# A Simple Squared Innovations Based Decision Generator for Detection/Selection Problems in Linear Stochastic Control Systems

**I. V. Semoushin***
Department of Mathematical Cybernetics and Informatics
Ulyanovsk State University, Ulyanovsk, Russia

**Abstract:** We consider the linear time-invariant state-space stochastic control system

$$
\begin{aligned}
x(t_{i+1}) &= \Phi_\theta x(t_i) + \Psi_\theta u(t_i) + w(t_i), \quad x \in \mathbb{R}^n && \text{(i)}\\
y(t_i) &= H_\theta x(t_i) + v(t_i), \quad y \in \mathbb{R}^m && \text{(ii)}\\
\hat{x}_0(t_{i+1}^-) &= \Phi_0 \hat{x}_0(t_i^+) + \Psi_0 u(t_i), \quad \hat{x}_0 \in \mathbb{R}^n && \text{(iii)}\\
\hat{x}_0(t_i^+) &= \hat{x}_0(t_i^-) + K_0 \nu(t_i), \quad \nu(t_i) = y(t_i) - H_0 \hat{x}_0(t_i^-) && \text{(iv)}\\
u(t_i) &= f_{\mathrm{R}}[\hat{x}_0(t_i^+)] = -G_0^\star \hat{x}_0(t_i^+), \quad u \in \mathbb{R}^q && \text{(v)}
\end{aligned}
$$

where $i \in \mathbb{Z}$, with a plant (i) and a sensor (ii), both parameterized by an uncertainty vector parameter $\theta$, and a feedback controller FC, (iii)–(v), designed to cascade the steady-state Kalman filter (iii)–(iv) with a regulator (v) for a nominal value $\theta_0$ of $\theta$. In (i)–(v), the initial state $x(t_{-s})$ is given at some $t_{-s} \in \mathbb{R}$ with expectation $\mathbf{E}\left\{\|x(t_{-s})\|^2\right\} < \infty$; $w(t_i)$ and $v(t_i)$ are zero-mean mutually orthogonal wide-sense stationary orthogonal sequences with $\mathbf{E}\left\{w(t_i)w(t_i)^T\right\} = Q_\theta \geq 0$ and $\mathbf{E}\left\{v(t_i)v(t_i)^T\right\} = R_\theta > 0$ for all $t_i \in \mathbb{R}$, with $\left[\begin{smallmatrix} w(t_i) \\ v(t_i) \end{smallmatrix}\right]$ orthogonal to $x(t_j)$ and $u(t_j)$ for all $j \geq i$; $u(t_i)$ is wide-sense stationary and $\mathbf{E}\left\{\|u(t_i)\|^2\right\} < \infty$ for all $t_i \in \mathbb{R}$.

Assuming that the true value $\theta^\dagger$ of $\theta$ can change abruptly from $\theta_0$ to some $\theta_1$ at an instant $t_{\mathrm{c}} \in (t_0, t_i)$, we detect $t_{\mathrm{c}}$ using the nominal covariance $C_0$ of $\nu(t_i)$ in (iv) and constructing a decision function of the cumulative sum form:

$$
S_k = \sqrt{m/(2k)} \sum_{i=1}^{k} \left[\nu^T(t_i) C_0^{-1} \nu(t_i)/m - 1\right] .
$$

**Keywords:** Empirical decision procedures; Hypothesis testing; Diagnostics; Stochastic systems and control; Stopping times.

**Subject Classification:** 62C12; 62F03; 62J20; 93E10; 60G40.

## 1. INTRODUCTION

The theory of abrupt changes is an intensively investigated topic in time series analysis and identification (Basseville and Nikiforov, 1993). Usually by abrupt changes, it is meant changes in characteristics that occur very fast with respect to the sampling period of measurements, if not instantaneously, at unknown time instants. Besides strong theoretical motivations, the detection of such changes is a problem of great practical interest.

The subject of abrupt changes basically grew up at the confluence of several disciplines, first of all, mathematical statistics and automatic control theory. In mathematical statistics, decision tools for detecting changes appeared first in the area of quality control, where the Shewhart 3-sigma control charts were introduced (Shewhart, 1931). As an alternative to them, Page set up a principle of cumulative sum control

*Address correspondence to I. V. Semoushin, Department of Mathematical Cybernetics and Informatics, Ulyanovsk State University, 42 L. Tolstoy Street, Ulyanovsk 432970, Russia; Fax: +7-842-241-2340; E-mail: SemushinIV@ulsu.ru .

charts (Page, 1954). In succeeding years the CUSUM has been proven to be a very effective tool to detect changes in a mean level of a process. The expanded CUSUM mathematical theory and its industrial applications were given in (Johnson and Leone, 1962), (Johnson and Leone, 1963), (Johnson and Leone, 1964) and then in (Siegmund, 1985), (Kenett and Zacks, 1998) and (Ghosh and Sen, 1991). Wide historical notes and references for seminars, survey papers, and books related to change detection both in mathematical statistics and automatic control theory, can be found in (Basseville and Nikiforov, 1993).

In (Basseville and Nikiforov, 1993), many applications of decision tools for detecting changes have been summarized, among them quality control, automatic segmentation of signals, fault detection and monitoring in industrial plants and navigation systems, seismic data processing, and some others. One of possible applications of change detection tools that has recently emerged may be in the Replication-Selection Systems (RSS) (Murgu, 2002). Genetic algorithms (GAs) used in the RSS require a fitness function that informs the GA of what is the best performer within a population of performers. Selection of the best performer is very close to discriminating between several contesting hypotheses, which is the core of any detection problem. By virtue of this fact, it is quite natural to search for a suitable fitness function for RSS amongst the well developed detection methods.

One of possible approaches to detect changes that is frequently used in the engineering literature, is based upon the intuitive idea of detecting a nonadditive change by monitoring the innovation sequence (IS) $\nu(t_i)$ of an autoregression (AR) model with the aid of a test for its variance $\sigma^2$. The CUSUM

$$S_k = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} s_i \quad \text{with} \quad s_i = \frac{1}{\sqrt{2}} \left( \frac{\nu(t_i)^2}{\sigma_0^2} - 1 \right) \tag{1.1}$$

is asymptotically, when $k$ goes to infinity, distributed as the standard Gaussian law: $S_k \sim \mathcal{N}(0;1)$ under the hypothesis $\mathbf{H}_0 : \ \sigma^2 = \sigma_0^2$.

This algorithm, which is monitoring $s_i$ termed the weighted squared innovations (WSI), was introduced in the AR case independently in (Jones et al., 1970), (Borodkin and Mottl, 1976), and (Segen and Sanderson, 1980) for the purpose of automatic segmentation of electroencephalogram signals. This test assumes that the IS is white, and thus a test for whiteness should be done first (Mehra and Peschon, 1971). It is well known that the IS is white within a filter if the filter is Kalman, i. e., optimal one (Maybeck, 1979), and whiteness of the IS is only necessary, not sufficient condition for filter optimality (Boozer, 1971). Autocorrelation properties of the IS after change may be very complex, depending on the change occurred. Because of this, monitoring squared innovations has received a criticism as having its poor behavior in many circumstances. The criticism, as well as the ease of implementation of this method, stems from the fact that there no information is used about the after change characteristics.

In the present paper we advocate the WSI method by showing that monitoring WSI will, in certain RSS systems, serve as a useful tool due to its tolerable performance and moderately simple computational cost, which promises the overall cost reduction when the size of population is large.

The paper formulates and develops the WSI method as applied to stochastic state estimation and control systems (Sections 2 and 3). In Section 4, the WSI test is shown to be a useful tool of GAs for the best feedback selection in view of improved performances and condition-based maintenance. Section 5 of the paper presents some experimental results of fault detection for an example relating to inertial navigation system modeling. Section 5 concludes the main content and Appendix completes the paper.

## 2. PROBLEM FORMULATION

Let the closed-loop control system be parameterized by a vector $\theta \in \mathbb{R}^p$ and the available data $z = \left[\begin{smallmatrix} y \\ u \end{smallmatrix}\right]$ be a vector composed of two parts: the control input $u \in \mathbb{R}^q$ and the measurement output $y \in \mathbb{R}^m$. The system is modeled for $i \in \mathbb{Z}$ by the equations

$$\begin{aligned}
x(t_{i+1}) &= \Phi_\theta x(t_i) + \Psi_\theta u(t_i) + w(t_i), \quad x \in \mathbb{R}^n && (2.1)\\
y(t_i) &= H_\theta x(t_i) + v(t_i), \quad y \in \mathbb{R}^m && (2.2)\\
\hat{x}_0(t_{i+1}^-) &= \Phi_0 \hat{x}_0(t_i^+) + \Psi_0 u(t_i), \quad \hat{x}_0 \in \mathbb{R}^n && (2.3)\\
\hat{x}_0(t_i^+) &= \hat{x}_0(t_i^-) + K_0 \nu(t_i), \quad \nu(t_i) = y(t_i) - H_0 \hat{x}_0(t_i^-) && (2.4)\\
u(t_i) &= f_{\mathrm{R}}[\hat{x}_0(t_i^+)] = -G_0^\star \hat{x}_0(t_i^+), \quad u \in \mathbb{R}^q && (2.5)
\end{aligned}$$

with $\{w(\cdot)\}$, $\{v(\cdot)\}$ being zero mean i.i.d. sequences of covariances $Q_\theta$ and $R_\theta$ respectively. The state difference equation (2.1) is propagated forward from the initial condition $x(t_{-s})$ of the finite expectation $\mathbf{E}\left\{\|x(t_{-s})\|^2\right\} < \infty$. The initial state is placed at some $t_{-s} \in \mathbb{R}$ where $s > 0$ determines what is termed settling time $T_s = t_0 - t_{-s}$ needed to think of all processes in (2.1)–(2.5) as wide-sense stationary for $i \geq 0$. As usual, equation (2.1) represents a plant, equation (2.2) a sensor, and (2.3)–(2.5) a feedback. The feedback is composed of a Kalman-like filter (2.3)–(2.4) cascaded with a regulator (2.5). The regulator is described by a function $f_{\mathrm{R}}[\cdot]$ of the measurement updated estimate $\hat{x}_0(t_i^+)$ or can be chosen according to the second equality in (2.5) with a matrix $G_0^\star$.

Matrices of the given system (2.1)–(2.2) are assumed to be known as $\Phi_0$, $\Psi_0$, $Q_0$, $H_0$ and $R_0$ for a *nominal mode*, i. e., for a *nominal value* $\theta_0$ of the uncertainty parameter $\theta$. For this mode, to guarantee existence of the steady-state filter with Kalman gain $K_0$, we assume that the pair $(\Phi_0, Q_0^{1/2})$ is *stabilizable*, the pair $(\Phi_0, H_0)$ is *observable*, and the pair $(\Phi_0, \Psi_0)$ is *controllable*. Matrix $G_0^\star$ can be taken then as a given one or designed to be LQG optimal for the nominal mode of operation (Maybeck, 1982), (Caines, 1988), (Mosca, 1995).

Parameter $\theta$ is subject to an abrupt change at an unknown time point $t_{\mathrm{c}} \in (t_0, t_i)$. This can be viewed as a switch of $\theta$ from $\theta_0$ to some other unknown value $\theta_1$; to detect the change point, a *decision generator*, $\mathcal{DG}$, is needed (Figure 1). Formalizing the problem of $\mathcal{DG}$ synthesis, we consider

$$\left.\begin{aligned}
x_j(t_{i+1}) &= \Phi_j x_j(t_i) + \Psi_j u(t_i) + w_j(t_i)\\
y_j(t_i) &= H_j x_j(t_i) + v_j(t_i)
\end{aligned}\right\} \ \mathcal{S}_0 \ \ (\text{for } j=0) \quad \text{or} \quad \mathcal{S}_1 \ \ (\text{for } j=1) \tag{2.6}$$

as the generalized description of the two systems: $\mathcal{S}_0$ the system with $\theta = \theta_0$; $\mathcal{S}_1$ the system with $\theta = \theta_1$. Filter (2.3)–(2.4), which is denoted by $\mathcal{F}_0$ in Figure 1, is designed as the Kalman filter for $\mathcal{S}_0$, i. e., satisfying equations

$$\left.\begin{aligned}
K_0 &= \widetilde{P_0} H_0^T C_0^{-1}, & C_0 &= H_0 \widetilde{P_0} H_0^T + R_0\\
\widehat{P_0} &= \widetilde{P_0} - \widetilde{P_0} H_0^T C_0^{-1} H_0 \widetilde{P_0}, & \widetilde{P_0} &= \Phi_0 \widehat{P_0} \Phi_0^T + Q_0
\end{aligned}\right\} \ \text{filter } \mathcal{F}_0 \ \ \text{for system } \mathcal{S}_0 \tag{2.7}$$

and processing data

$$y(t_i) = \begin{cases} y_0(t_i) & \text{if} \quad t_i < t_{\mathrm{c}} \quad (\text{data from } \mathcal{S}_0)\\ y_1(t_i) & \text{if} \quad t_i \geq t_{\mathrm{c}} \quad (\text{data from } \mathcal{S}_1)\end{cases} \tag{2.8}$$

The problem is to detect the change point $t_{\mathrm{c}}$ with a reasonable delay using a suitable decision rule $d_0(t_k) \in \{0,\ 1\}$ according to Figure 1.

## 3. WSI BASED DECISION GENERATING METHOD

Many systems, not only (2.1)–(2.5), have the IS, hence the WSI of $s_i$ type in (1.1) can be considered. Because covariance properties of the WSI after change may become very complex, depending on the change occurred, we simplify matters by approximating the covariance kernel function $\mathcal{K}_{ss}(j)$ of $s_i$ after change equal to the exponential function:

$$\mathcal{K}_{ss}(j) = D_s r^{|j|}, \quad j \in \mathbb{Z}, \quad r = \exp\{-\tau_{\mathrm{s}}/\tau_{\mathrm{c}}\}, \quad D_s \triangleq \mathbf{E}\left\{(s_i - \mathbf{E}\{s_i\})^2\right\} \tag{3.1}$$

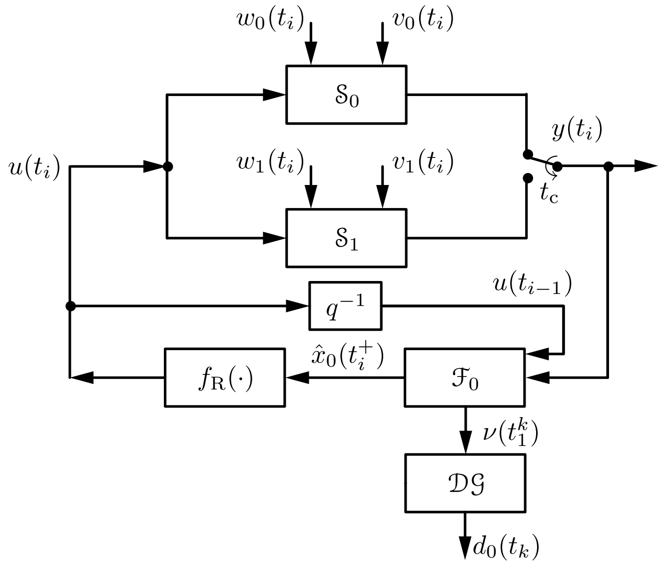


***Figure 1***. The general framework for the problem and solution. *Legend:* $\mathcal{S}_0$ is the system with $\theta = \theta_0$ while $\mathcal{S}_1$ is the system with $\theta = \theta_1$; $\mathcal{F}_0$ is the Kalman filter for $\mathcal{S}_0$; and $\mathcal{DG}$ is a decision generator; $q^{-1}$ denotes a one sample memory (unit delay); $\nu(t_1^k) \triangleq \mathrm{col}[][\nu(t_1), \nu(t_2), \cdots, \nu(t_k)]$ .

where $\tau_s$ is a sampling time interval and $\tau_c$ a correlation interval of $s_i$. Under this assumption, we state the WSI based method for detecting changes as follows.

**Theorem 3.1** *Let a dynamical system with the IS $\{\nu(t_i)\}$, $\nu(t_i) \in \mathbb{R}^m$, have two possible modes of operation termed "normal functioning" (hypothesis $\mathbf{H}_0$) and "a fault" (hypothesis $\mathbf{H}_1$). Under both hypotheses, $\{\nu(t_i)\}$ is a Gaussian sequence with $\mathbf{E}\{\nu(t_i)\} = 0$. Under hypothesis $\mathbf{H}_0$, $\{\nu(t_i)\}$ is an independent sequence with the known covariance $C_0 = L_0 L_0^T$ of each element $\nu(t_i)$ where $L_0$ is the square root of $C_0$ determined, for instance, by Cholesky decomposition. Under hypothesis $\mathbf{H}_1$, $\{\nu(t_i)\}$ is a correlated sequence with an unknown covariance $C_1 \neq C_0$ of each element $\nu(t_i)$. Define*

$$\mu_i \triangleq L_0^{-1}\nu(t_i)\,, \quad s_i \triangleq \sqrt{\frac{m}{2}}\left(\frac{1}{m}\mu_i^T \mu_i - 1\right)\,, \quad S_k \triangleq \frac{1}{\sqrt{k}}\sum_{i=1}^{k} s_i \tag{3.2}$$

*and suppose that under hypothesis $\mathbf{H}_1$, the covariance kernel function of this $s_i$ has been modeled (maybe rather approximately) by expression (3.1).*

*Then asymptotically, when $k$ goes to infinity, the following properties hold:*

1. *The probability laws $\mathcal{L}(\cdot)$ of $S_k$ under these hypotheses satisfy* †

$$\mathbf{H}_0 : \quad \mathcal{L}(S_k) \quad \rightsquigarrow \mathcal{N}(0,1) \tag{3.3}$$

$$\mathbf{H}_1 : \quad \mathcal{L}(S_k) \quad \rightsquigarrow \mathcal{N}(m_{S_k}, D_{S_k}) \tag{3.4}$$

*where $\mathcal{N}(0;1)$ stands for the Gaussian (normal) distribution with zero mean and unit variance, $\mathcal{N}(m_{S_k}, D_{S_k})$ is the normal distribution whose mean $m_{S_k}$ and variance $D_{S_k} = \sigma_{S_k}^2$ are given by*

†The notation $\rightsquigarrow$ corresponds to the weak convergence (Roussas, 1972).

*the below formulae*

$$\left.\begin{aligned} m_{S_k} &\triangleq \mathbf{E}\left\{S_k\right\} = m_s\sqrt{k}\,, \quad m_s \triangleq \mathbf{E}\left\{s_i\right\} = (1/\sqrt{2m})\,\mathrm{tr}\{\Delta\} \\ D_{S_k} &\triangleq \mathbf{E}\left\{(S_k - m_{S_k})^2\right\} = D_s\frac{1+r}{1-r}\,, \quad D_s = 1 + \frac{2}{m}\,\mathrm{tr}\{\Delta\} + \frac{1}{m}\|\Delta\|^2 \\ \Delta &= L_0^{-1}(C_1 - C_0)L_0^{-T}\,, \quad \|\Delta\|^2 \triangleq \sum_{i,j=1}^{m}|\Delta_{ij}|^2 \end{aligned}\right\} \tag{3.5}$$

2. *If $S_k$ of (3.2) is used as a decision function at a time instant $k$ in the decision rule*

$$|S_k| \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} h \tag{3.6}$$

*with $h$ being a conveniently chosen threshold (for example, $h = 3$), then probability of false alarm, $P_F$, and probability of detection, $P_D$, satisfy the following approximate expressions (they are exact if laws $\mathcal{L}(S_k)$ in (3.3), (3.4) are taken to be normal):*

$$\left.\begin{aligned} P_F &\simeq 1 - \phi(h) \\ P_D &\simeq 1 - \frac{1}{2}\left[\phi\left(\frac{m_{S_k}+h}{\sigma_{S_k}}\right) - \phi\left(\frac{m_{S_k}-h}{\sigma_{S_k}}\right)\right] > \frac{1}{2}\left[1 + \phi\left(\frac{m_{S_k}-h}{\sigma_{S_k}}\right)\right] \end{aligned}\right\} \tag{3.7}$$

*through the standard probability integral*

$$\phi(x) \triangleq \frac{2}{\sqrt{2\pi}}\int_0^x \exp\left(-t^2/2\right)\mathrm{d}t$$

3. *If the threshold $h$ in (3.6) is chosen as $h = \phi^{-1}(1-\alpha)$ to guarantee a given level $\alpha$ of $P_F$ in (3.7), $P_F = \alpha$, then hypothesis $\mathbf{H}_1$ is detected with probability $P_D = 1 - \beta$ (where $0 < \beta < 1/2$) not later than after*

$$k_a^\star \simeq \left[\phi^{-1}(1-\alpha) + \sigma_{S_k}\phi^{-1}(1-2\beta)\right]^2/m_s^2 \tag{3.8}$$

*disctere time instants where $x = \phi^{-1}(y)$ is the solution for $\phi(x) = y$.*

The proof of Theorem 3.1 is placed in Appendix A.

Clearly sequence $S_k$ can be used for any system having such an innovation $\nu(t_i)$ to detect the changes marked by matrix $\Delta$ with $\mathrm{tr}\{\Delta\} \neq 0$, with the aid of rule (3.6). In so doing, it is important to keep in mind the following merits of the rule:

1. Rule (3.6) contains the same quadratic forms as does the log-likelihood function

$$\ln p\left(\nu(t_1^k)\;\middle|\;\mathbf{H}_0\right) = -\frac{km}{2}\ln(2\pi) - \frac{k}{2}\ln|C_0| - \frac{1}{2}\sum_{i=1}^{k}\nu^T(t_i)C_0^{-1}\nu(t_i)$$

with $\nu(t_1^k) \triangleq \mathrm{col}[][\nu(t_1), \nu(t_2), \cdots, \nu(t_k)]$ because

$$S_k = \sqrt{\frac{m}{2k}}\sum_{i=1}^{k}\left(\frac{1}{m}\nu^T(t_i)C_0^{-1}\nu(t_i) - 1\right) \tag{3.9}$$

2. The sufficient statistic $\ell(\cdot) \triangleq \ell\left(\nu(t_1^k)\right)$ for detecting changes in the system parameters subject to the assumptions of Theorem 3.1, is (Van-Trees, 1968, Sec. 2.6.2)

$$\ell(\cdot) = \sum_{i=1}^{k} \nu^T(t_i) C_0^{-1} \nu(t_i) - \nu^T(t_1^k) \mathcal{C}_1^{-1}(k) \nu(t_1^k)\ , \ \mathcal{C}_1(k) \triangleq \mathbf{E}\left\{\nu(t_1^k)\nu^T(t_1^k) \mid \mathbf{H}_1\right\}$$

   The evident theoretical drawback of the method is that the decision function $S_k$, (3.2) or equivalently (3.9), is not a sufficient statistic; nevertheless it has found practical use in many applications since its invention, as working without any information about the model after the change, for example in checking filter optimality (Semoushin, 1979).

3. The type of parametric changes detectable by rule (3.6) is restricted to those for which $\mathrm{tr}\{\Delta\} \neq 0$.

4. To strictly satisfy the condition $\mathrm{tr}\{\Delta\} = 0$ in most real situations is hardly probable and because of this the restriction $\mathrm{tr}\{\Delta\} \neq 0$ should not be considered as critical from the practical point of view.

**Example 3.1.** Let $m$ be equal 3 and

$$C_0 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 8 & 2 \\ 3 & 2 & 14 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 2.0 & 1.8 & 2.4 \\ 1.8 & 9.0 & 1.0 \\ 2.4 & 1.0 & 13.0 \end{bmatrix}$$

Then

$$L_0 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 2 & 0 \\ 3 & -2 & 1 \end{bmatrix}, \qquad \Delta = \begin{bmatrix} 1.0 & -1.1 & -5.8 \\ -1.1 & 1.45 & 6.3 \\ -5.8 & 6.3 & 31.0 \end{bmatrix}, \qquad m_s \approx 13.7\,, \qquad D_s \approx 395.$$

Let $\alpha = 0.005$ and $\beta = 0.005$. We obtain, $h = \phi^{-1}(1-\alpha) = 2.8$ and $\phi^{-1}(1-2\beta) = 2.6$ .

If $r = 0.5$, then $k_a^\star \simeq 46$ and if $r = 0.95$, then $k_a^\star \simeq 562$. Hence, starting from the anticipated (for instance, from the most probable or most dangerous) faults as the base, and reasoning from the specified quality $\alpha$ and $\beta$ of the decision rule (3.6), it is possible to numerically predict the lapse of time[‡)] $k_a^\star$ after which the fault will be detected with confidence not worse than $P_D = 1 - \beta$.

For the above example, the condition of non-observable faults is readily obtained as the equation

$$\mathrm{tr}\{\Delta\} = 27a_{11} - 11a_{12} - 10a_{13} + \tfrac{5}{4}a_{22} + 2a_{23} + a_{33} = 0 \tag{3.10}$$

in the entries $a_{ij} = a_{ji}$ of $A = C_1 - C_0$ . Keeping $C_1$ within bounds of positive definiteness, one can choose $C_1$ to match the condition (3.10), for example

$$C_1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 8 & 1 \\ 3 & 1 & 16 \end{bmatrix}, \quad A \triangleq C_1 - C_0 = [a_{ij}] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

This simple example makes it apparent that only certain of the faults specially chosen to fit the equation $\mathrm{tr}\{\Delta\} = 0$ to a T cannot be detected by the method.

Theorem 3.1 assumes that changes in $\theta$ do not create a bias (i.e., non-zero expectation) in the innovation sequence $\{\nu(t_i)\}$ :

$$\mathbf{H}_0 : \ \nu(t_i) \sim \mathcal{N}(0, C_0) \qquad \text{and} \qquad \mathbf{H}_1 : \ \nu(t_i) \sim \mathcal{N}(0, C_1) \tag{3.11}$$

[‡)] In this example, the delay $k_a^\star$ should be increased by $k_{\mathrm{norm}} \gtrapprox 30$ needed to take distributions in (3.3)–(3.4) to be approximately normal, from the engineering point of view.

If this is not the case, one should turn from (3.11) to

$$\mathbf{H}_0 : \ \nu(t_i) \sim \mathcal{N}(0, C_0) \qquad \text{and} \qquad \mathbf{H}_1 : \ \nu(t_i) \sim \mathcal{N}(m_1, C_1) \tag{3.12}$$

with some $m_1 \neq 0$. This entails the only modification in Theorem 3.1, namely, in $\Delta$ of formulae (3.5):

$$\Delta = L_0^{-1}(C_1 + m_1 m_1^T - C_0)L_0^{-T} \qquad \text{instead of} \qquad \Delta = L_0^{-1}(C_1 - C_0)L_0^{-T} \tag{3.13}$$

Thus, for appreciable quantity of faults, the WSI based method boils down to a straightforward cumulation of $S_k$, (3.2), through the sequential algorithm

$$S_k = S_{k-1}\sqrt{1 - 1/k} + s_k/\sqrt{k}\ , \quad k \in \mathbb{N}\ , \quad S_0 = 0 \tag{3.14}$$

followed by the rule (3.6). When testing $\mathbf{H}_1$ against $\mathbf{H}_0$, let us denote $S_k$ in (3.14) by $S_k^{(0)}$ thus emphasizing by superscript $^{(0)}$ that the main hypothesis is $\mathbf{H}_0$. Then we have two schemes with the stopping rule $t_\mathrm{a}$ (the alarm time) for the alternative $\mathbf{H}_1$ detection:

**(A)** decision is made at any current time $t_k$:

$$d_0(t_k) = \begin{Bmatrix} 0 & \text{if} & \left|S_k^{(0)}\right| < h\ ; & \mathbf{H}_0 \ \text{is chosen} \\ 1 & \text{if} & \left|S_k^{(0)}\right| \geq h\ ; & \mathbf{H}_1 \ \text{is chosen} \end{Bmatrix}, \quad t_\mathrm{a} = \min\left\{t_k : \ d_0(t_k) = 1\right\}$$

**(B)** decision is made at the end of a sample number $l = 1, 2, \ldots, L$ each of size $N$:

$$d_0(l) = \begin{cases} 0 & \text{if } \forall\, k = 1, 2, \ldots, N : \ \left|S_{N(l-1)+k}^{(0)}\right| < h\ ; \quad \mathbf{H}_0 \ \text{is chosen} \\ 1 & \text{if } \exists\, k = 1, 2, \ldots, N : \ \left|S_{N(l-1)+k}^{(0)}\right| \geq h\ ; \quad \mathbf{H}_1 \ \text{is chosen} \end{cases}$$

$$t_\mathrm{a} = \tau_\mathrm{s}\left[N(l-1) + \min\left\{k : \ \left|S_{N(l-1)+k}^{(0)}\right| \geq h\right\}\right] \overset{r}{=} \tau_\mathrm{s} N \min\left\{l : \ d_0(l) = 1\right\}$$

where $\overset{r}{=}$ denotes the rounded up equality.

This result is of general significance because many stochastic dynamical systems have (or can be supplied with) a whitening filter that generates a process of innovations. It fits naturally into the systems of (2.1)–(2.5) type, in which the whitening filter exists in the Kalman filter form (2.3)–(2.4) with the IS $\{\nu(t_i)\}$.

## 4. WSI BASED FEEDBACK FILTER TESTING AND SELECTION

From the innovation process theory, two facts are known (Martin and Stubberud, 1976):

- **(i)** Case (3.11) holds if changes happen only in covariances $Q_\theta$ and $R_\theta$ .
- **(ii)** The only cause for Case (3.12) are changes in matrices $\Phi_\theta$, $\Psi_\theta$ and/or $H_\theta$ .

If Case (3.12) takes place, (3.13) with $m_1$ calculated by the general procedure described, for example in (Martin and Stubberud, 1976), should be used.

**Example 4.1.** Consider the simple case when $n = m = 1$, $u(t_i) \equiv 0$, $Q_1 = Q_0$, $R_1 = R_0$, $H_1 = H_0 = 1$ and $\Phi_1 \neq \Phi_0$ (both $|\Phi_j| < 1,\ j = 0, 1$). At $t_i \gg t_\mathrm{c}$ (i.e., under $\mathbf{H}_1$ when the change transient response effects are vanished) from (2.6) (for $j = 0$), (2.7), (2.8) and (2.3)–(2.4), we have

$$\nu(t_i) = x_1(t_i) + v_0(t_i) - \hat{x}_0(t_i) = e(t_i) + v_0(t_i)\ , \quad e(t_i) \triangleq x_1(t_i) - \hat{x}_0(t_i) \tag{4.1}$$

Denoting

$$\left.\begin{array}{ll} m_1(t_i) \triangleq \mathbf{E}\left\{\nu(t_i) \,\middle|\, \mathbf{H}_1\right\}, & m_e(t_i^-) \triangleq \mathbf{E}\left\{e(t_i)\right\} \\ m_{x_1}(t_i) \triangleq \mathbf{E}\left\{x_1(t_i)\right\}, & m_{\hat{x}_0}(t_i^{\pm}) \triangleq \mathbf{E}\left\{\hat{x}_0(t_i^{\pm})\right\} \end{array}\right\} \tag{4.2}$$

we obtain from (4.1), (2.6) (for $j = 1$) and (2.3)

$$\left.\begin{array}{ll} m_1(t_i) = m_e(t_i^-), & m_e(t_i^-) = m_{x_1}(t_i) - m_{\hat{x}_0}(t_i^-) \\ m_{x_1}(t_i) = \Phi_1 m_{x_1}(t_{i-1}), & m_{\hat{x}_0}(t_i^-) = \Phi_0 m_{\hat{x}_0}(t_{i-1}^+) \end{array}\right\} \tag{4.3}$$

Taking expectation of (2.4) yields

$$m_{\hat{x}_0}(t_i^+) = (1 - K_0) m_{\hat{x}_0}(t_i^-) + K_0 m_{x_1}(t_i)$$

Taking here a step backwards and then substituting this into the last formula of (4.3) give

$$m_{\hat{x}_0}(t_i^-) = \Phi_0\left[(1 - K_0) m_{\hat{x}_0}(t_{i-1}^-) + K_0 m_{x_1}(t_{i-1})\right]$$

Then we find consequently using (4.2) and (4.3)

$$\begin{aligned} m_e(t_i^-) &= m_{x_1}(t_i) - \Phi_0\left[(1 - K_0) m_{x_1}(t_{i-1}) - (1 - K_0) m_e(t_{i-1}^-) + K_0 m_{x_1}(t_{i-1})\right] \\ &= m_{x_1}(t_i) - \Phi_0\left[m_{x_1}(t_{i-1}) - (1 - K_0) m_e(t_{i-1}^-)\right] \\ &= (\Phi_1 - \Phi_0) m_{x_1}(t_{i-1}) + \Phi_0(1 - K_0) m_e(t_{i-1}^-) \\ m_1(t_i) &= \Phi_0(1 - K_0) m_1(t_{i-1}) + (\Phi_1 - \Phi_0) m_{x_1}(t_{i-1}) \end{aligned}$$

As $0 < \Phi_0 < 1$ and $0 < \Phi_0(1 - K_0) < 1$, there exist the steady-state values

$$m_1 \triangleq \lim_{t_i \to \infty}\{m_1(t_i)\} = \lim_{t_i \to \infty}\{m_1(t_{i-1})\}, \qquad m_{x_1} \triangleq \lim_{t_i \to \infty}\{m_{x_1}(t_{i-1})\}$$

and finally

$$m_1 = \frac{\Phi_1 - \Phi_0}{1 - \Phi_0(1 - K_0)} m_{x_1}$$

**Remark 4.1.** By virtue of $0 < \Phi_0 < 1$, $m_{x_1} = 0$. It is also worthy of note that although $\mathrm{tr}\{\Delta\} = 0$ is not a necessary or sufficient condition for filter optimality, the quantity of faults satisfying $\mathrm{tr}\{\Delta\} = 0$ is relatively thin.

As for the case of two contesting hypotheses $\mathbf{H}_0$ and $\mathbf{H}_1$ in Section 3, we can apply one of two schemes, (**A**) or (**B**), of decision making. However as differentiated from that case, here every singly taken $\mathbf{H}_j$ $(j = 0, 1, \ldots, K)$ contests individually against any other $\mathbf{H}_i$ $(i = 0, 1, \ldots, K; i \neq j)$ using its own decision function $S_k^{(j)}$ that equals to $S_k$ and is formed according to (3.14) supplied with the superscript $^{(j)}$. Again, we have two schemes:

**(A)** decision is made at any current time $t_k$. For any particular $\mathbf{H}_j$, $j \in \{0, 1, \ldots, K\}$, the particular decision rule is

$$d_j(t_k) = \begin{cases} 0 & \text{if } \left|S_k^{(j)}\right| < h\,; \quad \mathbf{H}_j \text{ is supported} \\ 1 & \text{if } \left|S_k^{(j)}\right| \geq h\,; \quad \mathbf{H}_j \text{ is rejected} \end{cases}$$

The over-all alarm time is

$$t_{\mathrm{a}} = \min\left\{t_k : \left(\exists \kappa : d_\kappa(t_k) = 0 \ \& \ \forall_{j \neq \kappa}\, d_j(t_k) = 1\right)\right\}$$

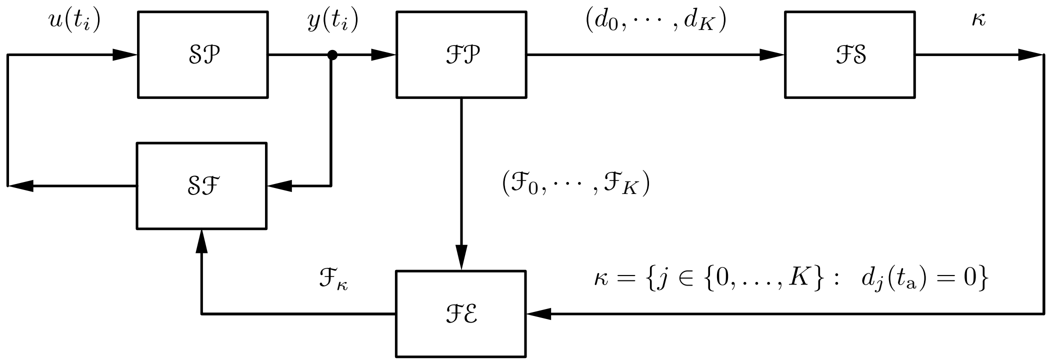


***Figure 2***. The best filter selection. *Legend:* $\mathcal{SP}$ stands for system population; $\mathcal{FP}$ stands for filter population, $\mathcal{FS}$ for filter selection, $\mathcal{FE}$ for feedback effector, and $\mathcal{SF}$ stands for the system feedback. The stopping rule is given by the alarm time $t_{\mathrm{a}} = \min\Big\{t_k : \; \big(\exists \kappa : \; d_\kappa(t_k) = 0 \;\; \& \; \underset{j\neq\kappa}{\forall} \; d_j(t_k) = 1\big)\Big\}$ . At this time, $\mathcal{FE}$ uploads the selected $\mathcal{F}_\kappa$ into $\mathcal{SF}$ if by then it is not in there.

with the over-all winner selection rule:

$$\kappa = \{j \in \{0, 1, \ldots, K\} : d_j(t_{\mathrm{a}}) = 0\} \;\; ; \quad \mathbf{H}_\kappa \;\text{ is chosen}$$

**(B)** decision is made at the end of a sample number $l = 1, 2, \ldots, L$ each of size $N$. For any particular $\mathbf{H}_j, \; j \in \{0, 1, \ldots, K\}$, the particular decision rule is

$$d_j(l) = \begin{cases} 0 & \text{if } \forall\, k = 1, 2, \ldots, N : \; \left|S^{(j)}_{N(l-1)+k}\right| < h \, ; \quad \mathbf{H}_j \;\text{ is supported} \\ 1 & \text{if } \exists\, k = 1, 2, \ldots, N : \; \left|S^{(j)}_{N(l-1)+k}\right| \geq h \, ; \quad \mathbf{H}_j \;\text{ is rejected} \end{cases} \tag{4.4}$$

$$t_{\mathrm{a}} \stackrel{r}{=} \tau_{\mathrm{s}} N \min\left\{ l : \; \left( \exists \kappa : d_\kappa(l) = 0 \;\; \& \; \underset{j\neq\kappa}{\forall} \; d_j(l) = 1 \right) \right\} \tag{4.5}$$

with the over-all winner selection rule:

$$\kappa = \{j \in \{0, 1, \ldots, K\} : d_j(l) = 0\} \;\; ; \quad \mathbf{H}_\kappa \;\text{ is chosen} \tag{4.6}$$

This method is realized in a framework corresponding to Figure 2.

## 5. SOME COMPUTATIONAL EXPERIMENT RESULTS

To have confidence in the freedom from failures of airborne equipment or changes in vehicle parameters is of critical importance for optimal navigation data processing.

As an example for the problem under study, we take the damped Shuler loop driven by the exponentially correlated noise, whose description was first given in (Gaines, 1971). In compliance with (Gaines, 1971), consider equations

$$\left.\begin{aligned} x_j(t_{i+1}) &= \Phi_j x_j(t_i) + \Gamma_j w_j(t_i) \\ y_j(t_i) &= H_j x_j(t_i) + v_j(t_i) \end{aligned}\right\} \; \mathcal{S}_0 \;\; (\text{for } j = 0) \quad \text{or} \quad \mathcal{S}_j \;\; (\text{for } j = 1, \cdots, K) \tag{5.1}$$

that have the form of (2.6) in the specific case of control input $u(t_i)$ being zero, with $K + 1$ modes of operation, each mode being associated with a system $\mathcal{S}_j$: $\mathcal{S}_0$ is the fault free system and every $\mathcal{S}_j$ (for $j =$

$1, \cdots, K$) is a faulty system. System $\mathcal{S}_0$ corresponds to

$$\Phi_0 = \begin{bmatrix} 0.75 & -1.74 & -0.3 & 0 & -0.15 \\ 0.09 & 0.91 & -0.0005 & 0 & -0.008 \\ 0 & 0 & 0.95 & 0 & 0 \\ 0 & 0 & 0 & 0.55 & 0 \\ 0 & 0 & 0 & 0 & 0.905 \end{bmatrix}, \quad \Gamma_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 24.64 & 0 & 0 \\ 0 & 0.835 & 0 \\ 0 & 0 & 1.83 \end{bmatrix} \tag{5.2}$$

$$Q_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad H_0 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}, \qquad R_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{5.3}$$

Disengaging ourselves at this stage from a physical interpretation of the given characteristics (5.2)–(5.3), we shall restrict our consideration in this section by introducing the set of faulty systems through the following formal relations:

$$\left.\begin{array}{l} \Phi_j = \Phi_0\ , \quad \Gamma_j = \Gamma_0\ , \quad Q_j = Q_0\ , \quad R_j = R_0 \\ H_j = \begin{bmatrix} 1-e & 0 & 0 & 0 & 1-f \\ 0 & 1-g & 0 & 1-h & 0 \end{bmatrix}, \quad \{efgh\} \triangleq j_{[2]} \end{array}\right\} \mathcal{S}_j \ \ (j = 1, \cdots, 15)$$

where $\{efgh\}$ stands for the binary code $j_{[2]}$ of the fault number $j$. For example, given $j = 2$ or $j = 9$, then $\{efgh\} = \{0010\}$ or, correspondingly $\{1001\}$, and in these cases

$$H_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad H_9 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

The change $\mathcal{S}_0 \stackrel{t_c}{\rightarrowtail} \mathcal{S}_j$ $(j = 1, \cdots, 15)$ occurs at $t_c = 2300$, however filter $\mathcal{F}_0$ "does not know" about that and continues to satisfy equations (2.7). Experimental data for all $K = 15$ faulty system cases are shown in Figures 3–4 together with the fault free system case, when $\{efgh\} = \{0000\}$. If the value of threshold $h$ in (3.6) is set to 3, then probability of false alarm, $P_F$, in the case of $\mathcal{S}_0$ will be not greater than 0.3% (due to property (3.3)), and in other cases probability of right detection, $P_D$, comes to 100% due to monotonicity of increasing $|\,\mathbf{E}\{S_k\}\,|$. The value of delay in right detection depends on the type of change and, as it can be seen from Figures 3–4, is in the region from 100 to 1500 sampling time intervals.

The reported simulation results as well as those obtained for many different samples from the sequences $w(t_i)$ and $v(t_i)$ in the model (2.1)–(2.2), show practicability of the approach for inertial navigation systems monitoring and thus cause us to anticipate that it will find a diversity of other applications. This assumption has received further support in our work through the similar results yielded by other experiments.

## 6. SIMULATION TOOL DEVELOPMENT

The first (above shown) experimental plots were made in MATLAB 6 after the results were obtained using ASPID – Adaptive System Parameter Identification – the soft package developed in Visual C 6.0 (service pack 6) to carry on research on Inertial Navigation System error budget (Semoushin and Polosenko, 2002). However, to perform a wide scale investigation of all the proposed detection/selection algorithms that would correspond Figure 2, a special software needs to be developed.

Such a tool, a very convenient one named LSCS (Linear Stochastic Control System), has been created by Mike Sunoplya, a fourth year Ulyanovsk State University student. It allows us to observe the behavior of $S_k$ before and after the system failure in full accordance with Figure 2.

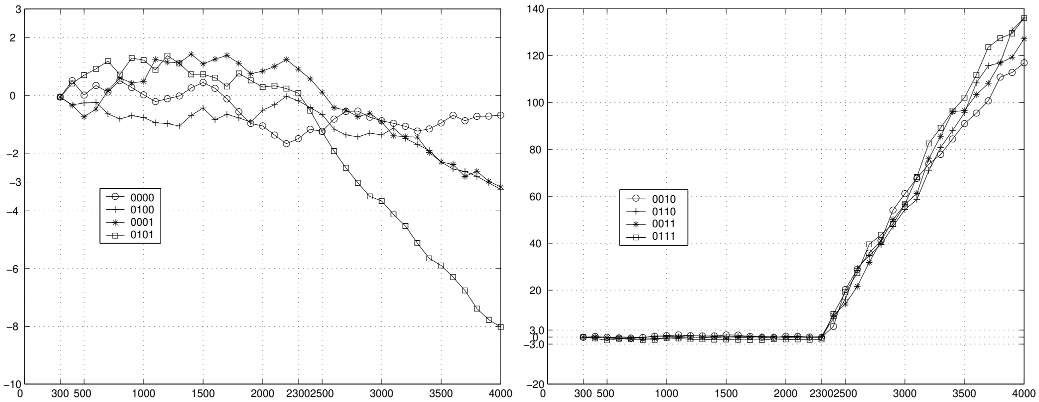


***Figure 3***. Behavior of $S_k$ before and after change at $t_\mathrm{c} = 2300$; *left:* $j = 0, 4, 1, 5$, *right:* $j = 2, 6, 3, 7$.

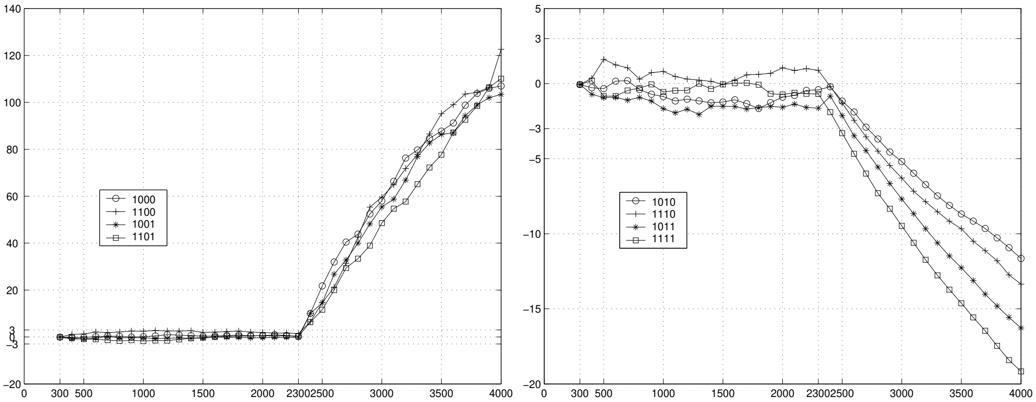


***Figure 4***. Behavior of $S_k$ before and after change at $t_\mathrm{c} = 2300$; *left:* $j = 8, 12, 9, 13$, *right:* $j = 10, 14, 11, 15$.

This LSCS application can work in 1 of 3 available modes:

*normal mode* — application calculates and draws (in real time mode on the work area) the $S_k^{(0)}$ changes, $S_k^{(0)}$ given by the formula (3.14) introduced in Section 3; only one filter $\mathcal{F}_0$ is available.

*mixed mode* — this mode supports four different faulty filters $\mathcal{F}_j$, $j \in \{1, 2, \ldots, K\}$ versus one (nominal) filter $\mathcal{F}_0$; four curves of $S_k^{(j)}$ changes can be taken from $\mathcal{F}_j$ and observed together like in Figures 3–4; if we write $\mathcal{S}_0 \overset{2300}{\rightarrowtail} \mathcal{S}_j \mid j \in \{2, 6, 3, 7\}$, this means that one of four changes $\mathcal{S}_0 \overset{t_\mathrm{c}}{\rightarrowtail} \mathcal{S}_j$ occurs at $t_\mathrm{c} = 2300$ for $j = 2, 6, 3$ *or* $7$, – this is the case of Figure 3, *right*.

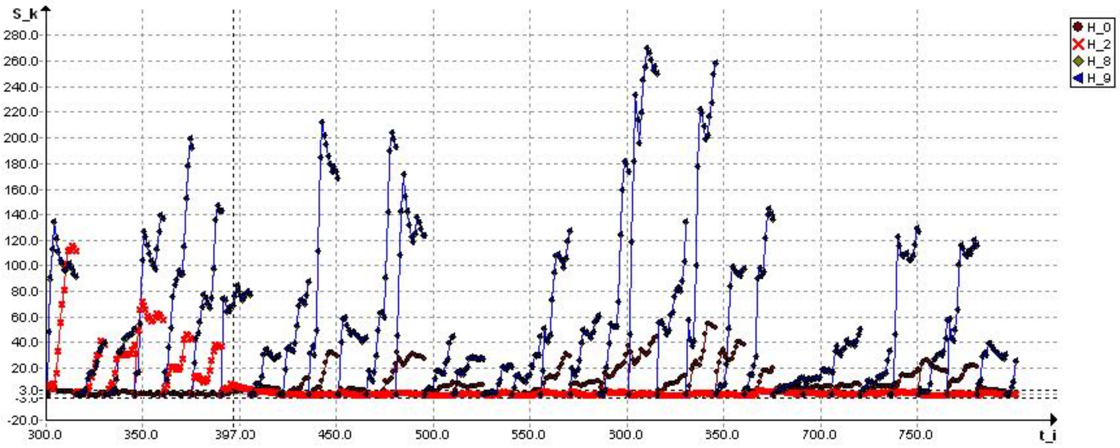


***Figure 5***. $\mathcal{S}_0 \overset{397}{\rightarrowtail} \mathcal{S}_2 \;\big|\; \mathcal{FP} = \{\mathcal{F}_0, \mathcal{F}_2, \mathcal{F}_8, \mathcal{F}_9\}$. Sample size $N = 15$ in Scheme B, (4.4)–(4.6). *Comment:* In this case the change $\mathcal{S}_0 \overset{t_c}{\rightarrowtail} \mathcal{S}_j$ occurs at $t_c = 397$ for $j = 2$. Scheme B, (4.4)–(4.6), is used with each $S_k^{(j)}$, $j \in \{0, 2, 8, 9\}$, reset to zero every $N$ steps of algorithm (3.14). The zoomed version of this plot can be obtained, – *cf.* Figure 6.

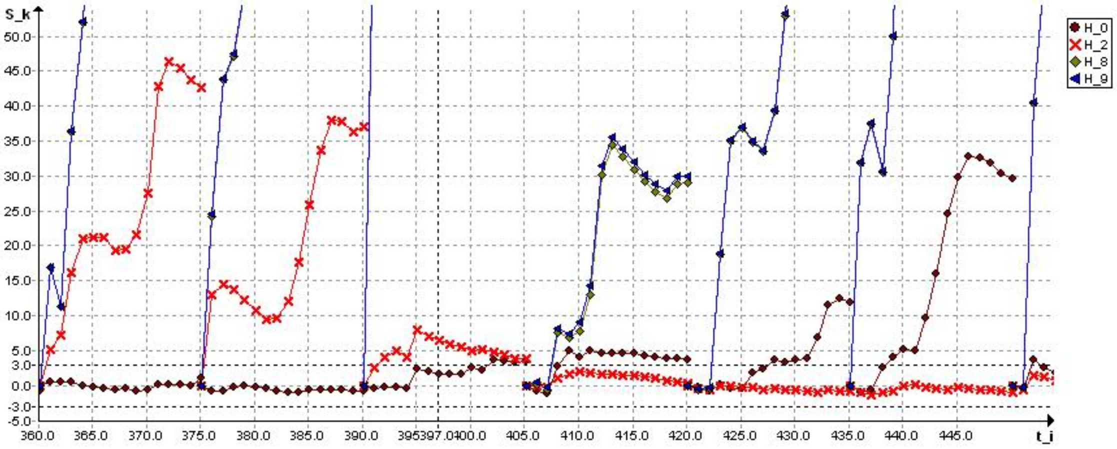


***Figure 6***. Zoom of the previous plot of Figure 5 with $\mathcal{S}_0 \overset{397}{\rightarrowtail} \mathcal{S}_2 \;\big|\; \mathcal{FP} = \{\mathcal{F}_0, \mathcal{F}_2, \mathcal{F}_8, \mathcal{F}_9\}$. Sample size $N = 15$ in Scheme B, (4.4)–(4.6). *Comment:* It can be seen from here that before the change – for $t_i < 397$ – only $S_k^{(0)}$ formed by (3.14) in $\mathcal{F}_0$, remains in the vicinity of zero. After the change from $\mathcal{S}_0$ to $\mathcal{S}_2$ at $t_i = 397$, behavior of the curves changes drastically: $S_k^{(2)}$ formed by (3.14) in $\mathcal{F}_2$, enters the vicinity of zero while $S_k^{(0)}$ leaves it as continue to do so all the other curves, – in this case $S_k^{(8)}$ and $S_k^{(9)}$ taken from $\mathcal{F}_8$ and $\mathcal{F}_9$, correspondingly. *Remark:* On insufficiently reduced scale, one curve can melt into another. In this plot, one can easily discriminate $S_k^{(8)}$ curve from $S_k^{(9)}$ curve between $t_i = 413$ and $t_i = 420$.

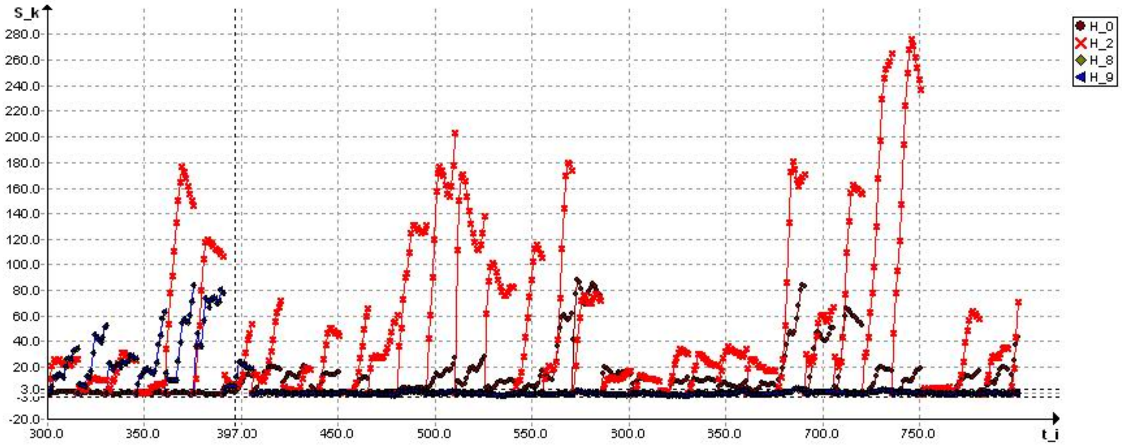


***Figure 7***. $\mathcal{S}_0 \overset{397}{\rightarrowtail} \mathcal{S}_9 \mid \mathcal{FP} = \{\mathcal{F}_0, \mathcal{F}_2, \mathcal{F}_8, \mathcal{F}_9\}$. Sample size $N = 15$ in Scheme B, (4.4)–(4.6). *Comment:* In this case the change $\mathcal{S}_0 \overset{t_c}{\rightarrowtail} \mathcal{S}_j$ occurs at $t_c = 397$ for $j = 9$. Scheme B, (4.4)–(4.6), is used with each $S_k^{(j)}$, $j \in \{0, 2, 8, 9\}$, reset to zero every $N$ steps of algorithm (3.14). The zoomed version of any plot can be obtained. For this case, it is shown in Figure 8 in the form of screenshot.

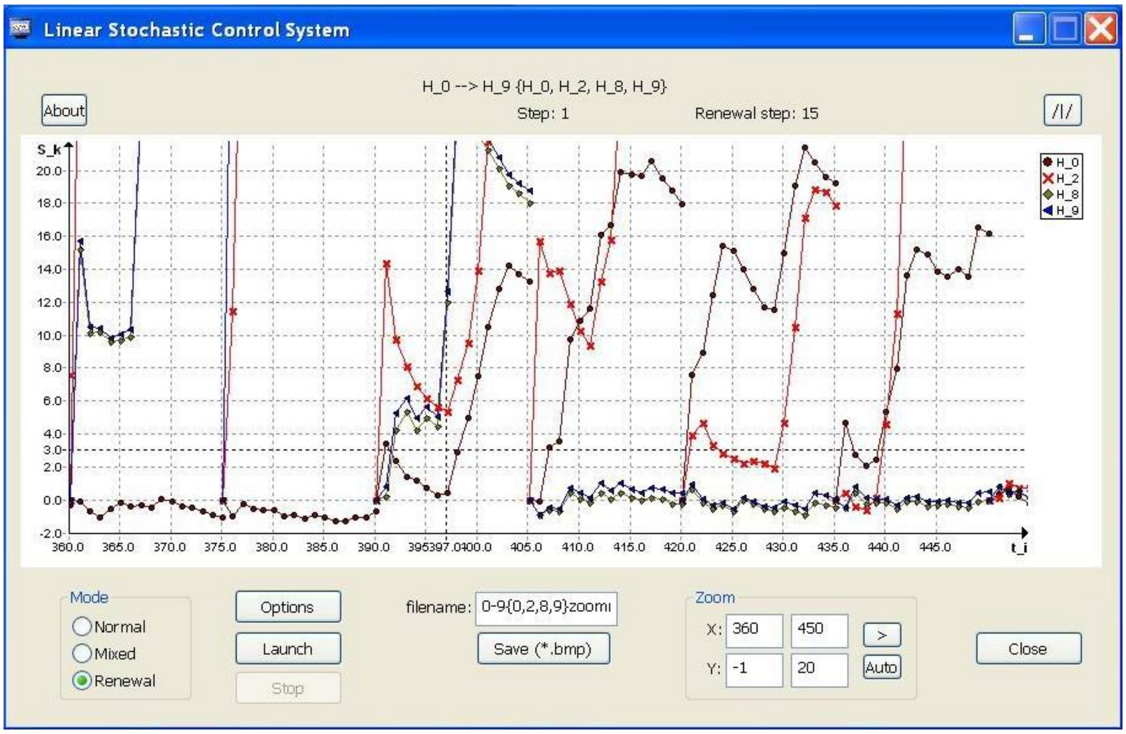


***Figure 8***. Application window. Zoom of the previous plot of Figure 7 with $\mathcal{S}_0 \overset{397}{\rightarrowtail} \mathcal{S}_9 \mid \mathcal{FP} = \{\mathcal{F}_0, \mathcal{F}_2, \mathcal{F}_8, \mathcal{F}_9\}$.

*renewal mode* application calculates and draws (in real time mode on the work area) the $S_k^{(j)}$ changes for all $j \in \{0, \ldots, K\}$; we identify this by the following notation: if we write $\mathcal{S}_0 \overset{397}{\rightarrowtail} \mathcal{S}_9 \,\big|\, \mathcal{FP} = \{\mathcal{F}_0, \mathcal{F}_2, \mathcal{F}_8, \mathcal{F}_9\}$ this means that the change $\mathcal{S}_0 \overset{t_c}{\rightarrowtail} \mathcal{S}_j$ occurs at $t_c = 397$ for $j = 9$; in this mode Scheme B, (4.4)–(4.6), is used with each $S_k^{(j)}$ reset to zero every $N$ steps of (3.14).

LSCS was developed using the Visual C++.NET development system and can work under operation system Windows 98/2000/XP. It provides:

- on-line plotting
- multi-mode support
- manual zoom
- friendly work area that can be saved to bmp-file
- user controlled system matrices $\Phi_0$, $\Gamma_0$ and $H_0$ input and change at any assigned $t_c$
- useful options to create the unique style of curves and legend at the work area
- handy interface
- and more...

With this tool switched onto the renewal mode, many computational experiments were made in this work (most intersing of them are presented in Figures 5–8) and many more are reserved for further research.

## 7. CONCLUSIONS

The weighted squared innovation method is formulated and tested as applied to stochastic state estimation and control systems. The findings of this paper may be of considerable practical value thanks to high detecting capacity of the method at not-too-high computational cost. The understanding of the exact capacity for work of the method awaits further investigation. Further studies will probe practical aspects of change point detecting with respect to changes not only in observation matrix $H$ but in other system matrices, too.

## ACKNOWLEDGEMENTS

The authors would like to thank Prof. Dr. A. Butov of Ulyanovsk State University, Russia, for his interest to this work and useful discussions.

## APPENDIX A: A PROOF OF THEOREM 3.1 FROM SECTION 3

**Lemma A.1.** $\xi \sim \mathcal{N}(0;1)$ implies $\mathbf{E}\left\{\xi^{2k}\right\} = (2k-1)!!$ .

*Proof.*

$$\begin{aligned}
\mathbf{E}\left\{\xi^{2k}\right\} &= \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} x^{2k} \exp\left\{-\tfrac{x^2}{2}\right\} \mathrm{d}x = \tfrac{2^k}{\sqrt{\pi}} \int\limits_{0}^{\infty} t^{k-\frac{1}{2}} \exp\{-t\} \,\mathrm{d}t \\
&= \tfrac{2^k}{\sqrt{\pi}} \Gamma\left(k+\tfrac{1}{2}\right) = \tfrac{2^k}{\sqrt{\pi}} \cdot \left(k-\tfrac{1}{2}\right) \cdot \Gamma\left(k-\tfrac{1}{2}\right) \\
&= (2k-1)\cdot(2k-3)\cdot\ldots\cdot 3\cdot 1
\end{aligned}$$

□

*Proof of Theorem 3.1.* Under hypothesis $\mathbf{H}_0$, $\mu_i \sim \mathcal{N}(0; I)$ is a white Gaussian sequence where $I$ is $m \times m$ identity matrix. If $\mu_{ji}$ denotes the $j$-th element of $\mu_i$, it follows that

$$\xi \triangleq \sum_{i=1}^{k} \sum_{j=1}^{m} \mu_{ji}^2 \sim \mathcal{F}_{\chi^2(mk)}(x)$$

where $\mathcal{F}_{\chi^2(mk)}(x)$ stands for $\chi^2$ distribution with $mk$ degrees of freedom. When $k \to \infty$, $\xi$ is distributed normally with center $mk$ and variance $2mk$, $\mathcal{F}_{\chi^2(mk)}(x) \rightsquigarrow \mathcal{N}(mk; 2mk)$. From (3.2) we have

$$S_k = \sqrt{mk/2} \left\{ \frac{1}{mk} \xi - 1 \right\}$$

therefore $\mathcal{L}(S_k) \rightsquigarrow \mathcal{N}(0; 1)$.

Under hypothesis $\mathbf{H}_1$ we have

$$\begin{aligned} C_\mu &\triangleq \mathbf{E}\left\{ \mu_i \mu_i^T \right\} = L_0^{-1} C_1 L_0^{-T} = I + L_0^{-1}(C_1 - C_0) L_0^{-T} = I + \Delta \\ m_s &\triangleq \mathbf{E}\left\{ s_k \right\} = \sqrt{m/2} \left[ (1/m) \operatorname{tr}\{C_\mu\} - 1 \right] = (2m)^{-1/2} \operatorname{tr}\{\Delta\} \end{aligned}$$

Let $[C_\mu]_{jk}$ be the $jk$-th element of $C_\mu$, then $[C_\mu]_{jk} = \sigma_j \sigma_k \rho_{jk}$ where $\sigma_j^2$, $\sigma_k^2$ are variances of the $j$-th and $k$-th elements of vector $\mu_i$, correspondingly, and $\rho_{jk}$ is the correlation between them. From (3.2) it follows that

$$D_s = \tfrac{1}{2m} \left[ \mathbf{E}\left\{ \left( \mu_i^T \mu_i \right)^2 \right\} - \left( \mathbf{E}\left\{ \mu_i^T \mu_i \right\} \right)^2 \right]$$

Lemma A.1 and straitforward calculations allow us to write

$$\mathbf{E}\left\{ \mu_{i,k}^4 \right\} = 3\sigma_k^4 \,, \quad \mathbf{E}\left\{ \mu_{ji}^2 \mu_{ki}^2 \right\} = \sigma_j^2 \sigma_k^2 \left( 1 + 2\rho_{jk}^2 \right)$$

where $j$ and $k$ are indices of the corresponding elements of $\mu_i$. It follows from here that

$$\begin{aligned} \mathbf{E}\left\{ \left( \mu_i^T \mu_i \right)^2 \right\} &= 3 \sum_{k=1}^{m} \sigma_k^4 + 2 \sum_{\substack{j,k=1 \\ j<k}}^{m} \sigma_j^2 \sigma_k^2 \left( 1 + 2\rho_{jk}^2 \right) \\ \left( \mathbf{E}\left\{ \mu_i^T \mu_i \right\} \right)^2 &= \sum_{k=1}^{m} \sigma_k^4 + 2 \sum_{\substack{j,k=1 \\ j<k}}^{m} \sigma_j^2 \sigma_k^2 \\ D_s &= \tfrac{1}{m} \sum_{j,k=1}^{m} [C_\mu]_{jk}^2 = \tfrac{1}{m} \|C_\mu\|^2 \end{aligned}$$

Because $[C_\mu]_{jj} = 1 + \Delta_{jj}$ and $[C_\mu]_{\substack{jk \\ j\neq k}} = \Delta_{jk}$ where $\Delta_{jj}$ and $\Delta_{jk}$ denote the corresponding entries of $\Delta$, then from the above expression we obtain

$$D_s = 1 + (2/m) \operatorname{tr}\{\Delta\} + (1/m) \|\Delta\|^2$$

Based on definition $s_i$ (3.2), obtain $m_{S_k}$ (3.5) and variance

$$D_{S_k} = D_s + 2 \sum_{j=1}^{k-1} (1 - j/k) \mathcal{K}_{ss}(j)$$

Under approximation (3.1), we have

$$\lim_{k \to \infty} D_{S_k} = D_s (1 + r)/(1 - r)$$

Due to the exponential correlated dependence (3.1), the mixing condition in Birkhoff-Khinchin theorem holds (Yoshihawa, 1992), (Ibragimov and Linnik, 1965), hence convergence (3.4) and $P_D$ in (3.7) are also true.

□

## REFERENCES


Basseville, M. and Nikiforov, I. (1993). *Detection of Abrupt Changes: Theory and Applications*, New York: Prentice-Hall.

Boozer, D. D. (1971). On Innovation Sequence Testing of the Kalman Filter, *Teledyne Brown Engineering* 1: 251–256.

Borodkin, L. I. and Mottl, V. V. (1976). Algorithm for Finding the Jump Times of Random Process Equation Parameters, *Automation and Remote Control* 37(6): 23–32.

Caines, P. E. (1988). *Linear Stochastic Systems*, New York: John Wiley & Sons.

Gaines, H. T. (1971). Flight Test of a Kalman Filter Doppler Aided Strapdown Inertial Navigator, in *Proc. National Aerospace Electronics Conference*, New York.

Ghosh, B. K. and Sen, P. K. (1991). *Handbook of Sequential Analysis*, New York: Marcel Dekker.

Ibragimov, I. A. and Linnik, Y. V. (1965). *Independent and Stationary Related Values*, Moscow: Nauka.

Johnson, N. L. and Leone, F. C. (1962). Comulative Sum Control Charts: Mathematical Principles Applied to Their Construction and Use, I, *Industrial Quality Control* 18: 15–21.

Johnson, N. L. and Leone, F. C. (1963). Comulative Sum Control Charts: Mathematical Principles Applied to Their Construction and Use, II, *Industrial Quality Control* 19: 29–36.

Johnson, N. L. and Leone, F. C. (1964). Comulative Sum Control Charts: Mathematical Principles Applied to Their Construction and Use, III, *Industrial Quality Control* 20: 22–28.

Jones, R. H., Crowell, D. H., and Kapuniai, L. E. (1970). Change Detection Model for Serially Correlated Multivariate Data, *Biometrics* 26(2): 269–280.

Kenett, R. S. and Zacks, S. (1998). *Modern Industrial Statistics: Design and Control of Quality and Reliability*, Pacific Grove: Duxbury Press.

Martin, W. C. and Stubberud, A. R. (1976). Innovation Process in Identification Problems, in *Control and Dynamic Systems – Advances in Theory and Applications*, C. T. Leondes, ed., pp. 121–176, New York: Academic Press.

Maybeck, P. S. (1979). *Stochastic Models, Estimation and Control*, vol. 1, New York: Academic Press.

Maybeck, P. S. (1982). *Stochastic Models, Estimation and Control*, vol. 3, New York: Academic Press.

Mehra, R. K. and Peschon, J. (1971). An Innovations Approach to Fault Detection and Diagnosis in Dynamic Systems, *Automatica* 7: 637–640.

Mosca, E. (1995). *Optimal, Predictive, and Adaptive Control*, Englewood Cliffs, New Jersey: Prentice Hall.

Murgu, A. (2002). Multiclass Service Dynamics Replication Using Evolutionary Selection Games, in *Proceedings of the* $10^{th}$ *International Symposium on Dynamic Games and Applications (ISDG-2002), St. Petersburg, Russia, 8–11 July 2002*, pp. 652–656.

Page, E. S. (1954). Continuous Inspection Schemes, *Biometrica* 41: 100–114.

Roussas, G. G. (1972). *Contiguity of Probability Measures. Some Applications in Statistics*, New York: Cambridge University Press.

Segen, J. and Sanderson, A. C. (1980). Detecting Changes in a Time Series, *IEEE Transactions on Information Theory* IT-26(2): 249–255.

Semoushin, I. V. (1979). Optimality Testing for the Adaptive Kalman Filter by the Use of a Realization of a Scalar Process, *Proceedings of National Academy of Sciences U.S.S.R. – Technicheskaya Cybernetika* 6: 195–198.

Semoushin, I. V. and Polosenko, V. P. (2002). Iterative Cooperative CAD for Avionics Navigation Systems, in *Advances in Concurrent Engineering, Proceedings of the* $9^{th}$ *International Conference on Concurrent Engineering, Cranfield, United Kingdom, 27–31 July 2002*, R. Jardim-Gonçalves, R. Roy, and A. Steigler-Garção, eds., pp. 753–761, Lisse, The Netherlands: A. A. Balkema.

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactures Product*, Toronto: Van Nostrand.

Siegmund, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*, New York: Springer-Verlag.

Van-Trees, H. I. (1968). *Detection, Estimation, and Modulation Theory. Part I: Detection, Estimation, and Linear Modulation Theory*, New York: John Wiley & Sons.

Yoshihawa, K. (1992). *Weakly Dependent Stochastic Sequences and Their Application*, vol. 1.4. Tokyo: World Scientific.