Proceedings of the 40th IEEE
Conference on Decision and Control
Orlando, Florida USA, December 2001

Change detection in the dynamics with recursive subspace identification*

H. Oku, G. Nijsse, M. Verhaegen and V. Verdult
University of Twente, The Netherlands,
{Oku, G.Nijsse, M.Verhaegen,
V.Verdult}@tn.utwente.nl

Abstract

In this paper, we propose a tracking mechanism to follow time-variations in the dynamics of a linear system with a recursive identification of the PI-MOESP scheme. The proposed mechanism consists of change detection scheme followed by a re-initialization of the recursive calculations. The change detection is based on the least-squares interpretation of the calculation in the subspace scheme and detects whether the estimates of the recursive solution without exponential forgetting lies in the confidence interval of the estimates obtained with a second finite-window length solution to the least-squares problem. When a change has been detected, the estimate by the recursive implementation is re-initialized via the solution of a constrained least-squares problem. One numerical example is presented to illustrate that our change detection and re-initialization scheme can detect incipient changes in the system dynamics without detecting changes in input dynamics.

1 Introduction

There is a vast amount of literature related to change detection and related to forgetting mechanism in (recursive) system identification schemes. The autoregressive-type model structure has often been used as models of systems under surveillance because there exist many kind of on-line parameter estimation algorithms, for example, RLS, RLMS and so on, which estimate parameters of such models recursively. Moreover, since it has a close connection with the probability theory, it is not difficult to introduce hypothesis tests using the statistics, e.g., mean and covariance values of estimates.

Suppose that we consider segmentation of data sampled from a practical plant at work as preprocessing for system identification. One possibility is to apply a change detection method to do data segmentation that searches for segments in the data batches during which the system dynamics are linear time-invariant (LTI) irrespective of whether the input properties change.

An interesting candidate for system identification is one of the existing subspace state-space system identification (4SID) methods. Recently, recursive algorithms of 4SID identification have been proposed [2, 9, 8, 12].

*Supported by the Dutch Technology Foundation (STW) under project number DEL55.3891.

It is expected that parameters estimated by recursive 4SID algorithms allow us to observe changes in the system dynamics without being disturbed by changes in the input signals. This is because the parameters are obtained based on the projection onto the orthogonal complement of a subspace defined by a given input sequence.

Our contribution is twofold. The first is to introduce a change detection mechanism in a recursive implementation of the PI-MOESP [11] scheme. In order to detect incipient changes as well as abrupt ones, our change detection scheme estimates a relevant system quantity by two 4SID algorithms *in parallel* and compares them using a statistical measure. One 4SID algorithm is to provide an accurate estimate recursively. The other can track a varying parameter; however it has low accuracy. The second contribution is to develop how to re-initialize the recursive accurate 4SID scheme once a change is detected. We apply an idea of a constrained least-squares problem proposed by Kulhavý and Zarrop [6] to our re-initialization procedure. The procedure is invoked only when a change is detected. Note that, potentially, our idea of the re-initialization can be applied to adaptive filters based on parallel parameter estimation.

In section 2, we introduce the notation used throughout the paper. Also an output error model is presented here as well as several assumptions related to the feasibility of PI-MOESP. In section 3, we outline the relationship between 4SID and a linear least squares problem in [4]. Section 4 describes the two contributions of this paper. A numerical example illustrates that our scheme can detect changes of a system properly without detecting input changes in section 5.

2 Notation

We consider the following output error model:

$$x_{k+1} = Ax_k + Bu_k, \tag{1a}$$

$$y_k = Cx_k + Du_k + v_k, \tag{1b}$$

where $y_k, v_k \in \mathbb{R}^l$, $u_k \in \mathbb{R}^m$ and $x_k \in \mathbb{R}^n$. The system matrix A is asymptotically stable and the pair (C, A) is observable. The output error v_k is a zero mean white noise sequence that is uncorrelated with u_k and has covariance matrix

$$\mathbf{E}\, v_i v_j^T = \begin{cases} R > 0, & i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Choose an integer s such that $N > s \geq n + 1$, where n is the dimension of the state vector. Let a sequence of input and output data $\{(u_j,\ y_j)\}$, $j = -s+1, \cdots, 0, 1, \cdots, N, \cdots$ be given. define the vector $u_{s(i)} \in \mathbb{R}^{ms}$ as follows:

$$u_{s(i)} := \begin{bmatrix} u_{i-s+1}^T & \cdots & u_i^T \end{bmatrix}^T. \tag{3}$$

The input Hankel matrix is defined as

$$U_N := \left[\begin{array}{ccc} \sqrt{\gamma}^{N-s} u_{s(s)} & \cdots & u_{s(N)} \end{array} \right], \quad (4)$$

where γ is a forgetting factor satisfying $\gamma \in \mathbb{R}$ and $\gamma \leq 1$. Similarly, the vector $y_{s(i)} \in \mathbb{R}^{ls}$ and the output Hankel matrix Y_N are defined as $u_{s(i)}$ and U_N from the sequence $\{y_j\}$, respectively. A "shifted" version of the input Hankel matrix is defined as

$$U_{P,N} := \left[\begin{array}{ccc} \sqrt{\gamma}^{N-s} u_{s(0)} & \cdots & u_{s(N-s)} \end{array} \right], \quad (5)$$

where the extra subscript "P" indicates the *past* with respect to the data in U_N. For the noise, $v_{s(i)}$, V_N and $V_{P,N}$ are defined similar to (3), (4) and (5), respectively. The matrix which consists of the state vectors at succeeding time instants is defined as

$$X_{N-s+1} := \left[\begin{array}{ccc} \sqrt{\gamma}^{N-s} x_1 & \cdots & x_{N-s+1} \end{array} \right]. \quad (6)$$

We define the following matrix which consists of the state vectors:

$$X_{P,N-s+1} := \left[\begin{array}{ccc} \sqrt{\gamma}^{N-s} x_{-s+1} & \cdots & x_{N-2s+1} \end{array} \right]. \quad (7)$$

We introduce the following projection matrices with respect to the input Hankel matrix U_N:

$$\Pi_{U_N} := U_N^T \left(U_N U_N^T\right)^{-1} U_N, \quad \Pi_{U_N}^{\perp} := I - \Pi_{U_N}. \quad (8)$$

Following the reference [5], the input is assumed to be persistently exciting, such that:

$$\operatorname{rank} \left(\lim_{N\to\infty} \frac{1}{N} \left[\begin{array}{c} X_{N-s+1} \\ U_N \end{array} \right] \left[\begin{array}{cc} U_{P,N}^T & U_N^T \end{array} \right] \right) = n+sm. \quad (9)$$

We assume that there exists an integer N such that

$$\frac{1}{N-s} \left[\begin{array}{c} U_{P,N-1} \\ U_{N-1} \end{array} \right] \left[\begin{array}{cc} U_{P,N-1}^T & U_{N-1}^T \end{array} \right] > 0. \quad (10)$$

For the case of $\gamma < 1$, we assume

$$\frac{1-\gamma}{1-\gamma^{N-s}} \left[\begin{array}{c} U_{P,N-1} \\ U_{N-1} \end{array} \right] \left[\begin{array}{cc} U_{P,N-1}^T & U_{N-1}^T \end{array} \right] > 0. \quad (11)$$

3 4SID as a linear least squares problem

3.1 The least squares problem

Jansson and Wahlberg [4] have introduced a cost function with a matrix linear regression model to interpret 4SID schemes. Consider the output error model (1). With the definition of the matrix $\mathcal{L} := \left[\begin{array}{ccc} A^{s-1}B, & \cdots & AB, \quad B \end{array} \right]$, we can denote the state vector x_k as

$$x_k = \mathcal{L} u_{s(k-1)} + A^s x_{k-s}, \quad (12)$$

On the other hand, let the extended observability matrix $\mathcal{O}$ and the lower triangular Toeplitz matrix $\mathcal{H}$ be defined, respectively, as

$$\mathcal{O} := \left[\begin{array}{cccc} C^T & (CA)^T & \cdots & (CA^{s-1})^T \end{array} \right]^T,$$

$$\mathcal{H} := \left[\begin{array}{cccc} D & & & 0 \\ CB & D & & \\ \vdots & & \ddots & \\ CA^{s-2}B & CA^{s-3}B & \cdots & D \end{array} \right],$$

then,

$$y_{s(k+s-1)} = \mathcal{O} x_k + \mathcal{H} u_{s(k+s-1)} + v_{s(k+s-1)}, \quad (13)$$

Substituting the expression for x_k given by (12) in the last equation yields:

$$y_{s(k+s-1)} = \mathcal{O}\mathcal{L} u_{s(k-1)} + \mathcal{H} u_{s(k+s-1)} + v_{s(k+s-1)} + \mathcal{O} A^s x_{k-s}, \quad (14)$$

and therefore:

$$Y_N = \left[\begin{array}{c|c} \mathcal{O}\mathcal{L} & \mathcal{H} \end{array} \right] \left[\frac{U_{P,N}}{U_N} \right] + V_N + \mathcal{O} A^s X_{P,N-s+1}, \quad (15)$$

Note that, since A is assumed to be asymptotically stable, the covariance matrix of the state sequence, i.e., $\lim_{N\to\infty} = \frac{1}{N} X_{P,N-s+1} X_{P,N-s+1}^T$ is bounded. Note also that, since the pair (C, A) is observable, the operator $\mathcal{O}$ (for $s \to \infty$) remains bounded. Therefore, let $\Phi_N = U_{P,N}$, the last term in the right hand side of (15) vanishes as s tends to infinity.

Now, we introduce the following matrix regression model:

$$Y_N = \left[\begin{array}{cc} \Omega & \mathcal{H} \end{array} \right] \left[\begin{array}{c} \Phi_N \\ U_N \end{array} \right] + \epsilon_N. \quad (16)$$

The parameter s is selected sufficiently large such that the error matrix ϵ_N approximates the properties of the matrix V_N arbitrarily close. The parameter Ω then can be thought of as the product of the extended observability matrix $\mathcal{O}$ and a kind of the controllability matrix $\mathcal{L}$. Let us denote that the k-th column vector of the modeling error ϵ_N by $\varepsilon_{s(k+s-1)} := \left[\begin{array}{ccc} \varepsilon_{k-s+1}^T & \cdots & \varepsilon_k^T \end{array} \right]^T$, and the k-th column vector of Φ_N by $\phi_{s(k-1)}$. Then, the k-th column vector of (16) can be described as

$$y_{s(k+s-1)} = \left[\begin{array}{cc} \Omega & \mathcal{H} \end{array} \right] \left[\begin{array}{c} \phi_{s(k-1)} \\ u_{s(k+s-1)} \end{array} \right] + \varepsilon_{s(k+s-1)}. \quad (17)$$

In [4] the following cost function has been defined:

$$J'_N(\Omega, \mathcal{H}) = \operatorname{Trace} \epsilon\epsilon^T, \quad (18a)$$

and the related least squares problem:

$$\widehat{\Omega}_N, \widehat{\mathcal{H}}_N = \arg\min J'_N(\Omega, \mathcal{H}). \quad (18b)$$

Using (16), the solution to (18b) is given by

$$\widehat{\Omega}_N := Y_N \Pi^{\perp}_{U_N} \Phi_N^T \Psi_N, \tag{19}$$

$$\widehat{\mathcal{H}}_N := Y_N \left(I - \Pi^{\perp}_{U_N} \Phi_N^T \Psi_N \Phi_N\right) U_N^T \left(U_N U_N^T\right)^{-1}, \tag{20}$$

where $\Psi_N := \left(\Phi_N \Pi^{\perp}_{U_N} \Phi_N^T\right)^{-1}$. The existence of the inverse is guaranteed by the persistency of excitation condition.

3.2 A recursive solution

The least-squares estimate $\widehat{\Omega}_N$ to (18b) can be updated recursively by making use of the recursive 4SID algorithm proposed in [8, 9]. Define $P_N := (U_N U_N^T)^{-1}$, and let at time instant N the pair of input and output data (u_N, y_N) be given. Then, an update of the least-squares estimate $\widehat{\Omega}_{N-1}$ is given by the following recursive equations:

$$\widehat{\Omega}_N = \widehat{\Omega}_{N-1} - \beta_N (e_N + \widehat{\Omega}_{N-1} q_N) q_N^T \Psi_{N-1} \tag{21}$$

$$\Psi_N = \frac{1}{\gamma}\left(\Psi_{N-1} - \beta_N \Psi_{N-1} q_N q_N^T \Psi_{N-1}\right), \tag{22}$$

$$P_N = \frac{1}{\gamma}\left(P_{N-1} - \alpha_N P_{N-1} u_{s(N)} u_{s(N)}{}^T P_{N-1}\right), \tag{23}$$

$$Y_N U_N^T = \gamma Y_{N-1} U_{N-1}^T + y_{s(N)} u_{s(N)}{}^T, \tag{24}$$

$$\Phi_N U_N^T = \gamma \Phi_{N-1} U_{N-1}^T + \phi_{s(N-s)} u_{s(N)}{}^T, \tag{25}$$

$$\alpha_N := \left(\gamma + u_{s(N)}{}^T P_{N-1} u_{s(N)}\right)^{-1}, \tag{26}$$

$$\beta_N := \left(\frac{1}{\alpha_N} + q_N^T \Psi_{N-1} q_N\right)^{-1}, \tag{27}$$

$$e_N := y_{s(N)} - Y_{N-1} U_{N-1}^T P_{N-1} u_{s(N)}, \tag{28}$$

$$q_N := \Phi_{N-1} U_{N-1}^T P_{N-1} u_{s(N)} - \phi_{s(N-s)}, \tag{29}$$

where $\gamma \leq 1$ is a forgetting factor Note that the derivation of the above equations except for (21) can be found in the references. Therefore, it is sufficient to prove (21). From the references, the product $Y_N \Pi^{\perp}_{U_N} \Phi_N^T$ in (19) can be updated by the following equation:

$$Y_N \Pi^{\perp}_{U_N} \Phi_N^T = \gamma Y_{N-1} \Pi^{\perp}_{U_{N-1}} \Phi_{N-1}^T - \gamma \alpha_N e_N q_N^T. \tag{30}$$

Hence, substitution of (22) and (30) into (19) gives the equation (21), using the fact that $1 - \beta_N q_N^T \Psi_{N-1} q_N = \frac{\beta_N}{\alpha_N}$. This equation can be easily derived from the definition of β_N.

Instead of starting the above recursion with the quantities $\widehat{\Omega}_{N-1}$, Ψ_{N-1}, $Y_{N-1}U_{N-1}^T$ and $\Phi_{N-1}U_{N-1}^T$, (which could have been determined off-line) the recursion could start with properly chosen initial values of the latter quantities. In section 4, a particular initialization mechanism to re-initialize the recursion when a change is detected will be proposed.

3.3 An alternative cost function

When we are only interested in the quantity Ω ($\approx \mathcal{OL}$), we introduce the following alternative to (18a):

$$J_N(\Omega) = \text{Trace}\, \epsilon \Pi^{\perp}_{U_N} \epsilon^T \tag{31}$$
$$= \text{Trace}\, (Y_N - \Omega\Phi_N) \Pi^{\perp}_{U_N} (Y_N - \Omega\Phi_N)^T.$$

Note that $\widehat{\Omega}_N$ in (19) also minimizes of the cost function (31) since

$$\frac{\partial J_N}{\partial \Omega} = -2 Y_N \Pi^{\perp}_{U_N} \Phi_N^T + 2\Omega \Phi_N \Pi^{\perp}_{U_N} \Phi_N^T,$$

and again we have (19) by solving $\frac{\partial J_N}{\partial \Omega} = 0$.

Using (19), the equation (31) can be recast into the following quadratic form:

$$J_N(\Omega) = \text{Trace}\left(\Omega - \widehat{\Omega}_N\right) \Psi_N^{-1} \left(\Omega - \widehat{\Omega}_N\right)^T \tag{32}$$
$$+\text{Trace}\, Y_N \Pi^{\perp}_{U_N} Y_N^T - \text{Trace}\, Y_N \Pi^{\perp}_{U_N} \Phi_N^T \Psi_N \Phi_N \Pi^{\perp}_{U_N} Y_N^T.$$

The last two terms in the right hand side of (32) are independent of the parameter Ω. Therefore, we notice that the first term of the right hand side of (32) has the property

$$\text{Trace}\left(\Omega - \widehat{\Omega}_N\right) \Psi_N^{-1} \left(\Omega - \widehat{\Omega}_N\right)^T \begin{cases} = 0, & \Omega = \widehat{\Omega}_N, \\ > 0, & \text{otherwise.} \end{cases}$$

The recursive update of the cost function, see appendix A, is

$$J_N(\Omega) = J_{N-1}(\Omega) + \alpha_N \|e_N + \Omega q_N\|^2. \tag{33}$$

3.4 Covariance estimate

Let the additive perturbation $\varepsilon_i \in \mathcal{R}^l$ defining the matrix ϵ_N be Gaussian and have the following property,

$$\mathbf{E}\varepsilon_i = 0, \qquad \mathbf{E}\varepsilon_i \varepsilon_j^T = \begin{cases} R > 0, & i = j, \\ 0, & i \neq j. \end{cases} \tag{34}$$

Assume that there exist true parameters Ω_0 and $\mathcal{H}_0$ such that

$$Y_N = \begin{bmatrix} \Omega_0 & \mathcal{H}_0 \end{bmatrix} \begin{bmatrix} \Phi_N \\ U_N \end{bmatrix} + \epsilon_N. \tag{35}$$

Suppose we are allowed to define the covariance of the matrix quantity $\widehat{\Omega}_N$ as

$$\text{Cov}(\widehat{\Omega}_N) := \mathbf{E}(\widehat{\Omega}_N - \Omega_0)^T (\widehat{\Omega}_N - \Omega_0)$$
$$= \left(\Phi_N \Pi^{\perp}_{U_N} \Phi_N^T\right)^{-1} \Phi_N \Pi^{\perp}_{U_N} \mathbf{E}\left[\epsilon^T \epsilon\right]$$
$$\cdot \Pi^{\perp}_{U_N} \Phi_N^T \left(\Phi_N \Pi^{\perp}_{U_N} \Phi_N^T\right)^{-1} \tag{36}$$
$$= \sigma_\epsilon^2 s \left(\Phi_N \Pi^{\perp}_{U_N} \Phi_N^T\right)^{-1}. \tag{37}$$

Taking account of the properties of the trace operator and the Gaussian properties of ε_i, the equation (37) shows that,

$$D(\widehat{\Omega}_N, \Omega_0) := \text{Trace}\, (\widehat{\Omega}_N - \Omega_0) \frac{1}{\sigma_\epsilon^2 s} \left(\Phi_N \Pi^{\perp}_{U_N} \Phi_N^T\right) (\widehat{\Omega}_N - \Omega_0)^T \tag{38}$$

has the χ^2-distribution with ls degrees of freedom, denoted by $\chi^2(ls)$ when $\widehat{\Omega}_N - \Omega_0$ is of rank ls. The probability that

$$D(\widehat{\Omega}_N, \Omega_0) \geq a \tag{39}$$

is $\chi^2_a(ls)$, the a-level of the $\chi^2(ls)$-distribution.

Since the variance of modeling error σ_ϵ^2 is usually unknown, our concern shifts to how to estimate it by the off-line least squares. Similarly to lemma II.1 of the book [7], we have the following lemma:

Lemma 1 *Let the criterion be given by (31) and suppose that (34) holds. Then,*

$$\widehat{\sigma}^2_{\epsilon N} := \frac{1}{s[(N-s+1)-2ms]} J(\widehat{\Omega}_N) \qquad (40)$$

is an unbiased estimate of σ_ϵ^2.
(The proof is omitted here.)

4 Change detection and re-initialization of recursive subspace identification

4.1 Change detection

Based on the covariance estimate derived in section 3.4, we can develop a hypothesis test for detecting a change in the parameter Ω in the least-squares problem (31).

Suppose that we have two pair of estimates $(\widehat{\Omega}_N, \widehat{\Psi}_N)$ and $(\overline{\Omega}_N, \overline{\Psi}_N)$ at the sampling instant N. The first pair is obtained with a recursive solution given by the equations (21–29) for $\gamma = 1$ and the second by solving a fixed windowed least squares problem

$$\overline{\Omega}_N = \arg\min \text{Trace} \left(\overline{Y}_N - \Omega\overline{\Phi}_N\right) \Pi^{\perp}_{\overline{U}_N} \left(\overline{Y}_N - \Omega\overline{\Phi}_N\right)^T$$

with

$$\overline{U}_N := \left[\begin{array}{ccc} u_{s(N-L+2s)} & \cdots & u_{s(N)} \end{array} \right], \qquad (41)$$

$$\overline{U}_{P,N} := \left[\begin{array}{ccc} u_{s(N-L+s)} & \cdots & u_{s(N-s)} \end{array} \right], \qquad (42)$$

$\overline{\Phi}_N := \left[\begin{array}{ccc} \phi_{s(N-L+s)} & \cdots & \phi_{s(N-s)} \end{array} \right]$ and the output Hankel matrix $\overline{Y}_N$ is defined similarly to (41). For brevity, we will indicate the scheme producing the estimates $\widehat{\Omega}_N$, $\widehat{\Psi}_N$ by the accurate 4SID and the one producing the estimates $\overline{\Omega}_N$, $\overline{\Psi}_N$.

Note that the following inequality

$$\overline{\Phi}_N \Pi^{\perp}_{\overline{U}_N} \overline{\Phi}^T_N =: \overline{\Psi}^{-1}_N \leq \widehat{\Psi}^{-1}_N = \Phi_N \Pi^{\perp}_{U_N} \Phi^T_N. \qquad (43)$$

is obained from the following lemma:

Lemma 2 *Suppose that wide matrices with the same row size $U_1 \in \mathbb{R}^{l\times n_1}$, $U_2 \in \mathbb{R}^{l\times n_2}$, $\Phi_1 \in \mathbb{R}^{l\times n_1}$ and $\Phi_2 \in \mathbb{R}^{l\times n_2}$ are row full rank, where $l < n_1$ and $l < n_2$. Define $U = [U_1\ U_2]$ and for $i = 1, 2$*

$$\Pi^{\perp}_{U_i} := I_{n_i} - U_i^T (U_i U_i^T)^{-1} U_i^T, \qquad (44)$$

$$\Pi^{\perp}_{U} := I_{n_1+n_2} - U^T (U U^T)^{-1} U^T, \qquad (45)$$

where the subscript of I denotes the size of the identity matrix. Assume that $\Phi_i \Pi^{\perp}_{U_i} \neq 0$ and $[\Phi_1\ \Phi_2]\Pi^{\perp}_U \neq 0$. Then, the following inequality holds:

$$\left[\begin{array}{cc} \Phi_1 & \Phi_2 \end{array} \right] \Pi^{\perp}_U \left[\begin{array}{c} \Phi_1^T \\ \Phi_2^T \end{array} \right] - \Phi_2 \Pi^{\perp}_{U_2} \Phi_2^T \geq 0. \qquad (46)$$

(The proof of the lemma is omitted.)

Following the results of section 3.4 and [6], it is proposed to use (38) to measure the distance from the estimated parameters to the true parameter. For the estimate $\overline{\Omega}_N$, this becomes

$$D(\overline{\Omega}_N, \Omega_0) := \text{Trace}\ (\overline{\Omega}_N - \Omega_0) \frac{1}{s\sigma_\epsilon^2} \overline{\Psi}^{-1}_N (\overline{\Omega}_N - \Omega_0)^T,$$

However, the true value Ω_0 is unknown. Therefore, instead of the true one, we use the estimate $\widehat{\Omega}_N$ given by the accurate 4SID as the reference value, and the proposed distance becomes

$$D(\overline{\Omega}_N, \widehat{\Omega}_N) = \text{Trace}\ (\overline{\Omega}_N - \widehat{\Omega}_N) \frac{1}{s\overline{\sigma}^2_{\epsilon N}} \overline{\Psi}^{-1}_N (\overline{\Omega}_N - \widehat{\Omega}_N)^T. \qquad (47)$$

Now, let a be a threshold designed according to the χ^2-distribution with ls degrees of freedom. Then, we perform the following test to detect a change:

$$\begin{array}{l} \text{if } D(\overline{\Omega}_N, \widehat{\Omega}_N) \leq a\text{: no change has occurred,} \\ \text{if } D(\overline{\Omega}_N, \widehat{\Omega}_N) > a\text{: a change has occurred.} \end{array} \qquad (48)$$

4.2 Re-initialization of the accurate 4SID

Once a change at time instant N is detected by the method proposed in the previous section, The recursive update equations are re-initialized at time instant N. The method is inspired by [6].

According to the *minimum distance principle* described in the reference, the following constrained minimization problem is proposed:

$$\min_{\Omega} D(\overline{\Omega}_N, \Omega) \qquad (49)$$

subject to,

$$D(\widehat{\Omega}_N, \Omega) - D(\overline{\Omega}_N, \Omega) < a, \qquad (50)$$

where a is the threshold used in the hypothesis test (48), and

$$D(\widehat{\Omega}_N, \Omega) := \text{Trace}\ (\widehat{\Omega}_N - \Omega) \frac{1}{s\overline{\sigma}^2_{\epsilon N}} \widehat{\Psi}^{-1}_N (\widehat{\Omega}_N - \Omega)^T,$$

$$D(\overline{\Omega}_N, \Omega) := \text{Trace}\ (\overline{\Omega}_N - \Omega) \frac{1}{s\overline{\sigma}^2_{\epsilon N}} \overline{\Psi}^{-1}_N (\overline{\Omega}_N - \Omega)^T.$$

By using a Lagrange multiplier μ, the constrained minimization problem can be solved by minimizing the following cost function:

$$\begin{aligned} Q(\Omega, \mu) :=& D(\overline{\Omega}_N, \Omega) + \mu \left(D(\widehat{\Omega}_N, \Omega) - D(\overline{\Omega}_N, \Omega) - a \right) \\ =& (1-\mu)\text{Trace}\ (\overline{\Omega}_N - \Omega) \frac{1}{s\overline{\sigma}^2_{\epsilon N}} \widehat{\Psi}^{-1}_N (\overline{\Omega}_N - \Omega)^T \\ &+ \mu \text{Trace}\ (\widehat{\Omega}_N - \Omega) \frac{1}{s\overline{\sigma}^2_{\epsilon N}} \widehat{\Psi}^{-1}_N (\widehat{\Omega}_N - \Omega)^T - \mu a \end{aligned}$$

Then, we have the following partial derivative of Q by Ω:

$$\frac{\partial Q}{\partial \Omega} = 2(1-\mu)\left(\Omega - \overline{\Omega}_N\right)\overline{\Psi}_N^{-1} + 2\mu\left(\Omega - \widehat{\Omega}_N\right)\widehat{\Psi}_N^{-1}. \tag{51}$$

Solving $\frac{\partial Q}{\partial \Omega} = 0$ for Ω yields the following re-initialization for the accurate 4SID:

$$\Omega_N^*(\mu) := \left((1-\mu)\overline{\Omega}_N\overline{\Psi}_N^{-1} + \mu\widehat{\Omega}_N\widehat{\Psi}_N^{-1}\right) \cdot \left((1-\mu)\overline{\Psi}_N^{-1} + \mu\widehat{\Psi}_N^{-1}\right)^{-1}. \tag{52}$$

The weighting factor μ is designed by solving the equation $f(\mu) = 0$ for μ, where the function f of μ is defined as

$$f(\mu) := s\overline{\sigma_\epsilon}_N^2\left(D(\widehat{\Omega}_N, \Omega_N^*(\mu)) - D(\overline{\Omega}_N, \Omega_N^*(\mu)) - a\right) \tag{53}$$

Note the fact that the function $f(\mu)$ is a decreasing function on $(0,1)$ and satisfies $f(0)f(1) < 0$. (The proof of the fact is omitted here.) This means that on the interval $(0,1)$ there exists one and only solution, denoted by μ_N^o, to the equation $f(\mu) = 0$.

Using the weighting factor $\mu_N^o \in (0,1)$, the re-initialization of the accurate 4SID can be performed according to the following procedure:

Re-initialization of the accurate 4SID *Suppose a change is detected at the sampling time N then, the quantities of the accurate 4SID are re-initialized according to the following replacement with the weighting factor $\mu_N^o \in (0,1)$:*

$$\widehat{\Psi}_N \longleftarrow \left((1-\mu_N^o)\overline{\Psi}_N^{-1} + \mu_N^o\widehat{\Psi}_N^{-1}\right)^{-1}, \tag{54}$$

$$\widehat{\Omega}_N \longleftarrow \Omega_N^*(\mu_N^o), \quad P_N \longleftarrow \left((1-\nu_N^o)\overline{P}_N^{-1} + \nu_N^o P_N^{-1}\right)^{-1},$$

$$Y_N U_N^T \longleftarrow (1-\nu_N^o)\overline{Y}_N\overline{U}_N^T + \nu_N^o Y_N U_N^T,$$

$$\Phi_N U_N^T \longleftarrow (1-\nu_N^o)\overline{\Phi}_N\overline{U}_N^T + \nu_N^o \Phi_N U_N^T,$$

where the left arrow "$\longleftarrow$" denote the replacement of the left hand side by the right hand side, and ν_N^o is determined by solving the following minimization problem derived from (52):

$$\nu_N^o := \arg\min_{0\le\nu\le 1} \Big\| \left((1-\nu)\overline{Y}_N\overline{U}_N^T + \nu Y_N U_N^T\right) \cdot \left((1-\nu)\overline{P}_N^{-1} + \nu P_N^{-1}\right)^{-1} \left((1-\nu)\overline{\Phi}_N\overline{U}_N^T + \nu\Phi_N U_N^T\right)^T - (1-\mu_N^o)\overline{Y}_N\overline{U}_N^T\overline{P}_N\overline{U}_N\overline{\Phi}_N^T - \mu_N^o Y_N U_N^T P_N U_N \Phi_N^T \Big\|.$$

Remark 1 Note that, since $\mu_N^o \in (0,1)$, the re-initialized (54) satisfies

$$(1-\mu_N^o)\overline{\Psi}_N^{-1} + \mu_N^o\widehat{\Psi}_N^{-1} \ge (1-\mu_N^o)\overline{\Psi}_N^{-1} + \mu_N^o\overline{\Psi}_N^{-1} = \overline{\Psi}_N^{-1}.$$

This means that the condition (43) holds even after re-initialization.

5 Numerical example

Let us consider the following 3rd-order time-varying SISO system

$$y_k = \frac{B(q)}{A(q)}u_k + v_k, \quad B(q) := 0.2q^{-1} + 0.08q^{-2},$$

$$A(q) := 1 - (\sqrt{3}r_1 + r_2)q^{-1} + (r_1^2 + \sqrt{3}r_1 r_2)q^{-2} - r_1^2 r_2 q^{-3},$$

where q^{-1} is the backward shift operator and the time-varying parameters, $r_1(k)$ and $r_2(k)$, displayed in figure 1, satisfying,

$$\left(r_1(k), r_2(k)\right) = \begin{cases} (0.95, 0.75), & 0 \le k < 2000, \\ (0.7, 0.9), & 3000 \le k < 5000, \\ (0.95, 0.75), & 6250 \le k \le 9000, \end{cases}$$

where, on the interval $[2000, 3000)$ (or $[5000, 6250)$), these parameters change linearly from the values at 2000 (or 5000) to 3000 (or 6250), respectively. During the whole time interval there are several changes of the dynamics of the input signal. For the time intervals, $[0, 1500]$ and $[3501, 7000]$, a zero-mean white sequence with unit variance is used as the input signal. For the intervals $[1501, 3500]$ and $[7001, 9000]$, the input equals the sum of a zero-mean white sequence with unit variance filtered with a 10th-order Butterworth filter (cutoff 0.6 times the Nyquist frequency) and a zero-mean white sequence with variance 0.01. The measured output is contaminated by a zero-mean white noise v_k which is uncorrelated with the input signal. The SNR is approximately 23.6dB. We adopt $s = 15$. Since the dimension of the output is $l = 1$, the degree of freedom of the χ^2-distribution should be taken as $ls = 15$. We use the recursive 4SID algorithm with the exponential forgetting factor $\gamma_L = 0.99$ as the tracking 4SID. This corresponds approximately to window size of 113.

Fig. 3 shows the result of the χ^2-test by (48). This figure clearly shows that only around the intervals $[2000, 3000)$ and $[5000, 6250)$ the estimate from the accurate 4SID is re-initialized. Our method can detect these system changes. Moreover, our hypothesis test does not detect a change when the input dynamics changes abruptly. This shows that our method is not sensitive to (known) changes of the input dynamics. It is reasonable because our decision rule takes account of the orthogonal subspace to the input Hankel matrix. This may be inferred by comparing figures 3 and 4. Although around the interval $[8000, 9000]$, the maximum singular value of the estimate by the tracking 4SID in the lower graph of figure 4 seems to be perturbed seriously, the perturbation does not influence the χ^2-test.

Figure 5 illustrates the weighting factor for the accurate 4SID which is designed by the procedure in subsection 4.2. Only after changes are detected, the weighting factor is activated, namely, it takes the number less than 1.

6 Conclusion

We have proposed a parameter tracking method using two PI-MOESP scheme subspace identification in parallel. The method can detect not only abrupt changes but also incipient changes in the dynamics of a linear system. The method consists of a change detection scheme and a re-initialization of the recursive identification algorithm which is invoked only when a change has been detected. A change is detected by checking whether the estimate given by the accurate algorithm lies in the confidence interval of the estimate obtained by the tracking algorithm. Once a change has been detected, the former algorithm is re-initialized via solving a constrained least squares problem. The extension of this scheme to PO-MOESP will be addressed in future research.

Due to space limitations, we have omitted the proofs from the paper. For the interested readers, please contact the first author at *oku@tn.utwente.nl* to get copies of the proofs.

References

[1] M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes.* Prentice Hall, New Jersey, 1993.

[2] T. Gustafsson, M. Lovera, and M. Verhaegen. A novel algorithm for recursive instrumental variable based subspace identification. In *Proceedings of the 37th CDC*, pages 3920–3925, Florida, December 1998.

[3] R. A. Horn and C. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.

[4] M. Jansson and B. Wahlberg. A linear regression approach to state-space subspace system identification. *Signal Processing*, 52:103–129, 1996.

[5] M. Jansson and B. Wahlberg. On consistency of subspace methods for system identification. *Automatica*, 34(12):1507–1519, 1998.

[6] R. Kulhavý and M. B. Zarrop. On a general concept of forgetting. *Int. J. Control*, 58(4):905–924, 1993.

[7] L. Ljung. *System Identification, 2nd Ed.* Prentice Hall, New Jersey, 1999.

[8] H. Oku. *Sequential subspace state-space system identification and state estimation of unknown multivariable systems.* PhD thesis, University of Tokyo, 2000.

[9] H. Oku and H. Kimura. A recursive 4SID from the input-output point of view. *Asian J. Control*, 1(4):258–269, 1999.

[10] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems.* Kluwer Academic Publishers, Massachusetts, 1996.

[11] M. Verhaegen. Subspace model identification Part 3: Analysis of the ordinary output-error state-space model identification algorithm. *Int. J. Control*, 58(3):555–586, 1993.

[12] M. Verhaegen and E. Deprettere. A fast, recursive MIMO state space model identification algorithm. In *Proceedings of the 30th CDC*, pages 1349–1354, Brighton, December 1991.

[13] M. Verhaegen and P. Dewilde. Subspace model identification Part I: The output-error state space model identification class of algorithm. *Int. J. Control*, 56(5):1187–1210, 1992.

[14] M. Verhaegen and P. Dewilde. Subspace model identification Part II: Analysis of the elementary output-error state-space model identification algorithm. *Int. J. Control*, 56(5):1211–1241, 1992.

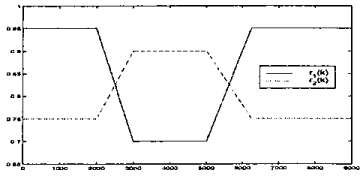

Figure 1: Time varying parameters $r_1(k)$ and $r_2(k)$.

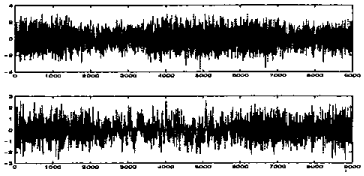

Figure 2: Input data, contaminated output data and measurement noise. The dynamics in the input data has changed at $N = 1500$, 3500 and 7000. The system has changed incipiently during the intervals [2000, 3000] and [5000, 6250].

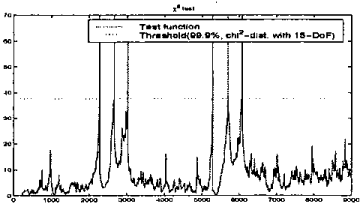

Figure 3: χ^2-test (48). Only the changes of the system have been detected.

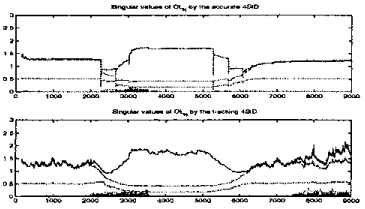

Figure 4: Singular values of the estimates by the accurate 4SID (upper) and the tracking 4SID (lower).

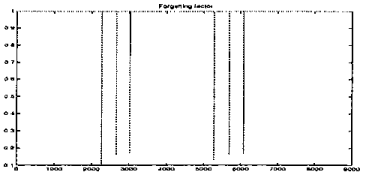

Figure 5: Weighting factor μ_N^o for the accurate 4SID. This weighting factor is activated only if a change is detected.