# On the Optimization of an Adaptive State Estimation of Dynamic Systems in the Presence of Time-varying Outliers

Rimantas PUPEIKIS

*Institute of Mathematics and Informatics, Vilnius Gediminas Technical University*
*Akademijos 4, 2600 Vilnius, Lithuania*
*e-mail: pupeikis@ktl.mii.lt*

**Abstract.** The problem of recursive estimation of a state of dynamic systems in the presence of time-varying outliers in observations to be processed has been considered. A learning phase used in the state estimation is investigated, assuming that the observations of a noisy output signal and that of a training one are given. A technique based on robust filtering by means of a bank of parallel Kalman filters and on the procedure of optimization of the state estimation itself is used, choosing, at each time moment, a current estimate, that ensures a minimal absolute deviation from the current value of the teaching signal. An approach, based on the relation between the mean squared deviation of state estimates from the true state and innovation sequence variance as well as on the fact that both variables achieve their minimum for the same filter from the respective Kalman filter bank, is proposed here for a working phase, where a training signal will be absent. The recursive technique based on an adaptive state estimation with optimization procedure is worked out. The results of numerical simulation of the linear discrete-time invariant (LTI) system (56) by computer using a bank, consisting of Kalman filters are given (Figs. 1–5).

**Key words:** system, Kalman filter, robustness, state estimation, optimization.

## 1. Introduction

The filtering techniques, based on an ordinary Kalman filter, completely lose their optimal properties because of outliers in observations to be processed. Therefore, various robust algorithms are worked out in order to increase the efficiency of such approaches where the distribution of signals, used for the state estimation, deviates from the Gaussian distribution. In such a case, the solution of the problem could be obtained using a bank or a network of parallel Kalman filters (Pupeikis and Huber, 1997; Pupeikis, 1998; Pupeikis, 1999; Marcos, 2000), respectively, and an approach that optimizes the estimation itself. It can be mentioned here that banks of Kalman filters have been used in fault detection since 1968 (Newbold and Ho, 1968). On the other hand, during the state estimation by means of the bank of Kalman filters in the presence of outliers – as a state, we understand here the current output of an observed dynamic system – there always arises a problem to choose some variable, that could be used for the optimization of state estimation. It can

be solved by separating the state estimation itself into two phases: a learning phase and a working one, according to Mulgrew and Cowan (1987) and Marcos (2000), assuming, as usual, that some teaching signal is given in the learning phase. Then, for each Kalman filter at each current iteration, a current value of deviation of the reconstructed output from the teaching signal could be calculated. Afterwards, at each current iteration the state estimate that ensures the minimal value of such a deviation and, correspondingly, some Kalman filters, producing such an estimate, from a bank of filters could be determined. After terminating the learning phase, the most appropriate filter or some filters will be chosen to estimate states in the working phase, where there is no teaching signal. Therefore, there arises a problem to choose a new function to be minimized during the state estimation in working phase. This problem is solved here at first analytically and afterwards it is verified by the simulation on PC, using the bank of Kalman filters.

## 2. Statement of the Problem

Assume that a linear discrete-time invariant (LTI) system can be described using state equations of the form

$$
\begin{aligned}
x[n+1] &= Ax[n] + Bu[n], \\
y[n] &= Cx[n],
\end{aligned}
\tag{1}
$$

with the $m \times 1$ state vector $x[n]$, the $p \times 1$ input vector $u[n]$, the $r \times 1$ output vector $y[n]$, and the $m \times m$, $m \times p$, $r \times m$ known beforehand matrices $A, B, C$, respectively. Here

$$
\begin{aligned}
x[n+1] &= \big(x_1[n+1], x_2[n+1], \ldots, x_m[n+1]\big)^T, \\
x[n] &= \big(x_1[n], x_2[n], \ldots, x_m[n]\big)^T
\end{aligned}
\tag{2}
$$

are vectors of states at time moments $n+1$ and $n$, respectively;

$$
u[n] = \big(u_1[n], u_2[n], \ldots, u_m[n]\big)^T
\tag{3}
$$

are the values of an input $\{u[n]\}$;

$$
y[n] = \big(y_1[n], y_2[n], \ldots, y_m[n]\big)^T
\tag{4}
$$

is unobserved output of the object.

The observations of a noisy output $\{z[n]\}$ described by

$$
z[n] = y[n] + v[n],
\tag{5}
$$

and of a training signal

$$
r[n] = \big(r_1[n], r_2[n], \ldots, r_m[n]\big)^T,
\tag{6}
$$

are given. Here

$$z[n] = \big(z_1[n], z_2[n], \ldots, z_m[n]\big)^T \tag{7}$$

are observations of output (5), $\{v[n]\}$ is a sequence of independent identically distributed variables with an "$\varepsilon$-contaminated" distribution of the form

$$p(v_n) = (1 - \varepsilon_n)N(0, \sigma_\xi^2) + \varepsilon_n N(0, \sigma_\varsigma^2), \tag{8}$$

and the variance

$$\sigma_v^2 = (1 - \gamma_n)\sigma_\xi^2 + \gamma_n \sigma_\varsigma^2;$$

$p(z_n)$ is a probability density distribution of the sequence $v[n] = \big(v_1[n], v_2[n], \ldots, v_m[n]\big)^T$ with the value

$$v_i[n] = (1 - \gamma_n)\xi_i[n] + \gamma_n \varsigma_i[n] \tag{9}$$

of additive noise at the time moment $n$; $\gamma_n$ is a random variable, taking values 0 or 1 with probabilities $p(\gamma_n = 0) = 1 - \varepsilon_n$, $p(\gamma_n = 1) = \varepsilon_n$; $\{\xi[n]\}$, $\{\varsigma[n]\}$ are sequences of independent Gaussian variables with zero means and variances $\sigma_\xi^2, \sigma_\varsigma^2$, besides, $\sigma_\xi << \sigma_\varsigma$; $0 \leqslant \varepsilon_n \leqslant 1$ is an unknown fraction of contamination varying in time.

The aim of the given paper is the development of an approach for a robust recursive estimation of states $y_1[n] = y_n, y_2[n] = y_{n-1}, \ldots, y_m[n] = y_{n-m+1}$ of an LTI object (1), using the observations of the noisy output $\{z[n]\}$ as well as a teaching signal $\{r[n]\}$ in the learning phase and only the observations of $\{z[n]\}$ in the working phase.

## 3. Discrete Kalman Filters

The ordinary Kalman filter used for the state estimation could be written as follows:

$$\begin{aligned}
\hat{x}[n|n] &= \hat{x}[n|n-1] + M[n]\big(z[n] - C\hat{x}[n|n-1]\big), \\
M[n] &= P[n|n-1]C^T\big(R[n] + CP[n|n-1]C^T\big)^{-1}, \\
P[n|n] &= \big(I - M[n]C\big)P[n|n-1], \\
\hat{x}[n+1|n] &= A\hat{x}[n|n], \\
P[n+1|n] &= AP[n|n]A^T.
\end{aligned} \tag{10}$$

Here $\hat{x}[n|n-1]$, $\hat{x}[n|n]$ are estimates of the vector $x[n]$ determined by processing observations including $z[n-1]$ and $z[n]$, respectively;

$$\begin{aligned}
R[n] &= E\{v[n]v^T[n]\}, \\
P[n|n] &= E\Big\{\big(x[n] - x[n|n]\big)\big(x[n] - x[n|n|n-1]\big)^T\Big\}, \\
P[n|n-1] &= E\Big\{\big(x[n] - x[n|n-1]\big)\big(x[n] - x[n|n-1]\big)^T\Big\}.
\end{aligned} \tag{11}$$

It is known (Masreliez and Martin, 1977; Schick and Mitter, 1994) that the Kalman filter turns out to be inefficient while estimating the state of the dynamic system (1) because of outliers in observations. Therefore, in such a case, according to Masreliez and Martin, 1977 and to Meyr and Spies, 1984, the simple robust versions of (10) could be obtained by changing the first recursive equation of (10) as follows

$$\hat{x}[n|n] = \hat{x}[n|n-1] + M[n]\psi\big\{z[n] - C\hat{x}[n|n-1]\big\}, \tag{12}$$

where

$$\psi(e[n]) = \begin{cases} -\Delta & \text{if} \quad e[n] < -\Delta, \\ e[n] & \text{if} \quad -\Delta \leqslant e[n] \leqslant \Delta, \\ \Delta & \text{if} \quad e[n] > \Delta \end{cases} \tag{13}$$

is Huber's $\psi$ – function with

$$e[n] = z[n] - C\hat{x}[n|n-1], \tag{14}$$

as an innovation at the time moment $n + 1$ and the threshold $\Delta$, which depends on the standard deviation $\sigma_\xi$ of the ground distribution and on a fraction of contamination $\varepsilon_n$.

The choice of $\Delta$ plays an important role for the estimation of an unknown parameter or state. In this connection, the existing propositions may be used to determine the threshold $\Delta$ (Huber, 1981; Hampel *et al.,* 1986; Ljung, 1991; Verboon, 1994; Stockinger and Dutter, 1987).

## 4. Optimization of a State Estimation in the Learning Phase

The aim of the learning phase consists in choosing from a bank of parallel Kalman filters a filter or several ones, such that ensure a maximal accuracy of state estimation in respect of other filters. These Kalman filters could be used in the working phase while the other filters would be rejected. In the learning phase, the bank of state estimates

$$\hat{x}_1[1], \ldots, \hat{x}_L[1]; \hat{x}_1[2], \ldots, \hat{x}_L[2]; \hat{x}_1[N], \ldots, \hat{x}_L[N] \tag{15}$$

is calculated by processing noisy output $\{z[n]\}$ observations, containing time-varying additive outliers, using the bank of parallel $L$ Kalman filters

$$\begin{aligned} \hat{x}_i[n|n] &= \hat{x}_i[n|n-1] + M_i[n]\psi_i\big(z[n] - C\hat{x}_i[n|n-1]\big), \\ M_i[n] &= P_i[n|n-1]C^T\big(R[n] + CP_i[n|n-1]C^T\big)^{-1}, \\ P_i[n|n] &= \big(I - M_i[n]C\big)P_i[n|n-1], \\ \hat{x}_i[n+1|n] &= A\hat{x}_i[n|n], \\ P_i[n+1|n] &= AP_i[n|n]A^T \quad \text{for} \quad i = 1, 2, \ldots, L. \end{aligned} \tag{16}$$

Here $\hat{x}_i[n|n-1]$, $\hat{x}_i[n|n]$ are estimates of the vector $x[n]$ determined by each $i$-th Kalman filter used for processing observations, including $z[n-1]$ and $z[n]$, respectively;

$$
\begin{aligned}
&R[n] = E\big\{v[n]v^T[n]\big\}, \\
&P_i[n|n] = E\Big\{\big(x_i[n] - x_i[n|n]\big)\big(x_i[n] - x_i[n|n|n-1]\big)^T\Big\}, \\
&P_i[n|n-1] = E\Big\{\big(x_i[n] - x_i[n|n-1]\big)\big(x_i[n] - x_i[n|n-1]\big)^T\Big\} \\
&\qquad\qquad \text{for} \quad i = 1, 2, \ldots, L,
\end{aligned}
\tag{17}
$$

$$
\psi_i(e_i[n]) = \begin{cases}
-\Delta_i & \text{if} \quad e_i[n] < -\Delta_i, \\
e_i[n] & \text{if} \quad -\Delta_i \leqslant e_i[n] \leqslant \Delta_i \\
\Delta_i & \text{if} \quad e_i[n] > \Delta_i \quad \text{for } i = 1, 2, \ldots, L,
\end{cases}
\tag{18}
$$

$$
e_i[n] = z[n] - C\hat{x}_i[n|n-1] \quad \text{for} \quad i = 1, 2, \ldots, L.
\tag{19}
$$

Thus, current state estimates (15) of LTI system (1) are calculated, using the bank of parallel Kalman filters of the form (16) with some initial values (17) and Huber's $\psi$-function (18) with an innovation (19). At each current time moment $n = 1, 2, \ldots, N$ the state estimate $\hat{x}_i^*[n]$, that guarantees the minimal absolute deviation $\delta^*[n]$ from the value of the training signal $\{r[n]\}$ is determined as follows: first, at the current time moment $n = 1, 2, \ldots, N$ the bank of deviations $\delta_1[n], \delta_2[n], \ldots, \delta_L[n]$ with $\delta_1[n] = \big|r[n] - \hat{x}_1[n]\big|$, $\delta_2[n] = \big|r[n] - \hat{x}_2[n]\big|, \ldots, \delta_L[n] = \big|r[n] - \hat{x}_L[n]\big|$ is calculated, second, $\delta^*[n]$, which is least of $\delta_1[n], \delta_2[n], \ldots, \delta_L[n]$ is determined. Thus, while estimating states in such a way at each current time moment $n = 1, 2, \ldots, N$, the state estimates $\delta^*[1], \delta^*[2], \ldots, \delta^*[N]$, ensuring the minimal deviation from the reference sequence $\{r[n]\}$ are found. It can be mentioned that the accuracy of state estimates depends on the number of filters in the Kalman filter bank as well as on thresholds in $\psi_i$-functions to be used. When the learning phase is over, some most appropriate Kalman filters, ensuring the maximal accuracy of prediction of a state, are established.

## 5. Determination of an Optimization Criterion for the State Estimation

In the working phase, there is no teaching signal used to solve the state estimation problem. Then the above-mentioned approach will fail and one has to choose a new variable for the optimization of state estimation. We will try to determine such a variable below.

**Theorem 1.** *If for large enough $N$ in the bank of Kalman filters* (16)–(19) *there exists such an $l$-th ($l \in L$) filter, ensuring that the variance $\sigma_{e_l}^2$ of the innovation sequence $\{e_l[n]\}$ of the form*

$$
\{e_l[n]\} = \psi_l\big\{z[n] - C\hat{x}_l[n]\big\} \quad \forall n
\tag{20}
$$

*is minimal as compared with the estimates of variance, calculated for the other $L - 1$ filters from the same bank, then for such an $l$-th filter, the mean squared deviation*

$$
E_l\big(x[n], \hat{x}[n]\big) = N^{-1}w_l^T[N]w_l[N]
\tag{21}
$$

*acquires a minimum. Here* $w_l[N] = \big(x[1] - \hat{x}_l[1], x[2] - \hat{x}_l[2], \ldots, x[N] - \hat{x}[N]\big)^T$ *is a vector of deviations of the values of the state estimates from the true state values, determined for the l-th filter.*

*Proof.* Assume that for the $l$-th filter

$$\hat{x}_l[1] = x[1] + \zeta_l[1], \hat{x}_l[2] = x[2] + \zeta_l[2], \ldots, \hat{x}_l[N] = x[N] + \zeta_l[N], \tag{22}$$

where $\{\zeta_l[n]\} \sim N(0, \sigma^2_{\zeta_l})$ and such that $\sigma^2_{\zeta_l}$ is lower than any $\sigma^2_{\zeta_i}$, which is determined for any $i$-th ($i \in L$) filter with

$$\hat{x}_i[1] = x[1] + \zeta_i[1], \hat{x}_i[2] = x[2] + \zeta_i[2], \ldots, \hat{x}_i[N] = x[N] + \zeta_i[N]. \tag{23}$$

Then the variance $\sigma^2_{e_l} = \sigma^2_v + \sigma^2_{\zeta_l}$ of the innovation sequence

$$\{e_l[n]\} = \psi_l\big\{z[n] - C\hat{x}_l[n]\big\} \quad \forall n \tag{24}$$

is lower than any estimate of variance, calculated for the other filters from the same bank. It also follows, that the mean square deviation (21) acquires on a minimum.

**Theorem 2.** *If in the bank of Kalman filters (16)–(19) there exists a j-th (j $\in$ L) filter, ensuring that the variance* $\sigma^2_{e_j}$ *of the innovation sequence* $\{e_j[n]\}$ *of the form*

$$\{e_j[n]\} = \psi_j\{z[n] - C\hat{x}_j[n]\} \quad \forall n \tag{25}$$

*is maximal as compared with the estimates of variance, calculated for the other filters from the same bank, then for such a j-th filter the mean square deviation*

$$E_j\big(x[n], \hat{x}[n]\big) = N^{-1}w_j^T[N]w_j[N] \tag{26}$$

*acquires a maximum. Here* $w_j[N] = \big(x_j[1] - \hat{x}_j[1], x_j[2] - \hat{x}_j[2], \ldots, x_j[N] - \hat{x}_j[N]\big)^T$ *is a vector of deviations of the values of the state estimates from the true state values, determined for the j-th filter.*

*Proof.* Assume that for the $j$-th filter

$$\hat{x}_j[1] = x[1] + \zeta_j[1], \quad \hat{x}_j[2] = x[2] + \zeta_j[2], \ldots, \quad \hat{x}_j[N] = x[N] + \zeta_j[N], \tag{27}$$

where $\{\zeta_j[n]\} \sim N(0, \sigma^2_{\zeta_j})$ and such that $\sigma^2_{\zeta_j}$ is more than any $\sigma^2_{\zeta_i}$, which is determined for any $i$-th ($i \in L$) filter with

$$\hat{x}_i[1] = x[1] + \zeta_i[1], \quad \hat{x}_i[2] = x[2] + \zeta_i[2], \ldots, \quad \hat{x}_i[N] = x[N] + \zeta_i[N]. \tag{28}$$

Then the variance $\sigma^2_{e_j} = \sigma^2_v + \sigma^2_{\zeta_j}$ of the innovation sequence

$$\{e_j[n]\} = \psi_j\big\{z[n] - C\hat{x}_j[n]\big\} \quad \forall n \tag{29}$$

is higher than any estimate of the variance, calculated for the other filters from the same bank. It also follows, that the mean squared deviation (26) acquires a maximum.

**Theorem 3.** *If in the bank of Kalman filters* (16)–(19) *there exists an $f$-th ($f \in L$) filter with*

$$\hat{x}_f[1] = x[1] + \zeta_f[1], \quad \hat{x}_f[2] = x[2] + \zeta_f[2], \dots, \quad \hat{x}_f[N] = x[N] + \zeta_f[N], \quad (30)$$

*where $\{\zeta_f[n]\} \sim N(0, \sigma^2_{\zeta_j})$ and such that $\sigma^2_{\zeta_f}$ is less than any $\sigma^2_{\zeta_i}$, which is determined for any $i$-th ($i \in L$) filter with*

$$\hat{x}_i[1] = x[1] + \zeta_i[1], \quad \hat{x}_i[2] = x[2] + \zeta_i[2], \dots, \quad \hat{x}_i[N] = x[N] + \zeta_i[N], \quad (31)$$

*then the variance $\sigma^2_{e_f}$ of the innovation sequence*

$$\{e_j[n]\} = \psi_f\{z[n] - C\hat{x}_f[n]\} \quad \forall n \tag{32}$$

*and the mean squared deviation*

$$E_f(x[n], \hat{x}_f[n]) = N^{-1} w_f^T[N] w_f[N] \tag{33}$$

*both are minimal as compared with the estimates of variance and mean square deviations, respectively, calculated for the other filters from the same bank.*

Here

$$w_f[N] = \big(x_f[1] - \hat{x}_f[1], x_f[2] - \hat{x}_f[2], \dots, x_f[N] - \hat{x}_f[N]\big)^T \tag{34}$$

*is a vector of deviations of the values of the state estimates from the true state values, determined for the $f$-th filter.*

Theorem 3 follows from Theorems 1 and 2.

**Theorem 4.** *If in the bank of Kalman filters* (16)–(19) *there exists an $l$-th ($l \in L$) filter with*

$$\hat{x}_l[1] = x[1] + \zeta_l[1], \quad \hat{x}_l[2] = x[2] + \zeta_l[2], \dots, \quad \hat{x}_l[N] = x[N] + \zeta_l[N], \quad (35)$$

*where $\{\zeta_l[n]\} \sim N(0, \sigma^2_{\zeta_l})$ and such that $\sigma^2_{\zeta_l}$ is lower than any $\sigma^2_{\zeta_i}$, which is determined for any $i$-th ($i \in L$) filter with*

$$\hat{x}_i[1] = x[1] + \zeta_i[1], \quad \hat{x}_i[2] = x[2] + \zeta_i[2], \dots, \quad \hat{x}_i[N] = x[N] + \zeta_i[N]. \quad (36)$$

*Then for an $l$-th filter the functions*

$$E_i(x[n], \hat{x}_i[n]) = N^{-1} w_i^T[N] w_i[N] \quad for \quad i = 1, 2, \dots, L, \tag{37}$$
$$Q_i(e[n], e_i[n]) = (N-1)^{-1} \alpha_i^T \alpha_i \quad for \quad i = 1, 2, \dots, L \tag{38}$$

*acquire minimal values.*

*Here*

$$\alpha_i = \big(e[1] - e_i[1], \ldots, e[N] - e_i[N]\big)^T \quad for \quad i = 1, 2, \ldots, L, \tag{39}$$

$e[k] \neq e_i[k] \ \forall k = 1, 2, \ldots, N$ *and* $\forall i = 1, 2, \ldots, L$; $e[1], e[2], \ldots, e[N]$ *are values of the sequence* $\{e[k]\}$ *of the form*

$$\{e[n]\} = \{z[n] - Cx[n]\} \quad \forall n = 1, 2, \ldots, N, \tag{40}$$

*determined in the absence of outliers in observations, substituting here true state values* $x[1], x[2], \ldots, x[N]$,

$$w_i[N] = \big(x_i[1] - \hat{x}_i[1], x_i[2] - \hat{x}_i[2], \ldots, x_i[N] - \hat{x}_i[N]\big)^T. \tag{41}$$

*Proof.* Assume that for the $l$-th filter

$$\hat{x}_l[1] = x[1] + \zeta_l[1], \quad \hat{x}_l[2] = x[2] + \zeta_l[2], \ldots, \quad \hat{x}_l[N] = x[N] + \zeta_l[N], \tag{42}$$

where $\{\zeta_l[n]\} \sim N(0, \sigma_{\zeta_l}^2)$. Moreover, $\sigma_{\zeta_l}^2$ is less than any $\sigma_{\zeta_i}^2$, which is determined for any $i$-th ($i \in L$) filter with

$$\hat{x}_i[1] = x[1] + \zeta_i[1], \quad \hat{x}_i[2] = x[2] + \zeta_i[2], \ldots, \quad \hat{x}_i[N] = x[N] + \zeta_i[N]. \tag{43}$$

Then the mean squared deviation

$$E_l(x[n], \hat{x}[n]) = N^{-1} w_l^T[N] w_l[N], \tag{44}$$

according to the Theorems 1, 3 is minimal. It follows that function (37) acquires a minimum. On the other hand, for the right-hand side of function (38) it can be written

$$\lim_{N \to \infty} N^{-1} \alpha_i^T \alpha_i = \lim_{N \to \infty} \sum_{k=1}^{N} \big(e^2[k] - 2e[k]e_i[k] + e_i^2[k]\big) = \sigma_e^2 + \sigma_{e_i}^2, \tag{45}$$
$$\forall i = 1, 2, \ldots, L,$$

while

$$\lim_{N \to \infty} N^{-1} \sum_{k=1}^{N} e[k]e_i[k] = \mathrm{cov}\big(e[k]e_i[k]\big) = 0, \quad \forall i = 1, 2, \ldots, L, \tag{46}$$

as $\{e[k]\} \sim N(0, \sigma_e^2)$ and $\{e_i[k]\} \sim N(0, \sigma_{e_i}^2) \ \forall i = 1, 2, \ldots, L$ are mutually uncorrelated for all $k$. Here $\mathrm{cov}\big(e[k]e_i[k]\big)$ is the covariance between $\{e[k]\}$ and $\{e_i[k]\}$

$\forall i = 1, 2, \ldots, L$ and $\forall k = 1, 2, \ldots, N$. It follows from (45) that function (38) achieves its minimum for the $i$-th filter, that is

$$Q_i\big(e[k], e_i[k]\big) < Q_j\big(e[k], e_j[k]\big) \quad \text{for} i \neq j, \tag{47}$$

if and only if

$$\sigma_{e_i}^2 < \sigma_{e_j}^2 \quad \text{for} \quad j = 1, 2, \ldots, L - 1. \tag{48}$$

REMARK 1. Functions (37) and (38) acquire their minimum for the same $i$-th filter.

REMARK 2. Functions (37) and (38) can not be calculated because of unknown $x[n]$ and $e[n]$ $\forall n$, respectively.

REMARK 3. Function (38) could be replaced by the calculable function

$$Q_i(e_i[n]) = (N - 1)^{-1} \beta_i^T \beta_i \quad (i \in L), \tag{49}$$

because for the each filter from the same bank component $\sigma_e^2$ is the same. Here

$$\beta_i = \big(e_i[1], \ldots, e_i[N]\big)^T \quad \text{for} \quad i = 1, 2, \ldots, L \tag{50}$$

vector of respective innovation sequence.

**Conclusion**. The existing relation between the mean squared deviation function (37) and the function (38) of variances of innovation sequences, calculated for each filter from the bank of Kalman filters (16)–(19), allowed us to replace the function (38) by the function (49), which components can be easy determined. Thus, the vector $\hat{x}_l[n] = (\hat{x}_l[1], \ldots, \hat{x}_l[N])^T$ of estimates of states $x[1], \ldots, x[N]$ may be defined using the criterion (49) and the condition

$$\hat{x}_l[n] : Q(e_l[n]) = \min_{e_i[n] \in D} Q_i(e_i[n]), \quad (i \in L). \tag{51}$$

Here $D$ is a restricted area of values of innovation sequences $e_i[n]$ $\forall i = 1, 2, \ldots, L$ generated by the bank of the Kalman filters.

## 6. Optimization of a State Estimation in a Working Phase

Thus, the current state estimates of a LTI system (1) are calculated by the bank of the $L$ parallel Kalman filters (16)–(19), that yield a bank of state estimates

$$\hat{x}_1[1], \ldots, \hat{x}_L[1]; \quad \hat{x}_1[2], \ldots, \hat{x}_L[2]; \quad \hat{x}_1[N], \ldots, \hat{x}_L[N], \tag{52}$$

and a respective bank of innovation sequences

$$e_1[1], \ldots, e_L[1]; \quad e_1[2], \ldots, e_L[2]; \quad e_1[N], \ldots, e_L[N], \tag{53}$$

using respective Huber's $\psi$-functions of the form (18). Then for $i = 1, 2, \ldots, L$ the estimates of the variance $\sigma_e^2$ are calculated by the formula

$$
\begin{bmatrix} \sigma_{e_1}^2 \\ \sigma_{e_2}^2 \\ \vdots \\ \sigma_{e_{L-1}}^2 \\ \sigma_{e_L}^2 \end{bmatrix} = \frac{1}{N-1} \begin{bmatrix} \sum_{n=1}^{N} \left(e_1[n] - \bar{e}_1\right)^2 \\ \sum_{n=1}^{N} \left(e_2[n] - \bar{e}_2\right)^2 \\ \vdots \\ \sum_{n=1}^{N} \left(e_{L-1}[n] - \bar{e}_{L-1}\right)^2 \\ \sum_{n=1}^{N} \left(e_L[n] - \bar{e}_L\right)^2 \end{bmatrix}, \tag{54}
$$

using, respectively, $L$ innovation sequences (53). Here $\bar{e}_1, \ldots, \bar{e}_L$ are means of innovation sequences.

If the variance $\sigma_{e_l}^2$ satisfies the condition (51), then the vector $\hat{x}_l[n] = (\hat{x}_l[1], \ldots, \hat{x}_l[N])^T$ of estimates of the state $x[n] = (x[1], \ldots, x[N])^T$ is an optimal one. It can be mentioned that formula (52) allows us to choose the optimal state estimates $\hat{x}_l[n] = (\hat{x}_l[1], \ldots, \hat{x}_l[N])^T$ only after processing the whole set of observations $z[1], z[2], \ldots, z[N]$. Really, it is important to choose the current optimal state estimate at each time moment $k$. Thus, it is necessary to rewrite (54) in such a way

$$
\begin{bmatrix} \sigma_{e_1}^2 \\ \sigma_{e_2}^2 \\ \vdots \\ \sigma_{e_{L-1}}^2 \\ \sigma_{e_L}^2 \end{bmatrix}_{k+1} = \begin{bmatrix} \sigma_{e_1}^2 \\ \sigma_{e_2}^2 \\ \vdots \\ \sigma_{e_{L-1}}^2 \\ \sigma_{e_L}^2 \end{bmatrix}_{k} + \begin{bmatrix} \vartheta_{e_1} \\ \vartheta_{e_2} \\ \vdots \\ \vartheta_{e_{L-1}} \\ \vartheta_{e_L} \end{bmatrix}_{k+1}
$$
$$
= \left(1 - \frac{1}{k}\right) \begin{bmatrix} \sigma_{e_1}^2 \\ \sigma_{e_2}^2 \\ \vdots \\ \sigma_{e_{L-1}}^2 \\ \sigma_{e_L}^2 \end{bmatrix}_{k} + \frac{1}{k} \begin{bmatrix} e_1^2[k+1] \\ e_2^2[k+1] \\ \vdots \\ e_{L-1}^2[k+1] \\ e_L^2[k+1] \end{bmatrix}. \tag{55}
$$

Now, at each time moment $k$, the minimal variance $\sigma_{e_l}^2$ as well as the state estimate from the bank of Kalman filters (16)–(19) can be chosen. It can be mentioned that in case such as in (Pupeikis, 1998) there could also appear here a false optimum if too small thresholds $\Delta_i \forall i = 1, 2, \ldots, L$ were used in (18).

## 7. Simulation Results

The LTI system is defined by the state-space equation system (1) with

$$A = \begin{bmatrix} 1.1269 & -0.4940 & 0.1129 \\ 1.0000 & 0 & 0 \\ 0 & 1.0000 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -0.3832 \\ 0.5919 \\ 0.5191 \end{bmatrix}, \quad C = [1\ 0\ 0]. \tag{56}$$

The impulse and step responses of system (1) with matrices (56), as well as Bode diagrams are shown in Fig. 1. The signals of the LTI system are presented in Fig. 2. As an input to the LTI system the autoregressive-moving average signal (ARMA)

$$u[n] = \frac{1 - 0.8q^{-1}}{1 - 1.5q^{-1} + 0.7q^{-2}} \zeta[n], \tag{57}$$

is applied, using the sequence $\{\zeta[n]\} \sim N(0,1)$ as an initial sequence. A noisy input $\{u[n] + 0.5\varsigma[n]\}$ with $\{\varsigma[n]\} \sim N(0,1)$ is shown in Fig. 2a. The output signal $\{z[n]\}$ (Fig. 2b) is calculated, substituting the noisy input into the first equation of (1). Afterwards, $\{z[n]\}$ is corrupted by additive noise $\{v[n]\}$ with outliers according to expression (5). The teaching signal $\{r[n]\}$ (Fig. 2c) is determined using equations (1), where instead of the noisy input the unnoisy one, i.e., $\{u[n]\}$ is substituted into the first equation. The unobserved output signal without an additive noise is presented in Fig. 2d. The results of the state estimation, performed in the learning phase by processing observations of the output signal with outliers (Fig. 2b) and of the teaching one $\{r[n]\}$ (Fig. 2c), are shown
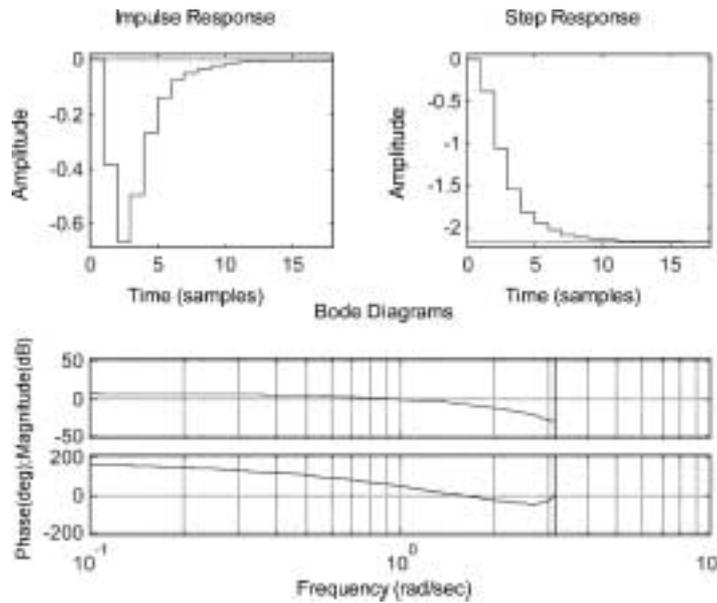


Fig. 1. Basic characteristics of the LTI system (1), (56) in time and frequency domains.
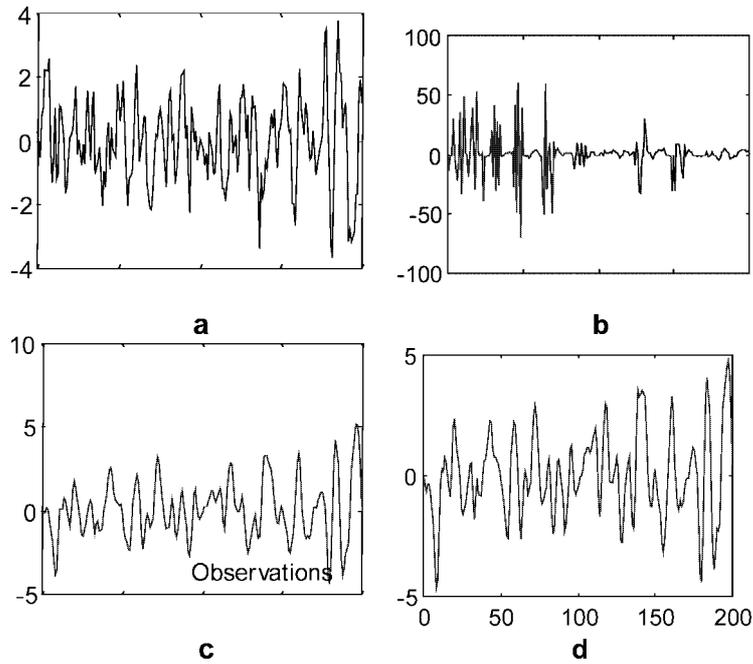
Fig. 2. Signals of the LTI system: noisy input (a), output with noise containing outliers (b), teaching signal (c), unobserved unnoisy output (d).

in Figs. 3, 4. The output, presented in Fig. 2b, is processed by one ordinary Kalman filter. The state estimates calculated by it dependent on the number of observations processed are shown in Fig. 3a. In Fig. 3b, we present the state estimates, calculated using the bank of 24 Kalman filters (16)–(19) with different thresholds $\Delta_i$ for $i = 1, 2, \ldots, 24$ in (18), which are chosen from the interval [0.25 5.8]. In Fig. 3c, the operation of the Kalman filters in time is shown. It can be mentioned that during the state estimation the numbers of the Kalman filters are choosen by the optimization technique, which is worked out here. From the simulation and state estimation results, presented in Fig. 3, it follows that the accuracy of state estimates, obtained using the bank of the Kalman filters by optimizing the estimation itself (Fig. 3b), is higher as compared to the accuracy of the estimates, obtained by the ordinary Kalman filter (Fig. 3a).

After the learning phase is terminated a single Kalman filter with $\Delta = 5.5$ in (18) from the bank of filters (16), (17) is chosen to estimate the states in a working phase. State estimation in a working phase is performed, using 300 observations of noisy output. In such a case, the teaching signal is absent. In Fig. 4a, the unobserved output and the one, corrupted by an additive noise with outliers are presented. In Fig. 4b, the unobserved output and the state estimates produced by one robust Kalman filter with $\Delta = 5.5$, selected in the learning phase, are shown.

Fig. 5 corresponds to the case, where the teaching signal is not available still in the learning phase. Therefore state estimates are calculated by processing the observations of noisy output with outliers (Fig. 2b) and a procedure, based on equations (51)–(55).
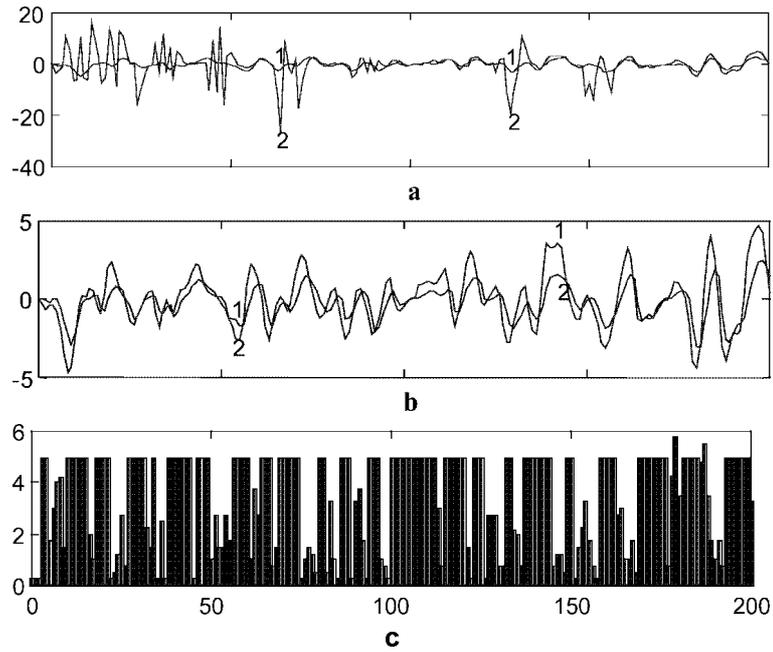
Fig. 3. State estimates and the operation of the Kalman filters. Estimates: obtained by an ordinary Kalman filter (a); obtained by the bank of Kalman filters (b), operation of the Kalman filters by processing observations (c). Curves: 1 – unobserved output, 2 – state estimates.
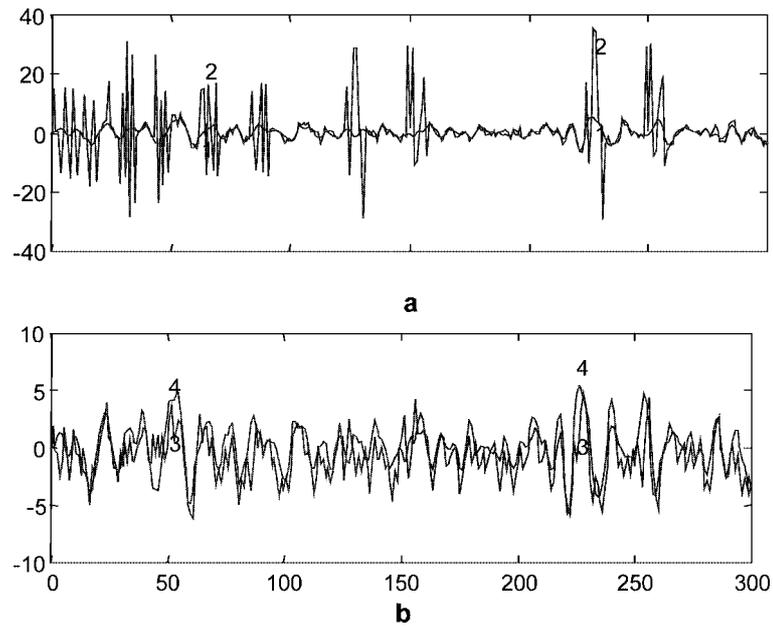


Fig. 4. Signals and state estimates in the working phase. Curves: 1, 3 – unobserved output, 2 – output to be processed, 4 – state estimates.
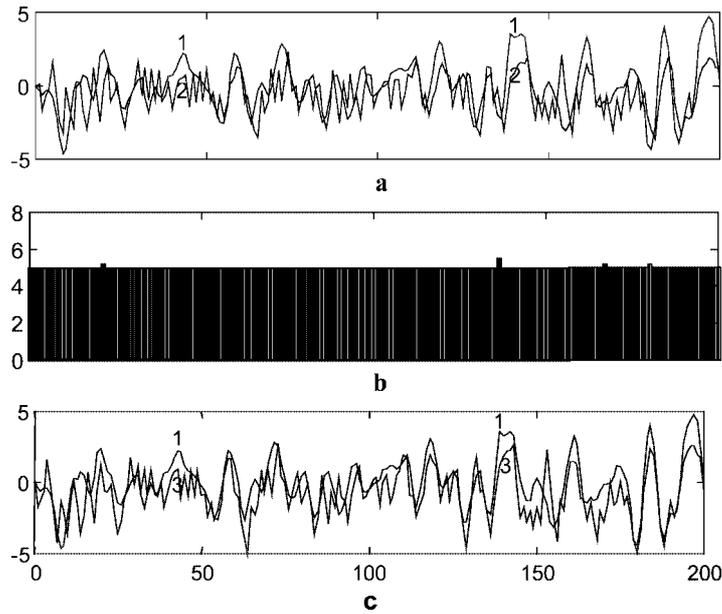
Fig. 5. State estimates and operation of the Kalman filters by processing observations. Estimates: obtained by
the bank of filters (a); obtained by one filter (c), operation of filters (b). Curves: 1 – unobserved output, 2, 3 –
state estimates.

In Fig. 5a, the state estimates calculated by the bank of 5 Kalman filters (16)–(19) with
different thresholds $\Delta_i$ for $i = 1, 2, \ldots, 24$ in (18), which are chosen from the interval
[5 5.8], are presented. In Fig. 5b, the operation of the Kalman filters in time is shown.
In Fig. 5c, the dependence of state estimates, found by one Kalman filter with $\Delta = 5$ in
(18), depending on the processed observations, is presented.

## 8. Conclusions

It follows that in the presence of time-varying outliers in observations to be processed
the accuracy of state estimates, obtained using the small bank of the Kalman filters (16)–
(19) by optimizing the estimation itself (Fig. 5a, b), is a little higher as compared to
the accuracy of estimates, obtained by one Kalman filter with $\Delta = 5$ in (18) (Fig. 5c)
or calculated employing large bank (Fig. 3b, c) using even the teaching signal. On the
other hand, as shown by the state estimation results presented in Figs. 4b and 5c, the
most important thing is choosing a $\Delta$ in (18). If $\Delta$ is chosen correctly (Fig. 4b), then the
accuracy of state estimates is higher as compared with that obtained by banks of filters
(Figs. 3b, 5a) even if the teaching signal is given (Fig. 3b). In spite of this, if the bank of
the Kalman filters is available, then a threshold $\Delta$ in (18) is chosen automatically. In the
opposite case, such a possibility is not valid, therefore one has to choose a threshold $\Delta$
arbitrary and often it is chosen incorrectly.

## References

Huber, P.J. (1981). *Statistics Robust*. Wiley, New York.

Hampel, F.R., E.M. Ronchetti, P., J. Rousseeuw, W.A. Stahel (1986). *Robust Statistics. The Approach Based on Influence Functions*. Wiley, New York.

Ljung, L. (1987). *System identification. Theory for the user*. Prentice-Hall, Inc.

Masreliez, C.J., R.D. Martin (1977). Robust bayesian estimation for the linear model and robustifying the Kalman filter. *IEEE Trans. on Automat. Contr.*, **22**(3), 361–371.

Meyr, H., G. Spies (1984). The structure and performance of estimators for real-time estimation of randomly varying time delay. *IEEE Trans. Acoust., Speech and Signal Process.*, **32**, 81–94.

Marcos, S.(2000). A network of adaptive Kalman filters for data channel equalization. *IEEE Trans. on Signal Process.*, **48**(9), 2620–2627.

Mulgrew, B., C.F.N. Cowan ( 1987). An adaptive Kalman equalizer: structure and performance. *IEEE Trans. on Acoust., Speech, and Signal Process.*, **35**(12), 1727–1735.

Newbold, P.M., and Y.C. Ho (1968). Detection of changes in the characteristics of a Gauss-Markov process. *IEEE Trans. on Aerospace and Electr. Syst.*, **4**(5).

Pupeikis, R., P.J. Huber (1997). Adaptiv-robuste Zustandsschätzung mit Optimierung des Schätzprozesses. *Bericht zur DFG Projekt "Echtzeit-Optimierunggrösser Systems"*.

Pupeikis, R. (1998). State estimation of dynamic systems in the presence of time-varying outliers in observations. *Informatica*, **9**(3), 325–342.

Pupeikis, R. (1999). Recursive optimization of state estimation of dynamic processes in the presence of patchy outliers in observations. *Informatica*, **10**(3), 297–312.

Schick, I.C, and S.K. Mitter (1994). Robust recursive estimation in the presence of heavy-tailed observation noise. *Ann. Math. Statist.*, **22**(2), 1045–1080.

Stockinger, N., and R. Dutter (1987). Robust time-series analysis. *Kybernetika*. **23**(1–5).

Verboon, P. (1994). *A Robust Approach to Nonlinear Multivariate Analysis*. Leiden University.

**R. Pupeikis** received Ph.D. 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 classical and robust approaches of dynamic system identification as well technological process control.

# Apie dinaminių sistemų adaptyviojo būsenos įvertinimo optimizavimą, taikant stebėjimus su nestacionaraus srauto išmetomis

Rimantas PUPEIKIS

Straipsnyje nagrinėjamas dinaminių sistemų (1)–(4) būsenos rekurentinis įvertinimas, apdorojant stebėjimus (5) su išmetomis (8), (9), kurių srauto intensyvumas kinta laikui bėgant. Būsenos įvertinimui taikomi Kalmano tipo filtrų bankai (16)–(19), generuojantys patvarius išmetų atžvilgiu įverčius. Teoriškai (teoremos 1–4) parodyta, kad esti priklausomybė tarp vidutinio kvadratinio būsenos ir jos įverčio nuokrypio funkcijos (37) bei atnaujinančių sekų dispersijų, gautų kiekvienam Kalmano filtrui iš minėto filtrų banko, funkcijos (38). Tai leidžia mums funkciją (38) pakeisti funkcija (49), kurios komponentės gali būti randamos, taikant kriterijų (49) ir optimizuojant būsenos įvertinimą pagal (51) išraišką. Pasiūlytas naujas metodas sistemos būsenai įvertinti, grindžiamas patvariąja filtracija ir optimizavimo procedūra (54), (55), kurią taikant kiekvienu laiko momentu iš būsenos įverčių banko (52) išrenkamas įvertis, optimalus minimalaus nuokrypio prasme. Dinaminės sistemos ((56), 1 pav.) su įėjimu (57) modeliavimo rezultatai (2–4 pav.) patvirtina teorinių išvadų pagrįstumą.