers in a correlated noise (see Fig. 6). Other values and markings are the same as in Fig. 5. The estimates are obtained using the ordinary RML.

timates of the parameter vector θ of the LTI system $G_0(q^{-1};\theta)$ simultaneously with the current parameter vector α of the LQG controller $G_R(q^{-1};\alpha)$, obtained using the H-algorithm with a version of Huber's M-estimator and S-algorithm with version of Shweppe's GM-estimator. We have used the Monte Carlo simulation with 10 data sets, each containing 400 input-output observation pairs in the case of additive correlated noise $\{v(k)\}$, having large outliers and corrupting the output $\{y(k)\}$ (see Fig. 12). In
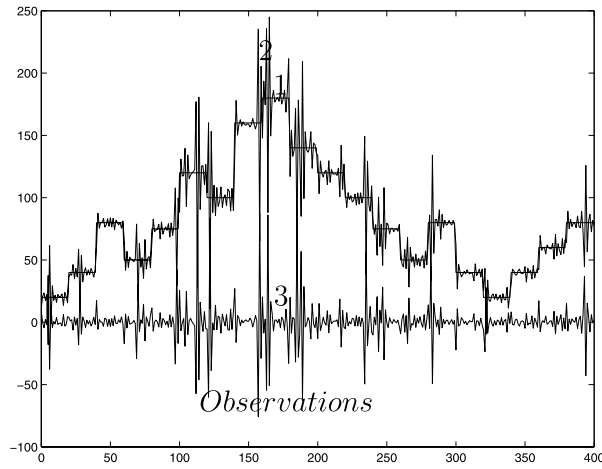
Fig. 8. Signals of a noisy closed-loop system in the presence of outliers in $v(k)(\epsilon=0.1)$ in (39): 1) the reference signal $r(k)$, 2) the output $y(k)$ corrupted by an additive noise with outliers in it, 3) the input $u(k)$.
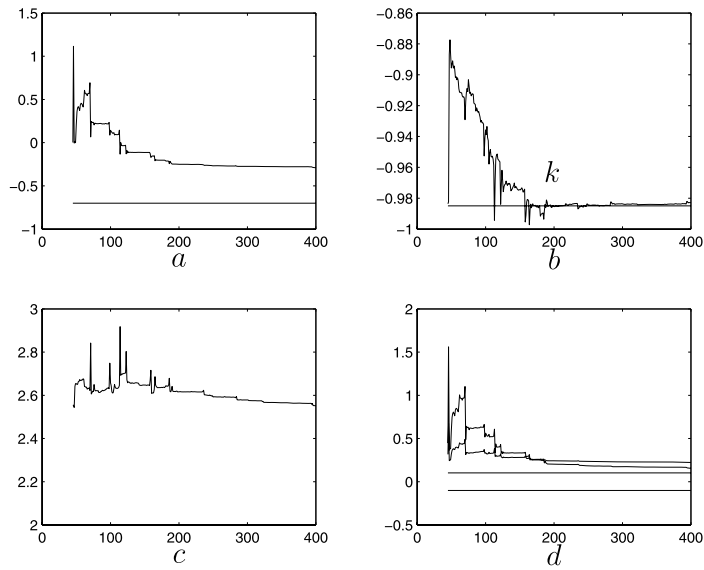

Fig. 9. Estimates of noisy closed-loop system parameters in the case of outliers in a correlated noise (see Fig. 8). Other values and markings are the same as in Fig. 5. The estimates are obtained using the ordinary RML.

each ith experiment the estimates of parameters $a_1 = -0.985$, $b_1 = 2$, $c_1 = -0.7$, and $w_1 = 0.1005$, $w_2 = -0.1016$ have been determined. Table 1 illustrates the values $\bar{b}_1, \bar{a}_1, \bar{c}_1$ of estimates $\hat{b}_1(k), \hat{a}_1(k), \hat{c}_1(k)$, (averaged by 10 experiments), and their
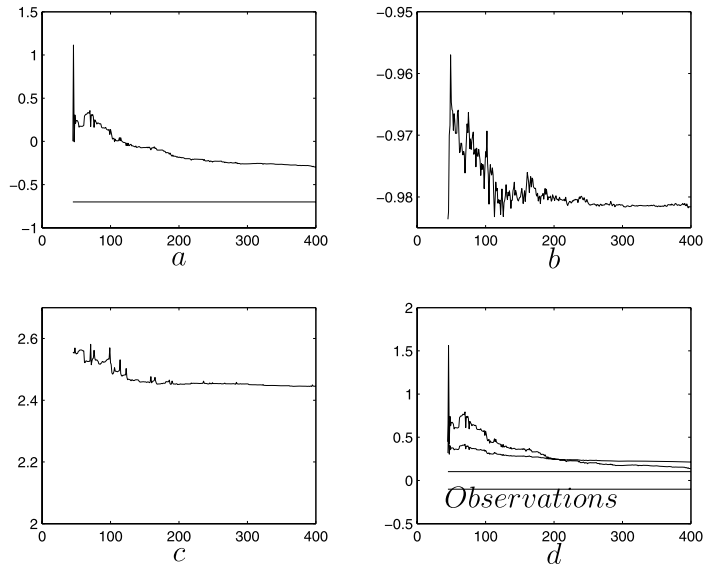
Fig. 10. Estimates of noisy closed-loop system parameters in the case of outliers in a correlated noise (see Fig. 8). Other values and markings are the same as in Fig. 5. The estimates are obtained using the H-algorithm with a version of Huber's M-estimator.
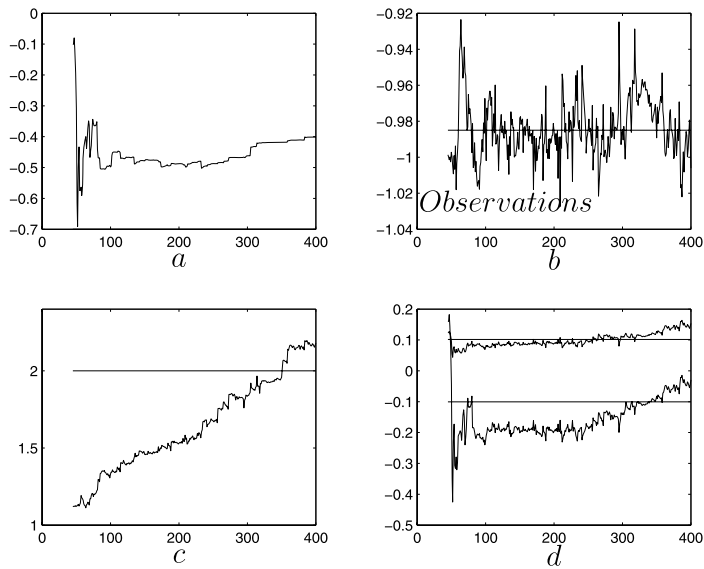

Fig. 11. Estimates of noisy closed-loop system parameters in the case of outliers in a correlated noise (see Fig. 8). Other values and markings are the same as in Fig. 5. The estimates are obtained using the S-algorithm with Shweppe's GM-estimator.
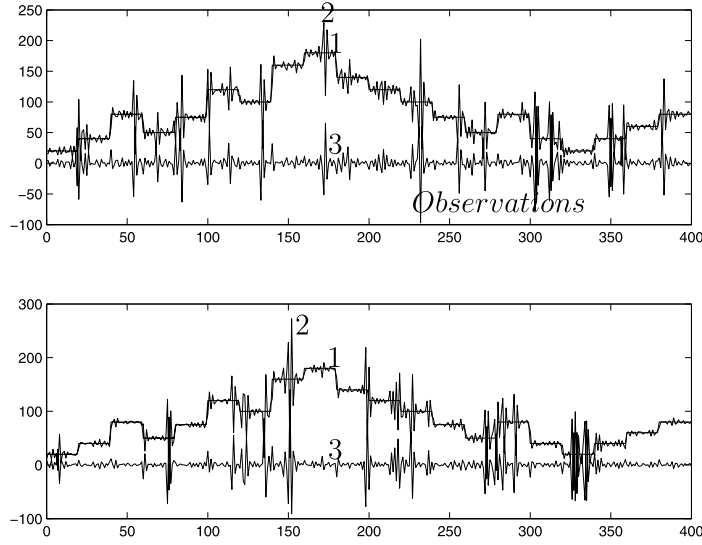
Fig. 12. Two different cases of realizations of respective signals in the case of outliers in a correlated noise $\{v(k)\}$, used in the Monte Carlo simulation. Other values and notation are the same as in Fig. 3.

confidence intervals

$$\Delta_1 = \pm t_\alpha \frac{\hat{\sigma}_{b_1}}{\sqrt{N}}, \quad \Delta_1 = \pm t_\alpha \frac{\hat{\sigma}_{a_1}}{\sqrt{N}}, \quad \Delta_2 = \pm t_\alpha \frac{\hat{\sigma}_{c_1}}{\sqrt{N}} \ \forall \, k = \overline{1,400}. \tag{67}$$

Here $\hat{\sigma}_{b_1}, \hat{\sigma}_{a_1}, \hat{\sigma}_{c_1}$ are estimates of the standard deviations $\sigma_{b_1}, \sigma_{a_1}, \sigma_{c_1}$, respectively; $\alpha = 0.05$ is the significance level; $t_\alpha = 2.26$ is the $100(1 - \alpha)\%$ point of Student's distribution with $L - 1$ degrees of freedom; $L = 10$ is the number of experiments. Table 2 illustrates the values $\bar{w}_1, \bar{w}_2$ of estimates $\hat{w}_1(k), \hat{w}_2(k)$, (averaged by 10 experiments), and their confidence intervals

$$\Delta_1 = \pm t_\alpha \frac{\hat{\sigma}_{w_1}}{\sqrt{N}}, \quad \Delta_2 = \pm t_\alpha \frac{\hat{\sigma}_{w_2}}{\sqrt{N}} \quad \forall \, k = \overline{1,400}. \tag{68}$$

Here $\hat{\sigma}_{w_1}, \hat{\sigma}_{w_2}$ are estimates of the standard deviations $\sigma_{w_1}, \sigma_{w_2}$, respectively; Note that in both tables the first line of each k corresponds to the averaged estimates and their confidence intervals which were calculated using the S-algorithm with Shweppe's GM-estimator while the second one – to the same values calculated by the H-algorithm with a version of Huber's M-estimator. The analysis of the estimates, presented in Tables 1, 2, implies that the results obtained by the S-algorithm with Shweppe's GM-estimator corroborate the fact that it is more appreciable than the H-algorithm with a version of Huber's M-estimator because of a higher accuracy of recursive estimates.

Table 1

The averaged estimates of parameters $a_1 = -0.985$, $b_1 = 2$, $c_1 = -0.7$ and their confidence intervals for different k

Observations	The averaged estimates of parameters		
k	$\bar{a}_1$	$\bar{b}_1$	$\bar{c}_1$
45	-0.997 ± 0.003	1.081 ± 0.015	-0.110 ± 0.008
	-0.997 ± 0.003	1.081 ± 0.015	-0.110 ± 0.008
100	-0.993 ± 0.006	1.371 ± 0.168	-0.396 ± 0.089
	-0.966 ± 0.029	2.453 ± 0.074	-0.169 ± 0.251
200	-0.988 ± 0.010	1.689 ± 0.213	-0.418 ± 0.054
	-0.975 ± 0.012	2.397 ± 0.098	-0.223 ± 0.241
300	-0.984 ± 0.007	1.925 ± 0.213	-0.400 ± 0.004
	-0.977 ± 0.010	2.384 ± 0.094	-0.234 ± 0.228
400	-0.984 ± 0.009	2.075 ± 0.196	-0.367 ± 0.003
	-0.977 ± 0.009	2.374 ± 0.092	-0.263 ± 0.216

Table 2

The averaged estimates of parameters $w_1 = 0.1005$, $w_2 = -0.1016$ and their confidence intervals for different k

Observations	The averaged estimates of parameters	
k	$\bar{w}_1$	$\bar{w}_2$
45	-0.154 ± 0.003	-0.119 ± 0.002
	-0.154 ± 0.003	-0.119 ± 0.002
100	0.115 ± 0.096	-0.097 ± 0.039
	-0.194 ± 0.237	-0.178 ± 0.094
200	0.110 ± 0.067	-0.116 ± 0.018
	-0.142 ± 0.239	-0.196 ± 0.061
300	0.072 ± 0.100	-0.128 ± 0.020
	-0.122 ± 0.224	-0.191 ± 0.058
400	0.028 ± 0.062	-0.144 ± 0.022
	-0.102 ± 0.217	-0.186 ± 0.057

7. Conclusions

Despite that the RML technique has been worked out for an additive correlated noise generated by the filter (31) from the statistically independent and stationary sequence with (3), it appears that the ordinary RML is also applicable in the presence of large, but rare outliers in output observations (see Figs. 5, 7). If the number of outliers is increasing,

then the robustified versions of the RMLs gain an advantage over the ordinary RML procedure (Figs. 9, 10, 11). It seems that the recursive estimates, obtained by the S-algorithm with Shweppe's GM-estimator, and the estimates of controller coefficients, recalculated in each current iteration, approach the true values of parameters and coefficients, respectively, more rapidly than that calculated by the recursive H-technique with a version of Huber's M-estimator (see Figs. 10, 11 and Tables 1, 2). On the other hand, the estimates of controller coefficients are not robust even for Shweppe's GM-estimator.

References

Åström, K.J. (1987). Adaptive feedback control. *Proc. of the IEEE*, **75**(2), 185–217.

Åström, K.J., and T. Bohlin (1965). Numerical identification of linear dynamic systems from normal operating records. In *IFAC Symp. on Theory of Selfadaptive Control Systems*, Teddington, N.Y. Plenum Press. pp. 96–106.

Åström, K.J., and B. Wittenmark (1987). *Computer Controlled Systems*. Mir, Moscow (in Russian).

Clarce, D.W. (1967). Generalized least-squares estimation of the parameters of a dynamic model. In *IFAC Symp. on Identification in Automatic Control Systems*, Paper 3.17. Prague.

Denby, L., and R.D. Martin (1979). Robust estimation of the first-order autoregressive parameter. *J. Amer. Statist. Assoc.*, **74**, 140–146.

Eykhoff, P. (1975). *System Identification. Parameter and State Estimation*. Mir, Moscow (in Russian).

Forsell, U., and L. Ljung (1999). Closed-loop identification revisited. *Automatica*, **35**, 1215–1241.

Genov, D.G., N.R. Atanasov and R. Pupeikis (2006). Robust M- and GM-estimators for closed-loop identification using the direct approach. In *Proceedings of Int. Conf. on Automatic and Informatics*, Sofia. pp. 193–196 (in Bulgarian).

Glad, T., and L. Ljung (2000). *Control Theory. Multivariable and Nonlinear Methods*. Taylor & Francis, N.Y.

Halwass, M. (1988). Selbsteinstellende LQG-regelung. *MSR*, Teil 1, 1, 2–6.

Halwass, M. (1988). Selbsteinstellende LQG-regelung. *MSR*, Teil 2, 2, 61–63.

Hastings–James, R., and M.W. Sage (1969). Recursive generalized least-squares procedure for on-line identification of process parameters. *Proc. IEE*, **116**, 2057–2062.

Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.*, **35**, 73–101.

Huber, P.J. (1984). *Robust Statistics*. Mir, Moscow (in Russian).

Isermann, R. (1974). *Prozessidentifikation*. Springer Verlag, Berlin.

Isermann, R. (1984). *Digital Control Systems*. Springer Verlag, Mir, Moscow (in Russian).

Isermann, R., U. Baur, W. Bamberger, P. Kneppo and H. Siebert (1974). Comparison of six on-line identification and parameter estimation methods. *Automatica*, **10**(1), 81–103.

Ljung, L. (1977). Analysis of recursive stochastic algorithms. *IEE Trans. AC.*, **22**(4), 551–575.

Lucas, A. (1996). *Outlier Robust Unit Root Analysis*. Thesis Publishers, Amsterdam.
    http://staff.feweb.vu.nl/alucas/thesis/default.htm

Proakis, J.G., and D.G. Manolakis (1996). *Digital Signal Processing*: *Principles, Algorithms and Applications*, Prentice-Hall, London.

Novovičova, J. (1987). Recursive computation of M-estimates for the parameters of the linear dynamical system. *Problems of Control and Information Theory*, **16**(1), 19–59.

Petersen, I.R. (2006). Minimax LQG control. *Int. J. Appl. Math. Comput. Sci.*, **16**(3), 309–323.

Polyak, B.T., and Ya.Z. Tsypkin (1980). Robust identification. *Automatica*, **16**(1), 53–65.

Pupeikis, R. (1991). Recursive robust estimation of dynamic systems parameters. *Informatica*, **2**(4), 579–592.

Pupeikis, R. (2000). Closed-loop robust identification using the direct approach. *Informatica*, **11**(2), 163–178.

Saridis, G. (1974). Stochastic approximation methods for identification and control – a survey. *IEEE Trans. AC.*, **19**(6), 798–809.

Steiglitz, K., and L.E. McBride (1965). A technique for the identification of the linear systems. *IEEE Trans. AC.*, **10**(4), 461–464.

Stoica, P., and T. Söderström (1981). The Steiglitz–McBride identification algorithm revisited-convergence analysis and accuracy aspects. *IEEE Trans. AC.*, **26**(3), 712–717.

Van den Boom, A.J.W. (1981). ????? Doctoral thesis. Technishe hogeshool Eindhofen.

N. Atanasov, PhD is an assistent professor at the Department of Automation, Technical University of Varna. His main research interests are in robust identification, robust estimation, control theory, adaptive control.

R. Pupeikis received PhD degree from the Kaunas Polytechnic Institute, Kaunas, Lithuania, 1979. He is a senior researcher at the Process Recognition Department of the Institute of Mathematics and Informatics and an associated professor at the Gediminas Technical University. His research interest include the digital signal processing, robust approaches of dynamic system identification as well technological process control.

Apie rekurentinį M- ir GM- įverčių skaičiavimą, taikant tiesioginį LQG valdymo sistemų identifikavimą

Nasko ATANASOV, Rimantas PUPEIKIS

Straipsnyje vystomas parametrinio LQG (tiesinis kvadratinis Gauso) valdymo sistemų identifikavimo metodas, kai tiesinės pastovių koeficientų sistemos bei LQG reguliatoriaus parametrai esti nežinomi ir turi būti skaičiuojami. Čia taikomos rekurentinės parametrų įvertinimo išraiškos, grindžiamos M- ir GM- įverčių skaičiavimo algoritmais, kadangi LQG valdymo sistemos išėjimas stebimas triukšme su didelėmis to triukšmo reikšmių išmetomis. Kiekvienoje skaičiavimų iteracijoje gauti sistemos parametrų įverčiai taikomi LQG reguliatoriaus koeficientų reikšmėms perskaičiuoti. Pateikti LQG valdymo sistemos modeliavimo bei jos parametrinio identifikavimo rezultatai.