# Jointly Performed Computational Tasks in the Multi-mode System Identification⋆

Innokenti Semoushin

Ulyanovsk State University, 42 Leo Tolstoy Str., 432970 Ulyanovsk, Russia
SemoushinIV@ulsu.ru   http://staff.ulsu.ru/semoushin/

**Abstract.** In this paper, we propose a new closed-loop multi-mode system identification setup that allows the jointly performed change point detection, fault diagnosis, supervised training of a bank of signature generators, unsupervised identification of the optimal filter, and monitored system accommodation on the basis of incomplete noisy data. We validate the setup by simulating numerical examples.

## 1   Introduction

In the literature on modern control systems, more emphasis is placed on detection [1] and identification [2] as separate problems. At the same time, reality provides evidence that in many cases fault detection-diagnosis must be performed together with system identification and accommodation. In adaptive multi-mode or hybrid stochastic systems with an uncertain mode switching mechanism, identification is to be launched repeatedly every time when a mode switch is detected. In the equivalent manner, identification must be stopped when the identification algorithm provides the highest (or tolerable) pitch of agreement with the observed data. Thus, FDD-SIA is a composite challenging problem needing not only a theoretical but also computational research.

In this work, we propose a unified approach to perform FDD-SIA computational tasks simultaneously in a single system. The outline of the paper is as follows: Section 2 describes the monitored system (MS). In Sect. 3 we formulate the Statistical Orthogonality Principle (SOP) and refer to the Accessible Indirect Performance Index (AIPI). The SOP-based FDD is described in Sect. 5. Section 6 shows one of identification algorithms (IA) intended to identify the optimal Steady-State Kalman Filter (SSKF) as a substitute for the Feedback Filter (FF). The MS accommodation is briefly described in Sect. 7. Some experimental results are shown in Sect. 8. Finally, Sect. 9 concludes the paper.
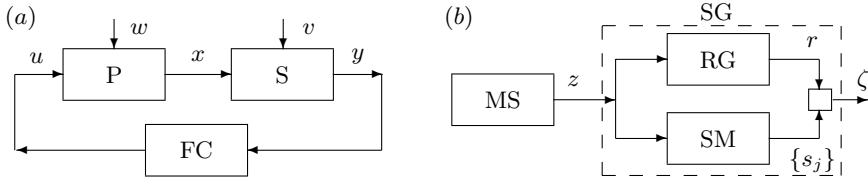
## 2   Monitored System

For the monitored system we consider an errors-in-variables (EIV) control system with the observed data $z(t_i) = \begin{bmatrix} y(t_i) \\ u(t_{i-1}) \end{bmatrix}$, $i = 1, 2, \ldots$, composed of the control

---

input $u \in \mathbb{R}^q$ and the measurement output $y \in \mathbb{R}^m$. The system is parameterized by an uncertainty vector $\theta \in \mathbb{R}^l$ and consists of a plant (P), a sensor (S) and a stabilizing feedback controller(FC) (Fig.1($a$)). The P state equation is



**Fig. 1.** ($a$) Monitored System (MS), and ($b$) Statistical orthogonality. Legend: P – plant, S – sensor, FC – feedback controller, RG – residual generator, SM – sensitivity model, SG – signature generator

$$x(t_{i+1}) = \Phi_\theta x(t_i) + \Psi_\theta u(t_i) + w(t_i), \quad i = 0, 1, \ldots \quad (1)$$

with $x \in \mathbb{R}^n$ and a noise $w(t_i)$ where the random initial $x(t_0)$ at some $t_0$ has a finite mean $\bar{x}_0$ and a finite covariance $P_0 \geq 0$ and $\{u(t_i)\}$ is wide-sense stationary and so $\mathrm{E}\{||u(t_i)||^2\} < \infty$ for all $t_i$. The S equation is

$$y(t_i) = Hx(t_i) + v(t_i), \quad i = 1, 2, \ldots \quad (2)$$

with $\mathrm{rank}(H) < n$ and a noise $v(t_i)$. Both $\{w(t_i)\}$ and $\{v(t_i)\}$ are i.i.d. zero mean mutually independent wide-sense stationary sequences whose covariances are $\mathrm{E}\{w(t_i)w(t_i)^T\} = Q_\theta \geq 0$ and $\mathrm{E}\{v(t_i)v(t_i)^T\} = R_\theta > 0$ for all $t_i$. The FC, in compliance with the well-known separation property, is formed by a feedback filter (FF) cascaded with a feedback regulator (FR). The FF equations are

$$\tilde{x}(t_{i+1}^-) = \bar{\Phi}\tilde{x}(t_i^+) + \Psi u(t_i), \quad \tilde{x}(t_i^+) = \tilde{x}(t_i^-) + \bar{K}[y(t_i) - H\tilde{x}(t_i^-)] \quad (3)$$

with some initials $\tilde{x}(t_0^+)$, $u(t_0)$. Constant transition matrix $\bar{\Phi}$ and constant gain $\bar{K}$ are chosen in a sense arbitrarily or designed to be optimal in a steady state (as $t_0 \to -\infty$) w.r.t. a *fixed* (fault-free) system mode specified by a nominal value $\theta_0$ of $\theta$. The FR equation is

$$u(t_i) = f_{\mathrm{FR}}[\tilde{x}(t_i^+)] = -\bar{G}_{\mathrm{FR}}\tilde{x}(t_i^+) \quad (4)$$

with a function $f_{\mathrm{FR}}[\cdot]$ which can be chosen to satisfy the second equality in (4) with a constant gain matrix $\bar{G}_{\mathrm{FR}}$. By (3), (4), it is predetermined that $u(t_i)$ depends on all available data $z(t_1^i) = [z(t_1), \cdots, z(t_i)]$, the latter is a notation for a composite (stackable) column vector. Note that $\begin{bmatrix} w(t_i) \\ v(t_j) \end{bmatrix}$ is independent of $u(t_k)$ and $x(t_k)$ for all $t_i \geq t_k$, $t_j > t_k$.

The overall closed-loop control system (1), (2), (3), (4) is assumed to be asymptotically stable in all modes of operating referred to by the subscript $\theta$,

and so all the processes within the system are wide-sense stationary at every $t_i$ as $t_0 \to -\infty$ (*the main assumption*). The system is designed to operate with a minimum expected control cost

$$J_c = \lim_{j \to -\infty} \mathrm{E}\{x^T(t_{N+1})X_f x(t_{N+1}) + \sum_{i=j}^{N}[x^T(t_i)Xx(t_i) + u^T(t_i)Uu(t_i)]\} \quad (5)$$

for $N > 0$, $X_f \geq 0$, $U > 0$, $X \geq 0$, and in all modes holds the *main properties*:

* $(\Phi_\theta, Q_\theta^{1/2})$ is stabilizable, $(\Phi_\theta, X^{1/2})$ is detectable, $(\Phi_\theta, H)$ is completely observable, and $(\Phi_\theta, \Psi_\theta)$ is completely controllable.

to guarantee the existence of the optimal steady-state parameters for (3), (4).

The system itself can operate in several (*finitely* or *infinitely* many) modes. This brings the problem to the realm of hybrid or multi-mode systems. The mode characterized by $\theta_0$ can be treated as the main one. The other modes characterized by some $\theta_1, \theta_2, \ldots, \theta_M$ – for the case of *finitely* many modes, – can be viewed as some alternative modes not obligatory (albeit possibly) faulty. In this case we have to test the $M$ pairs of hypotheses, in each pair between

$$\mathrm{H}_0 = \{\theta^\dagger = \theta_0\} \quad \text{and} \quad \mathrm{H}_\mu = \{\theta^\dagger = \theta_\mu = \theta_0 + \Upsilon_\mu\} \quad (6)$$

where $\theta^\dagger$ is the *true* value of $\theta$, $\mu = 1, 2, \ldots, M$, and $\Upsilon_\mu$ is the $\mu$-th change on $\theta$ (the $\mu$-th alternative mode of operation). Allowing for arbitrary faults, which are modelled by any sizable change $\Upsilon$ on $\theta$ not violating the above main assumption, we come to the case of *infinitely* many modes. By this, we generalize the problem as we have to test the continuum set of pairs of hypotheses

$$\mathrm{H}_0 = \{\theta^\dagger = \theta_0\} \quad \text{and} \quad \mathrm{H}_\Upsilon = \{\theta^\dagger = \theta_\Upsilon = \theta_0 + \Upsilon\} \ . \quad (7)$$

In both cases (6) and (7) we do not assume a specific mode switching behavior (for example, a Markovian one) and view it as deterministic (albeit unknown to the observer) like controlled by an independent actor.

## 3   Statistical Orthogonality Principle

**Definition 1.** *Given $l$ pairs $\{r, s_j\}$, $j = 1, 2, \ldots, L$, each formed by column vectors $r = r(t_i)$ and $s_j = s_j(t_k)$. Then $r(t_i)$ and $s_j(t_k)$ are said to be* statistically orthogonal *at distance $\Delta t = t_k - t_i$, that is denoted as $r(t_i) \perp s_j(t_k)$, if*

$$\forall t_i, t_k, \Delta t = t_k - t_i = \text{const}: \quad \mathrm{E}\{s_j^T(t_k)r(t_i)\} = 0 \quad (8)$$

*and $r(t_i)$ is said to be statistically orthogonal to $S(t_k) = [s_1(t_k) \,|\, \cdots \,|\, s_L(t_k)]$ at $\Delta t = t_k - t_i$ if (8) holds for all given pairs, that is denoted as $r(t_i) \perp S(t_k)$.*

Let $r$ and $S$ in Definition 1 be computed by $L$ pairs of parallel blocks RG and SM as shown in Fig.1($b$). By RG and SM are meant *Residual Generator* and *Sensitivity Model* both parameterized by a design parameter $\hat{\theta} \in \mathbb{R}^L$ not fully but partly being an exact replica of $\theta$, the MS parameter, in its elements' numbering and sense. Thus, $r(t_i) = r(t_i, \hat{\theta})$ and $S(t_k) = S(t_k, \hat{\theta})$.

**Definition 2.** *Let* RG *and* SM *be designed in such a way that* in case (6) *for each* $\mu = 0, 1, 2, \ldots, \mathcal{M}$, *parameter* $\hat{\theta}$ *takes a specified value* $\hat{\theta}_\mu$ *such that*

$$\forall \mu : \quad \{\hat{\theta} = \hat{\theta}_\mu\} \equiv \{ r(t_i, \hat{\theta}) \perp S(t_i, \hat{\theta}) \; \leftrightarrows \; \theta^\dagger = \theta_\mu \} \tag{9}$$

*or* in case (7) *for each* $\Upsilon$, *parameter* $\hat{\theta}$ *takes a specified value* $\hat{\theta}_\Upsilon$ *such that*

$$\forall \Upsilon : \quad \{\hat{\theta} = \hat{\theta}_\Upsilon\} \equiv \{ r(t_i, \hat{\theta}) \perp S(t_i, \hat{\theta}) \; \leftrightarrows \; \theta^\dagger = \theta_\Upsilon \} \tag{10}$$

*where* $\leftrightarrows$ *stands for "iff" (if and only if). Then (9) and (10) is termed* Statistical Orthogonality Principle.

The parallel blocks in Fig.1(*b*), RG and SM, may be variously designed to satisfy Definitions 1 and 2. However as system (1), (2), (3), (4) is stochastic, the presence of an *(adaptive) estimator* (AE), denote it by $g(t_{i|i-1}, \hat{\theta})$, is requisite for obtaining a residual $r(t_i, \hat{\theta})$. With it, writing $r(t_i, \hat{\theta}) \perp S(t_i, \hat{\theta})$ is straightforward as a necessary (and on occasion, sufficient) condition for minimum of

$$J(\hat{\theta}) = 1/2 \lim_{t_j \to -\infty} \mathrm{E}\{\|r(t_i, \hat{\theta})\|^2 \mid y(t_j^{i-1})\} \;. \tag{11}$$

Thus, SOP implementation boils down to the construction of (11) so as to obtain

$$S(t_i, \hat{\theta}) = \left[ \left. \frac{\partial r(t_i, \hat{\theta})}{\partial \hat{\theta}_1} \right| \cdots \left| \frac{\partial r(t_i, \hat{\theta})}{\partial \hat{\theta}_L} \right. \right] \;. \tag{12}$$

## 4   Identification Performance Indices

Minimum Prediction Error methods [3] offer to predict by $g(t_{i|i-1}, \hat{\theta})$ the system output $y(t_i)$ and use $r(t_i, \hat{\theta}) = y(t_i) - g(t_{i|i-1}, \hat{\theta})$ as a residual. Minimizing (11) with such residual leads to the biased parameter estimates for the given EIV system [4], [5]. Had $r(t_i, \hat{\theta})$ in (11) instead been in one of the following forms

$$r(t_i, \hat{\theta}) = \{ x(t_i) - g(t_{i|i-1}, \hat{\theta}) \text{ or } \hat{x}_{\mathrm{OPT}}(t_{i|i-1}) - g(t_{i|i-1}, \hat{\theta}) \} \tag{13}$$

where $\hat{x}_{\mathrm{OPT}}(t_{i|i-1})$ is the one-step SSKF predictor for $x(t_i)$, minimizing (11) would have helped to avoid the bias. Obviously, (13) provides two clear examples of *Direct* (but) *Inaccessible Residual* (DIR) and in this case (11) is an example of *Direct* (but) *Inaccessible Performance Index* (DIPI).

*Remark 1.* In the limit as $t_0 \to -\infty$, FF equations (3) assure coincidence

$$\tilde{x}(t_i^+) = \hat{x}_{\mathrm{OPT}}(t_{i|i}) \quad \text{and} \quad \tilde{x}(t_i^-) = \hat{x}_{\mathrm{OPT}}(t_{i|i-1}) \tag{14}$$

if optimally designed for the fault-free system mode specified by $\theta^\dagger = \theta_0$.

*Remark 2.* If a set of AEs $g(t_{i|i-1}, \hat{\theta})$ in (13) contains SSKF, then $g(t_{i|i-1}, \hat{\theta})$ minimizing (11) coincides with $\hat{x}_{\mathrm{OPT}}(t_{i|i-1})$. If a fault occurs, we consider SSKF a *hidden* object to be identified to substitute FF and be used to modify FR.

*Problem 1.* To implement both SOP-based FDD and SSKF identification, we need to construct the *Accessible* albeit *Indirect Residual* (AIR) as a substitute for the DIR in order to change from the DIPI to the *Accessible* albeit *Indirect Performance Index* (AIPI) in (11), the latter having the same minimizing estimator as the DIPI.

*Solution 1.* Being the system completely observable enables the solution to be found in the form:

$$\text{AIR}(t_i, \hat{\theta}) = \text{DIR}(t_i, \hat{\theta}) + \text{SP}(t_i), \quad \text{SP}(t_i) \perp \text{DIR}(t_i, \hat{\theta}), \ \text{SP}(t_i) \neq f(\hat{\theta}) \quad (15)$$

$$\text{AIPI}(\hat{\theta}) = \text{DIPI}(\hat{\theta}) + \text{Const}, \quad \text{Const} = 1/2 \, \text{E}\{\|\text{SP}(t_i)\|^2 \mid y(t_j^{i-1})\} \quad (16)$$

where $\text{SP}(t_i)$ is a *slack process* fully independent of $\hat{\theta}$ $[\text{SP}(t_i) \neq f(\hat{\theta})]$ and (16) is the half mean squared Euclidean norm of (15) in the limit as $t_j \to -\infty$.

*Remark 3.* The solution is given in [6], [7], [8] and thoroughly tested in [9] for identification purposes only. In what follows, we denote the $\text{AIR}(t_i, \hat{\theta})$ by $\varepsilon(t_i, \hat{\theta})$ and use it here for the jointly performed FDD-SIA computational tasks. So from now on, $\varepsilon(t_i, \hat{\theta})$ serves instead of $r(t_i, \hat{\theta})$ in Formulas (8) to (12).

## 5    Fault Detection and Diagnosis

In accordance with Definition 2 and Remark 3, being the $k$-th expectation

$$\zeta_j(t_i, \hat{\theta}) = \text{E}\{\left( \frac{\partial \varepsilon(t_i, \hat{\theta})}{\partial \hat{\theta}_j}, \varepsilon(t_i, \hat{\theta}) \right) \mid \hat{\theta} = \hat{\theta}_\mu \text{ or } \hat{\theta}_\Upsilon\}, \quad j = 1, 2, \dots, L \quad (17)$$

of the dot product $(\cdot, \cdot)$ in the immediate vicinity of zero is equipollent to the fact that $\theta^\dagger = \theta_\mu$ or $\theta_\Upsilon$. On this basis we use it as a signature of the fact as it is shown schematically in Fig. 1(b) where $\zeta$ is composed of (17), $\zeta \in \mathbb{R}^L$.

In FDD, when we test between (6), the theoretical signature

$$\forall \mu \text{ and } j = 1, 2, \dots, L: \quad \zeta_j(t_i, \hat{\theta}_\mu) = 0 \quad (18)$$

is checked by the following approximate decision rule:

$$\forall \mu = 1, 2, \dots, \mathcal{M} \text{ and } j = 1, 2, \dots, L: \quad \left| 0.5 - \frac{n_j^\mu}{N} \right| \begin{matrix} \text{H}_\mu \\ \gtrless \\ \text{H}_0 \end{matrix} \gamma_j \quad (19)$$

where $n_j^\mu$ is the number of negative outliers of the process $\hat{\zeta}_j(t_i, \hat{\theta}_\mu)$ in the current sample of size $N$. Here $\hat{\zeta}_j(t_i, \hat{\theta}_\mu)$ is an estimator (for which we recommend the exponential smoothing) of $\zeta_j(t_i, \hat{\theta}_\mu)$ defined for (18) according to (17) for each $\mu$-th *Signature Generator* (SG) within the *Bank of Signature Generators* (BSG), $\mu = 1, 2, \dots, \mathcal{M}$ (Fig. 2(a)). Signature Evaluator and Selector (SES) in this setup operates according to (19), where the sample size $N$ and thresholds $\gamma_j$ should

be determined experimentally to guarantee the maximum decision power with the limited error probability. If several hypotheses among $H_\mu$, $\mu = 1, 2, \ldots, \mathcal{M}$, have been selected, we discriminate between them by the additional rule: select the $H_{\mu_*}$ with

$$\mu_* = \arg\min_\mu \sum_{j=1}^{L} \left| 0.5 - \frac{n_j^\mu}{N} \right| . \tag{20}$$

If none $H_\mu$, $\mu = 1, 2, \ldots, \mathcal{M}$, has been selected, we adopt a decision that $H_0$ is true.
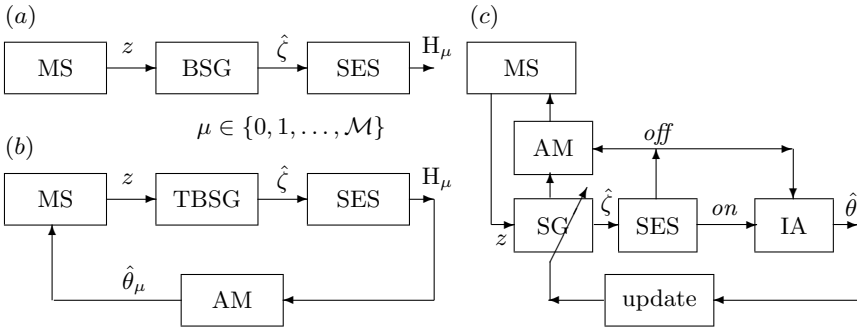


**Fig. 2.** (*a*) Fault detection and diagnosis (discrimination between $\mathcal{M} + 1$ hypotheses). (*b*) Fault diagnosis and system accommodation using TBSG. (*c*) Fault point detection to switch IA and AM. Legend: MS – monitored system, BSG – Bank of Signature Generators, TBSG – Trained BSG, SES – Signature Evaluator and Selector, AM – Accommodation Mechanism, IA – Identification Algorithm

## 6   Optimal Filter Identification and BSG Training

To reason what change (or what fault) has occurred, one of two strategies may be used: parallel or sequential. The first one is adequate for Case (6), while the latter for Case (7). The parallel strategy implies the existence of the BSG as discussed in Sect. 5. Each SG can be trained beforehand to match the corresponding system mode, i.e. to satisfy (18). Training algorithms for a single ($\mu$-th) SG are standard identification algorithms for $\hat{\theta} = \hat{\theta}_\mu$. For example, one can use the following SLS:

*Simplified Least Squares.* For $k = 1, 2, \ldots$ and $\Lambda = \text{diag}[\lambda^{(j)}]$, $j = 1, 2, \ldots, L$, with $\Lambda_1 = I$ do (21) and then (22) where $k$ is the step number:

$$\lambda_{k+1}^{(j)} = \lambda_k^{(j)} + \left\| \frac{\partial \varepsilon(t_i, \hat{\theta}_k)}{\partial \hat{\theta}_j} \right\|^2 \tag{21}$$

$$\hat{\theta}_{k+1} = \hat{\theta}_k - \Lambda_{k+1}^{-1} \mathcal{E}\{ S^T(t_i, \hat{\theta}_k) \varepsilon(t_i, \hat{\theta}_k) \} \tag{22}$$

where $\mathcal{E}\{\cdot\}$ is a smoothing algorithm. In the case of exponential smoothing we obtain $\hat{\zeta}(t_i, \hat{\theta}) = \mathcal{E}\{\zeta(t_i, \hat{\theta})\}$ from the signature vector $\zeta(t_i, \hat{\theta}) = S^T(t_i, \hat{\theta})\varepsilon(t_i, \hat{\theta})$ by

$$\hat{\zeta}(t_i, \hat{\theta}) = \beta\hat{\zeta}(t_{i-1}, \hat{\theta}) + (1 - \beta)\zeta(t_i, \hat{\theta}), \qquad 0 \leq \beta < 1 . \tag{23}$$

In Case (7) we have to use the sequential strategy of Fig. 2(c) with one SG, one SES and one identification algorithm (IA). In this setup, SG is intended to detect and track all possible abrupt changes as they occur in time in the monitored system. The SG performs this by means of IA, which may be of (21)–(22) type or another. The SES tests between hypotheses (7) so that the rule (19) is implemented with $\mu = 1$. Decision to adopt $H_1$ starts up the IA; decision to adopt $H_0$ stops it and starts the accommodation mechanism (AM) in order to bring the feedback of MS and the real situation after change into better agreement and, in doing so, to adapt the system feedback to suit the new (after-change) conditions. Figure 3 shows how it is implemented except AM. Accommodation mechanism is discussed briefly in Sect. 7.

```
begin i := 0;
      for j := 1 to L do
            begin n_j := 0; s_j :=off; λ_j := 1 end;
      while i ≤ imax do
            for j := 1 to l do
                  begin i := i + 1;
                        if ζ̂_j(t_i, θ̂) < 0 then n_j := n_j + 1;
                        f_j := 0.5 - n_j/N;
                        if i mod N = 0 then
                              begin n_j := 0;
                                    if abs(f_j) < γ_j then
                                          s_j :=off else s_j := on
                              end;
                        if s_j = on then
                              begin
                                    λ^(j) := λ^(j) + ‖∂ε(t_i,θ̂)/∂θ̂_j‖²;
                                    θ̂^(j) := θ̂^(j) - ζ̂_j(t_i, θ̂)/λ^(j)
                              end else
                              if i mod N = 0 then λ^(j) := 1
                  end
      end.
```

**Fig. 3.** Pseudocode sketching the inter-independent parameter identification algorithm

## 7   Monitored System Accommodation

To accommodate the MS to the newly developed (maybe, faulty) mode, the accommodation mechanism (AM) is expected to be capable to replace parameters $\bar{\Phi}$ and $\bar{K}$ of (3) with the newly identified matrix $\Phi$ and optimal SSKF gain $K$, and $\bar{G}_{\mathrm{FR}}$ of (4) with the $G_{\mathrm{FR}}$ computed by solving the *Discrete Algebraic Riccati Equation*, DARE, in compliance with the well known LQG control law using matrices of (5).

## 8   Approach Validation

*I. System.* Dimensions are $n = 2$, $m = 1$, $q = 1$, $l = 2$. Matrices are as follows:

$$\Phi_\theta = \begin{bmatrix} 0 & 1 \\ \theta_1 & \theta_2 \end{bmatrix}, \qquad \Psi_\theta = \Psi = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad Q_\theta = Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad R_\theta = R = \begin{bmatrix} 0.1 \end{bmatrix}, \qquad \theta^T = \begin{bmatrix} \theta_1 & \theta_2 \end{bmatrix}$$

$$\theta_0^T = \begin{bmatrix} 0.30 & 0.68 \end{bmatrix}, \quad \theta_\Upsilon^T = \begin{bmatrix} 0.40 & 0.10 \end{bmatrix}, \quad \Upsilon^T = \begin{bmatrix} 0.10 & -0.58 \end{bmatrix}$$

$$\bar{\Phi} = \begin{bmatrix} 0 & 1 \\ 0.2 & 0.2 \end{bmatrix}, \qquad \bar{K} = \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}, \qquad \bar{G}_{\mathrm{FR}} = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix} \ .$$

*II. Residual Generator.* According to Remark 3, RG is composed of AE and AIR:

$$g(t_{i+1|i}, \hat{\theta}) = A[g(t_{i|i-1}, \hat{\theta}) + D\eta(t_i, \hat{\theta})] + \Psi u(t_i) \tag{24}$$

$$\eta(t_i, \hat{\theta}) = y(t_i) - Hg(t_{i|i-1}, \hat{\theta}) \ . \tag{25}$$

For AE (24), (25), we have here $L = 4$ and

$$A = \begin{bmatrix} 0 & 1 \\ a_1 & a_2 \end{bmatrix}, \qquad D = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \qquad \hat{\theta}^T = \begin{bmatrix} a_1 & a_2 & d_1 & d_2 \end{bmatrix} \ . \tag{26}$$

With reference to Remark 3, we construct AIR as

$$\varepsilon(t_i, \hat{\theta}) = \mathcal{N}(D)\eta(t_{i-s+1}^i, \hat{\theta}), \quad \mathcal{N}(D) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \ddots & 1 & 0 & 0 \\ d_{n-1} & \ddots & 1 & 0 \\ d_n & d_{n-1} & \ddots & 1 \end{pmatrix} \ . \tag{27}$$

In the first formula of (27), $s$ is the greatest partial observability index of the system and the second formula holds if MS with $m = 1$ is taken in the standard observable form [7]. Such is here the case and $s = 2$,

$$\eta(t_{i-s+1}^i, \hat{\theta}) = [\eta(t_{i-1}, \hat{\theta}) \quad \eta(t_i, \hat{\theta})]^T \ . \tag{28}$$

*III. Sensitivity Model.* To obtain (12) from $\varepsilon(t_i, \hat{\theta})$, (27), in view of Remark 3, we have to generate the sensitivity functions having defined them by

$$\mu^{(j)}(t_{i+1|i}, \hat{\theta}) = \frac{\partial}{\partial \hat{\theta}_j} g(t_{i+1|i}, \hat{\theta}), \quad j = 1, \ldots, L \; . \tag{29}$$

*IV. Pattern for Computational Experiments.* The composite algorithm we test is outlined by Figs. 2(c) and 3. Besides the MS with the above matrices, it involves: RG – (24), (25), (26), (27) and (28); SM – (12) and (29); SG – (17); SES – (19) and (20); and IA – (21), (22) and (23). The algorithm is started up with the non-optimal parameters $\bar{\Phi}$, $\bar{K}$ and $\bar{G}_{\mathrm{FR}}$. The SES is expected to detect this initial non-optimality and to start up the IA precisely in respect to the parameters which require optimisation. The IA is expected to eliminate the initial non-optimality (this is the SG training). After that, at the middle time instant $t_{8000}$, the MS switches from mode $\mathrm{H}_0$ to $\mathrm{H}_\Upsilon$, cf. (7). The algorithm is expected to detect this mode switch, to identify the new optimal parameters and accommodate the MS feedback to the after-switch conditions.

*V. Experimental Results.* An example of the results is shown in Figs. 4 and 5. This is one of many results obtained for many different conditions [10]. In these experiments, the following values were chosen: $\gamma_j = 0.05$, $N = 2000$ and $\beta = 0.1$. The estimates formed by the composite algorithm are thus seen to be satisfactory enough. This lends support to the validity of the proposed approach.
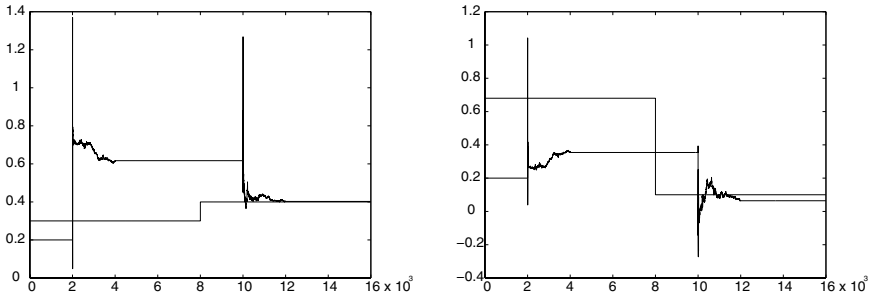


**Fig. 4.** Tracking the true parameters $\theta_1$ and $\theta_2$ by $a_1$ (*left*) and $a_2$ (*right*) of (26)

# 9   Conclusions and Future Work

We have shown that statistical orthogonality principle and indirect identification performance index provide the basis for the jointly performed computational tasks in the multi-mode stochastic system identification. Further investigations of this approach should be made to study all the aspects, both theoretical and
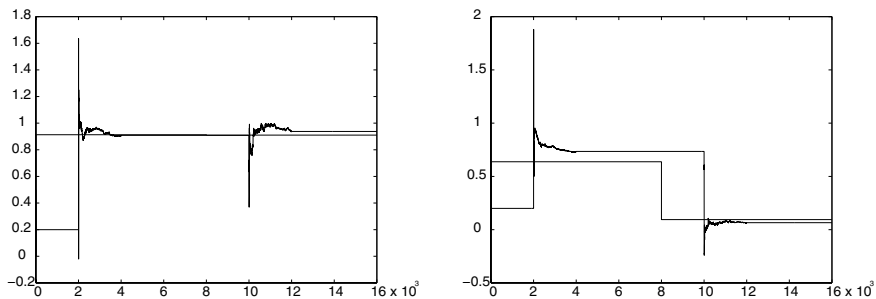
**Fig. 5.** Tracking the SSKF parameters $k_1$ and $k_2$ by $d_1$ (*left*) and $d_2$ (*right*) of (26)

applied, for instance – to evaluate it in terms of false alarm and missed detection rates.

## References

1. Basseville, M., Nikiforov, I.: Detection of Abrupt Changes: Theory and Applications. Prentice-Hall Inc., New York (1993)
2. Landau, I.D. (ed.): Identification des Systemes. Les Bases de l'Identification des Systemes. Hermes, Paris (2001)
3. Caines, P.: Linear Stochastic Systems. John Willey & Sons, New York Chichester Brisbane Toronto Singapore (1988)
4. Soderstrom, T.: Identification of stochastic linear systems in presence of input noise. Automatica **17** (1981) 713–725
5. Ansay, P., Gevers, M., Wertz, V.: Closed-loop or open-loop models in identification for control? In: Proc. of the European Control Conference (1999) CD-ROM F544
6. Semoushin, I.V.: Adaptive Identification and Fault Detection Methods in Random Signal Processing. Saratov University Publishers, Saratov (1985) [*in Russian*]
7. Semoushin, I.V.: Adaptive control for stochastic linear plants under conditions of uncertainty. Nonlinear Dynamical Systems - Qualitative Analysis and Control: Collected Papers, The Institute for System Analysis of the Russian Academy of Sciences **2** (1994) 104–110
8. Semoushin, I.V., Tsyganova, J.V.: Indirect error control for adaptive filtering. In: Neittaanmaki, P., Tiihonen, T., Tarvainen, P. (eds.): Proc. of the 3rd European Conference on Numerical Mathematics and Advanced Applications. World Scientific, Singapore New Jersey London Hong Kong (2000) 333–340
9. Semoushin, I., Gorokhov, O.: Computational processes in iterative control design. In: Sloot, P.M.A., Kenneth Tan, C.J., Dongarra, J.J., Hoekstra, A.G. (eds.): Computational Science – ICCS 2002. Lecture Notes in Computer Science, Vol. 2329. Springer-Verlag, Berlin Heidelberg New York Barcelona Hong Kong London Milan Paris Tokyo (2002) 186–195
10. Kondratiev, A.E., Fatyanova, O.A.: Start-stop algorithms for adaptive filtering. In: Boulyarski, S.V. (ed.): Young Researchers' Collected Papers of the Ulyanovsk State University. UlSU Publishers, Ulyanovsk (2001) 6–7