A signal for online whiteness testing generated by recursive subspace model identification from closed-loop data

Hiroshi Oku
Dept. of Robotics, Osaka Institute of Technology
5-16-1, Omiya, Asahi-ku, Osaka 535-8585, Japan
E-mail: oku@elc.oit.ac.jp

Abstract
In this paper, we will develop an online statistical change detection method, which is used for detection of changes in a system under surveillance in closed-loop. A recursive algorithm of closed loop subspace model identification is presented, which is based on the matrix inversion lemma. The relation between the proposed recursive algorithm and the predictor-based subspace identification method (PBSID) is clarified from the viewpoint of a matrix-valued least squares problem. A test signal generated sequentially by the recursive algorithm is studied and its asymptotic whiteness is proved explicitly. The proposed online change detection method is based on a likelihood ratio test to examine the change in the covariance of the aforementioned test signal.

1 Introduction
Recently, change (or fault) detection methods of monitoring multivariable systems have been in great demand [1]. For more than a decade, subspace-based statistical change detection algorithms have been studied by several research groups, e.g., [2, 3, 4, 5, 6, 7] and the references therein. Advantages of subspace-based change detection methods are twofold. First, they are useful for multivariable systems. Second, algorithms based on subspace model identification work not only as system identification tools but also as whitening filters more precisely than conventional least squares and least mean squares algorithms do [3, 4, 5].

In this paper, we will develop an online statistical change detection method which is useful for multivariable systems in closed-loop. The contributions of the paper are as follows. First, we derive a matrix-inversion-lemma-based recursive algorithm for updating matrix-valued parameters, and clarify the mechanism of sequentially generating a test signal used for statistical change detection from the proposed recursive algorithm. Second, we show the relation between the predictor-based subspace identification (PBSID, [8, 9, 10, 11, 12]) method and the proposed recursive algorithm in terms of a matrix-valued least squares problem. Third, we show the asymptotic whiteness of the test signal explicitly. Fourth, using the test signal, we develop an online statistical change detection method based on a likelihood ratio test and a GMA-based recursive algorithm [13].

We must emphasize that the proposed recursive algorithm is derived from the recursification of the solution to a matrix-valued least-squares problem with the Frobenius norm introduced for measuring the distance between two matrices [14], and it is different from the recursive algorithm presented in [15]. We must mention that the preceding studies [6, 7] have also considered subspace based fault detection for systems in close-loop. However, their test statistics are generated by batch-processing using data from a sliding window of a fixed interval. It is different from our online recursive update approach.

The paper is organized as follows. In section 2, we propose a recursive algorithm of closed-loop subspace model identification, which is based on the matrix inversion lemma. The relation between PBSID and the proposed recursive algorithm is investigated in terms of a matrix-valued least squares problem. In section 3, asymptotic whiteness of a sequence generated sequentially by the proposed recursive algorithm is clarified. Finally, an online statistical change detection method for systems in closed-loop is developed in section 4. The aforementioned sequence is adopted as the test signal for the online change detection. The proposed method is based on a likelihood ratio test to examine the change in the covariance of the test signal. A numerical simulation is presented in section 5 to demonstrate the effectiveness of the proposed method.

2 Recursive algorithm of closed-loop subspace model identification

2.1 Problem formulation for recursive closed-loop subspace model identification
Let us assume that a system to be identified can be described by the innovations form as follows ([16]):

\[ x_{k+1} = Ax_k + Bu_k + K e_k, \]  
\[ y_k = Cx_k + Du_k + e_k, \]  

(1a)  
(1b)
where $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^l$, $u_k \in \mathbb{R}^m$. The innovation $e_k \in \mathbb{R}^l$ is an independently identically distributed Gaussian noise. $A, B, C, D, K$ are real matrices with appropriate dimensions. Feedback from the output $y_k$ to the input $u_k$ is allowed. Henceforth, for the sake of simplicity, $D = 0$ is assumed to ensure the well-posedness of the feedback loop ([17, 18]).

The goal of this section is to obtain a recursive identification algorithm to estimate parameters related to $(A, B, C, K)$ in (1) from closed-loop data $\{(u_k, y_k)\}$.

2.2 Notations

Given a finite integer $N$, define the matrix, denoted by $U_{t,N}$, whose columns consist of a finite sequence of $u_t$ as follows:

$$U_{t,N} = \begin{bmatrix} u_t & u_{t+1} & \cdots & u_{t+N-1} \end{bmatrix}. \quad (2)$$

The matrices $X_{t,N}$, $Y_{t,N}$ and $E_{t,N}$ are defined in the same manner as $U_{t,N}$ in (2).

For $s > n$, define the following block Hankel matrix, $U_{k,N}$, whose block rows consist of a bunch of consecutive matrices $U_{k,N}$ ($t \leq k \leq t + s - 1$) as follows,

$$U_{k,N} = \begin{bmatrix} U_{t,N} & U_{t+1,N} & \cdots & U_{t+N-1,N} \\ U_{t+s-2,N} & \vdots & \cdots & \vdots \\ U_{t+s-1,N} \end{bmatrix} = \begin{bmatrix} u_t & u_{t+1} & \cdots & u_{t+N-1} \\ u_{t+1} & u_{t+2} & \cdots & u_{t+N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{t+s-2} & u_{t+s-1} & \cdots & u_{t+N+s-3} \\ u_{t+s-1} & u_{t+s} & \cdots & u_{t+N+s-2} \end{bmatrix} = \begin{bmatrix} u_s(t) & u_s(t+1) & \cdots & u_s(t+N-1) \end{bmatrix},$$

where the size of $U_{k,N}$ is of $m \times 2N$, and $u_s(t + k)$ denotes the vector whose block entries consist of a bunch of $s$ consecutive data samples from $u_{t+k}$ to $u_{t+k+s-1}$. Similarly, define the matrix $U_{k,N}$ with the size of $m(s-1) \times N$ as follows:

$$U_{k,N} = \begin{bmatrix} U_{t,N} & U_{t+1,N} & \cdots & U_{t+N-1,N} \\ U_{t+s-2,N} & \vdots & \cdots & \vdots \\ U_{t+s-1,N} \end{bmatrix} = \begin{bmatrix} u_t & u_{t+1} & \cdots & u_{t+N-1} \\ u_{t+1} & u_{t+2} & \cdots & u_{t+N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{t+s-2} & u_{t+s-1} & \cdots & u_{t+N+s-3} \\ u_{t+s-1} & u_{t+s} & \cdots & u_{t+N+s-2} \end{bmatrix} = \begin{bmatrix} u_s(t) & u_s(t+1) & \cdots & u_s(t+N-1) \end{bmatrix}.$$ 

The block Hankel matrices $Y_{t,N}$, $X_{t,N}$, $Y_{t,N}$ and $E_{t,N}$ are defined similarly to $U_{k,N}$ and $U_{k,N}$, respectively.

The joint input-output data matrices, denoted by $Z_{t,N}$ and $Z_{t,N}$, are respectively defined as follows:

$$Z_{t,N} = \begin{bmatrix} U_{t,N} & Y_{t,N} \\ \hat{U}_{t,N} & \hat{Y}_{t,N} \end{bmatrix} = \begin{bmatrix} z_s(t) & z_s(t+1) & \cdots & z_s(t+N-1) \end{bmatrix},$$

$$Z_{t,N} = \begin{bmatrix} \hat{U}_{t,N} & \hat{Z}_{t,N} \\ \hat{Y}_{t,N} \end{bmatrix} = \begin{bmatrix} z_s(t) & z_s(t+1) & \cdots & z_s(t+N-1) \end{bmatrix},$$

where joint input-output vectors are defined $z_s(k) = \left[ u_s(k)^T y_s(k)^T \right]^T$, $z_s(k) = \left[ u_s(k)^T y_s(k)^T \right]^T$.

For future reference the following matrices are defined, where $\bar{A} := A - KC$:

$$\mathcal{O} = \begin{bmatrix} C^T (CA)^T \cdots (CA^{n-1})^T \end{bmatrix}^T,$$

$$\mathcal{L} = \begin{bmatrix} \bar{L}_B & \bar{L}_K \end{bmatrix},$$

$$\mathcal{H} = \begin{bmatrix} \begin{bmatrix} 0 & \cdots & 0 \\ CB & 0 & \cdots \end{bmatrix} & \begin{bmatrix} C A^{-2} B & CA^{-3} B \cdots & CB \end{bmatrix} \\ \begin{bmatrix} C A & 0 \cdots \end{bmatrix} & \begin{bmatrix} 0 & \cdots & 0 \\ CK & 0 & \cdots \end{bmatrix} \\ \begin{bmatrix} C A^{-2} K & CA^{-3} K \cdots & CK \end{bmatrix} \end{bmatrix}$$

2.3 Input-output relation of data Hankel matrices

Assumption 1 $\bar{A} = A - KC$ is stable.

Note that assumption 1 implies that all the eigenvalues of $\bar{A}$ lie strictly inside the unit circle in the complex plane.

Assumption 2 $(\bar{A}, C)$ is observable and $(\bar{A}, [B K])$ is reachable.

Substitute $e_k = y_k - C x_k$ for $e_k$ in (1a), and we have

$$x_{k+1} = A x_k + B u_k + K (y_k - C x_k) = A x_k + B u_k + K y_k. \quad (3)$$

Using the finite sequences with $k$ from 1 to $N + 2s - 1$, recursive regression of (1b) and (3) yields the following matrix input-output relation:

$$Y_{k,N} = \mathcal{O} \bar{X}_{1,N} + \mathcal{L} Z_{1,N} + \mathcal{H} Z_{k,N} + \mathcal{E}_{k,N} \approx \begin{bmatrix} \Theta \mathcal{H} \end{bmatrix} \begin{bmatrix} Z_{1,N} \\ Z_{k,N} \end{bmatrix} + \mathcal{E}_{k,N}, \quad (4)$$

\[ \text{• 156 •} \]
where \( k_0 = s + 1 \) and \( \Theta = \tilde{\Theta}C \). Note that, with respect to the approximation in (4), the term \( \tilde{\Theta}A^*X_1 \) can be negligible for a sufficiently large \( s \) due to assumption 1.

2.4 Least squares estimate of \( \Theta \)

Assumption 3 Given a positive integer \( N_0 \). The closed-loop signals are persistently exciting in the sense that the following relations hold for \( \forall N > N_0 \) including infinity:

\[
\frac{1}{N}Z_{k_0,N}^T \begin{bmatrix} Z_{1,N} \\ Z_{k_0,N} \end{bmatrix} > 0, \\
\frac{1}{N} \begin{bmatrix} Z_{1,N} \\ Z_{k_0,N} \end{bmatrix} ^T \begin{bmatrix} Z_{1,N}^T \\ Z_{k_0,N}^T \end{bmatrix} > 0.
\]

Consider the following cost function w.r.t. \( \Theta \) and \( \tilde{\Theta} \):

\[
J(\Theta, \tilde{\Theta}) = \left\| Y_{k_0,N} - \begin{bmatrix} \Theta & \tilde{\Theta} \end{bmatrix} \begin{bmatrix} Z_{1,N} \\ Z_{k_0,N} \end{bmatrix} \right\|_F^2, (5)
\]

where the Frobenius norm of a matrix \( X \) is defined as \( \|X\|_F = \sqrt{\text{trace}(X^T X)} \) [14]. Under the assumption 3, similarly to the open-loop case [19], the minimizer of the cost function \( J(\Theta, \tilde{\Theta}) \), denoted by \( (\hat{\Theta}_N, \tilde{\Theta}_N) \), can be given as follows:

\[
\hat{\Theta}_N = \frac{1}{N}Y_{k_0,N} \Pi_{Z_{k_0,N}} \begin{bmatrix} Z_{1,N}^T \\ Z_{k_0,N} \end{bmatrix} \left( \frac{1}{N}Z_{1,N} \Pi_{Z_{k_0,N}} Z_{1,N}^T \right)^{-1}, (6)
\]

\[
\tilde{\Theta}_N = \frac{1}{N} \left( Y_{k_0,N} - \hat{\Theta}_N Z_{1,N} \right) Z_{k_0,N}^{-T} \left( \frac{1}{N}Z_{k_0,N} Z_{k_0,N}^{-T} \right)^{-1},
\]

where \( \Pi_{Z_{k_0,N}} \) denotes a projection matrix defined as

\[
\Pi_{Z_{k_0,N}} = I - \frac{1}{N}Z_{k_0,N}^{-T} \left( \frac{1}{N}Z_{k_0,N} Z_{k_0,N}^{-T} \right)^{-1} Z_{k_0,N}^{-T}. (7)
\]

Note that the inverse matrices as above exist due to assumption 3.

Remark 1 Using the QR factorization of

\[
\begin{bmatrix} Z_{k_0,N}^- \\ Z_{1,N}^- \\ Y_{k_0,N}^- \end{bmatrix} = \begin{bmatrix} L_{11} \\ L_{21} & L_{22} \\ L_{31} & L_{32} & L_{33} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \\ Q_3^T \end{bmatrix}, (8)
\]

(6) can be rewritten as follows [20]:

\[
\hat{\Theta}_N = L_{32} L_{22}^{-1}. (9)
\]

2.5 Relation between \( \hat{\Theta}_N \) and the predictor-based subspace identification

A procedure of the PBSID (Predictor-Based Subspace Identification) [8, 9, 10, 11, 12] method is briefly reviewed as follows:

(1) Compute the oblique projection of \( \hat{Y}_{k_0,N} \) onto the space generated by the rows of \( \bar{Z}_{1,N} \) along the space generated by the rows of \( \bar{Z}_{k_0,N} \), denoted by \( \bar{Y}_{k_0,N} \):

\[
\bar{Y}_{k_0,N} = \tilde{\Pi}_{\bar{Z}_{k_0,N}} \left[ Y_{k_0,N} | \bar{Z}_{1,N} \right], (10)
\]

where \( \bar{Y}_{k_0,N} \) denotes a bunch of estimated output sequences. For the definition of the oblique projection, see appendixA.

(2) Note that \( \bar{Y}_{k_0,N} \approx \tilde{\Theta}X_{k_0,N} \). The singular value decomposition of \( W^{-1}Y_{k_0,N} \) with an adequate weighting matrix \( W \) gives an estimate of extended observability matrix and an estimated sequence of state vectors.

(3) Estimates of coefficient matrices \( (A, B, C, K) \) up to a similarity transform can be derived in the same way as those are derived in N4SID[21].

Note that the key point of the PBSID method is to obtain \( \bar{Y}_{k_0,N} \) in (10). Some calculation using the definition of the oblique projection and eqs. (6) and (8) leads to the relation

\[
\bar{Y}_{k_0,N} = \hat{\Theta}_N Z_{1,N}. (11)
\]

The relation shows that the PBSID method is interpretable as an least squares problem with the cost function of (5). This fact implies that the recursive algorithm derived in the following subsection can be thought of as a recursive version of predictor-based closed-loop subspace identification.

Remark 2 In the literature [8, 9, 10, 11, 12], the joint input-output data matrices are defined as such that input and output data are arrayed alternately as block row entries, e.g., \( Z_{k_0,N} = [U_{k_0,N} Y_{k_0,N}^T \cdots U_{k_0,N+1-N} Y_{k_0,N+1-N}^T]^T \), and they look very different from our corresponding matrices. Note that however, the difference between them is just in row-permutation and permutation matrices are orthogonal. Therefore, one will know that, in derivation of equations, the difference is not significant at all.

2.6 Recursive formula of updating \( \hat{\Theta}_N \)

Let us consider derivation of recursive formula of updating the following equations instead of eqs. (6) and (7) since in practice system identification is performed with finite-length data:

\[
\hat{\Theta}_N = Y_{k_0,N} \Pi_{Z_{k_0,N}} \begin{bmatrix} Z_{1,N}^T \\ Z_{k_0,N} \end{bmatrix} \left( \frac{1}{N}Z_{1,N} \Pi_{Z_{k_0,N}} Z_{1,N}^T \right)^{-1}, (12)
\]

\[
\Pi_{Z_{k_0,N}} = I - Z_{k_0,N}^{-T} \left( Z_{k_0,N} Z_{k_0,N}^{-T} \right)^{-1} Z_{k_0,N}^{-T}.
\]

Note that the matrices, \( Y_{k_0,N}, \bar{Z}_{k_0,N} \) and \( Z_{1,N} \) are composed of the data sampled during the interval
[1, \ N + 2s − 1]. Note also that, using the data sampled during the interval [1, \ N + 2s], we can describe the matrices \( Y_{k_0,N+1}, Z_{1,N+1} \) and \( Z_{k_0,N+1}^{-} \) as follows:

\[
\begin{align*}
\mathcal{Y}_{k_0,N+1} &= [Y_{k_0,N} \ y_s(N+k_0)], \\
Z_{1,N+1} &= [Z_{1,N} \ z_s(N+1)], \\
Z_{k_0,N+1}^{-} &= [Z_{k_0,N}^{-} \ z_{s-1}(N+k_0)].
\end{align*}
\]

Now, a recursive algorithm of closed-loop subspace model identification is given as follows:

**Proposition 1** The derivation of the latest estimate, \( \hat{\Theta}_{N+1} \), from the previous estimate, \( \hat{\Theta}_N \), and the vectors including the latest data, \( y_s(N+k_0), z_s(N+1) \) and \( z_{s-1}(N+k_0) \), is given by \(^1\):

\[
\hat{\Theta}_{N+1} = \hat{\Theta}_N - b_{N+1}\varepsilon_s(N+k_0)\xi_{N+1}^T \Psi_N^{-1}
\]

where the notations are defined as follows:

\[
\begin{align*}
\Psi_N &= Z_{1,N} \Pi_{Z_{k_0,N}}^{-1} Z_{1,N}^T, \\
\varepsilon_s(N+k_0) &= \begin{bmatrix} \varepsilon_s^{T} & \cdots & \varepsilon_s^{T+2s|N+2s-1} \end{bmatrix}^T, \\
\xi_{N+1} &= Z_{k_0,N}^{-} \left( Z_{k_0,N}^{-} \right)^{-1} z_{s-1}(N+k_0), \\
a_{N+1} &= \left( 1 + \xi_{N+1} \right)^{-1}, \\
b_{N+1} &= \left( a_{N+1}^2 + \xi_{N+1}^{-2} \right)^{-1}.
\end{align*}
\]

The recursive updating equations auxiliary to (13) are given as follows:

\[
\begin{align*}
\Psi_{N+1}^{-1} &= \Psi_N^{-1} - b_{N+1}\varepsilon_s(N+k_0)\xi_{N+1}^T \Psi_N^{-1}, \\
\left(Z_{t,N+1}^{-} Z_{t,N+1}^{-} \right)^{-1} &= \left(Z_{k_0,N}^{-} Z_{k_0,N}^{-} \right)^{-1}, \\
\cdot &= \left(Z_{k_0,N}^{-} Z_{k_0,N}^{-} \right)^{-1} z_{s-1}(N+k_0), \\
\sum_{t=1}^{N} y_s(N+k_0) &= \Theta_s z_s(N+1) + \hat{\Theta}_s \xi_{N+1} + \hat{\Theta}_s \xi_{N+1}, \\
\sum_{t=1}^{N} y_s(N+k_0) &= \Theta_s z_s(N+1) + \hat{\Theta}_s \xi_{N+1} + \hat{\Theta}_s \xi_{N+1}.
\end{align*}
\]

**Remark 3** One may prefer the recursive update of the QR-factorization-based estimate (9) to that of the estimate based on the matrix inversion lemma as above. In such a case, similarly to the literature, e.g., [22, 23], pivoting using Givens rotations [14] are helpful for derivation of a relevant recursive updating formula.

**Remark 4** Similarly to the conventional recursive least squares algorithm, an exponential forgetting factor can easily be installed in the recursive formula in Proposition 1.

\(^1\)The proposed recursive formula is based on the matrix inversion lemma. The derivation of the formula is very similar to one presented in [19].

## 3 On asymptotic whiteness of a sequence

When the system (1) to be identified stays at a stationary point and \( N \) is sufficiently large, the recursion formula (13) implies that the second term on the right hand side of (13) is a perturbation which is driven by \( \varepsilon_s(N+k_0) \). Here, we must emphasize that the vector \( \varepsilon_s(N+k_0) \) can be generated on-line at every sampling period from the recursive algorithm composed of (13), (14) and their auxiliary update equations. In this section, asymptotic whiteness of a sequence derived from the recursive algorithm is investigated.

**Assumption 4** An integer \( s \) is sufficiently large but finite.

**Assumption 5** (Inverse boundness) For \( \forall N > N_0 \) including infinity there exists a positive number \( \beta < \infty \) such that

\[
\left\| \frac{1}{N} Z_{k_0,N}^{-} Z_{k_0,N}^{-} \right\|_2 \leq \beta.
\]

**Assumption 6** There exists a stationary point \( (\Theta_s, \hat{\Theta}_s) \) such that for \( \forall N > N_0 \) the following equations hold:

\[
\mathcal{Y}_{k_0,N} = \Theta_s Z_{1,N} + \hat{\Theta}_s Z_{k_0,N}^{-} + \varepsilon_{k_0,N},
\]

\[
y_s(N+k_0) = \Theta_s z_s(N+1) + \hat{\Theta}_s z_{s-1}(N+k_0) + \varepsilon_s(N+k_0). \tag{16}
\]

Substitute eqs. (15) and (16) into (14), and for sufficiently large \( N \) we have

\[
\varepsilon_s(N+k_0) = y_s(N+k_0) - \mathcal{Y}_{k_0,N} \xi_{N+1} + \Theta_s \xi_{N+1} = e_s(N+k_0) - \frac{1}{N} \varepsilon_{k_0,N} Z_{k_0,N}^{-} T \left( \frac{1}{N} \left( Z_{k_0,N}^{-} Z_{k_0,N}^{-} \right)^{-1} z_{s-1}(N+k_0) \right) \tag{17}
\]

Note that, unlike the open-loop case [4, 5], we have

\[
\lim_{N \to \infty} \frac{1}{N} \varepsilon_{k_0,N} Z_{k_0,N}^{-} T \not= 0
\]

due to the existence of feedback, and therefore the second term on the right hand side of (17) does not converge to 0.

However, when we focus on the last block entry of the vector \( \varepsilon_s(N+k_0) \) in (14), i.e., \( \varepsilon_s(N+2s) \not= 0 \), called the residual henceforth, we have

\[
\frac{1}{N} E_{k_0,N} Z_{k_0,N}^{-} T \left( \frac{1}{N} Z_{k_0,N}^{-} Z_{k_0,N}^{-} \right)^{-1} z_{s-1}(N+k_0), \tag{18}
\]
and due to the uncorrelation\footnote{Note that the word "uncorrelation" is used here with a slight abuse of terminology [24, 25].} between the relatively future innovation and the relatively past joint input-output, that is,

\[
\lim_{N \to \infty} \frac{1}{N} \mathbf{E}_{s,N} z_{k_0,N}^T T = 0,
\]

the second term on the right hand side of (18) vanishes as \( N \) tends to infinity. Hence, the residual \( \tilde{e}_{N+2s|N+2s-1} \) can be thought to be asymptotically equivalent to the innovation \( e_{N+2s} \).

In summary of this section, we have the following theorem, which is one of the main contribution of this paper:

**Theorem 1** For \( \forall \delta > 0 \),

\[
\lim_{N \to \infty} \mathbf{P} \left( \| \tilde{e}_{N+2s} - \tilde{e}_{N+2s|N+2s-1} \| \geq \delta \right) = 0,
\]

where \( \mathbf{P}(\Omega) \) denotes the probability of an event \( \Omega \) occurring.

**Proof:** From (18) and assumption 5,

\[
\lim_{N \to \infty} \mathbf{P} \left( \| \tilde{e}_{N+2s} - \tilde{e}_{N+2s|N+2s-1} \| \geq \delta \right)
= \lim_{N \to \infty} \mathbf{P} \left( \left\| \frac{1}{N} \mathbf{E}_{s,N} z_{k_0,N} - z_{s-1}(N + k_0) \right\| \geq \delta \right)
\leq \lim_{N \to \infty} \mathbf{P} \left( \left\| \frac{1}{N} \mathbf{E}_{s,N} z_{k_0,N} - z_{s-1}(N + k_0) \right\| \right)
= 0
\]

Thus, the proof is completed. \( \square \)

### 4 Online change detection of a system in closed-loop

#### 4.1 Problem formulation for online change detection

In this section, let us assume that a system under surveillance can be described by an innovation-type model with an abrupt change in the coefficients at an unknown time instant \( k_c \) as follows:

\[
x_{k+1} = (A + A\Delta)x_k + (B + B\Delta)u_k + Ke_k, \quad (19a)
\]

\[
y_k = (C + C\Delta)x_k + e_k, \quad (19b)
\]

\[
\begin{cases}
(A\Delta, B\Delta, C\Delta) \equiv (0, 0, 0), & k < k_c, \\
(A\Delta, B\Delta, C\Delta) \neq (0, 0, 0), & k \geq k_c,
\end{cases}
\]

where \( x_k \in \mathbb{R}^n, y_k \in \mathbb{R}^l, u_k \in \mathbb{R}^m \). Note that \( A\Delta, B\Delta \) and \( C\Delta \) correspond to a change in the dynamics, an actuator fault and a sensor fault, respectively. Define \( A_\Delta \equiv A_\Delta - KC\Delta \). The innovation \( e_k \) is a zero-mean independently identically distributed Gaussian noise with its covariance matrix equal to \( \sigma_e^2 I \), that is, \( \mathbf{E} [e_k] = 0 \) and \( \mathbf{E} [e_k e_k^T] = \sigma_e^2 \mathbf{I} \), where \( \mathbf{E} \) denotes expectation and \( \sigma_e \) Kronecker’s delta. The coefficient matrices are assumed to be real matrices with appropriate dimensions. Assumptions 1 and 2 are assumed. Feedback from the output \( y_k \) to the input \( u_k \) is allowed. The signals are assumed to be bounded, pseudo-stationary [16] and satisfies the relevant persistence of excitation condition given in assumption 3.

The goal of this section is to develop an online change detection method based on whiteness testing to detect, as soon as possible, the abrupt change in the system which occurs at unknown time instant \( k_c \).

#### 4.2 Discussion of the test statistics

As is discussed in section 3, the residual \( \tilde{e}_{N+2s} \) is a promising candidate for a test signal used for whiteness testing. Taking account of similarity in derivation from recursive subspace model identification between the open-loop [4, 5] and closed-loop cases, the proposed residual is also expected to be superior in accuracy to conventional test signals generated by, e.g., well-known RLS and RLM algorithms. To convince ourselves of its usefulness, it is investigated here what comes to the residual when an abrupt change in the system occurs.

Suppose \( k_c = N + 2s \), in other words, an abrupt change in the system under surveillance occurs between \( N + 2s - 1 \) and \( N + 2s \). Assume that \( N \) is sufficiently large, and that the signal to noise ratio, \( \{ (u_k, y_k) \} \) to \( \{ e_k \} \) is significant. Then, from eqs. (18) and (19), we have

\[
\tilde{e}_{N+2s|N+2s-1} \approx \tilde{e}_{N+2s} + \Delta_1 A^{s-2} x_{N+k_0} + \Delta_2 z_{s-1}(N + k_0)
\]

where \( \Delta_1 \equiv C\Delta A + C\Delta A\Delta + C\Delta A_\Delta \) and \( \Delta_2 \equiv \Delta_3 A^{s-K} z_{s-1}(N + k_0) \).

Due to the abrupt change at \( k_c = N + 2s \), the presence of the second and third terms of (20) leads to the covariance of \( \tilde{e}_{N+k_c} \), i.e., \( \sigma^2_\epsilon I \). Similarly, for \( \tilde{e}_{N+i} \), with \( i \geq 1 \), the asymptotic equivalence between \( \tilde{e}_{N+i|k_c-i+1} \) and \( \tilde{e}_{N+i} \) does not hold any more after the abrupt change occurs at \( k_c \).

The discussion as above supports the usefulness of the proposed residual as a test signal, with high accuracy, for a whiteness test.

#### 4.3 Likelihood ratio test

In this subsection, making use of the real-time observation of the residual, denoted briefly by \( \tilde{e} := \tilde{e}_{N+k_c} \)
henceforth, an on-line change detection scheme is developed. The key point of the change detection scheme presented here is to decide whether or not the covariance of the residual \( \{\varepsilon_k\} \) is significantly larger than that of \( \{\varepsilon_k\} \) (i.e., \( \sigma_0^2 I \)). In other words, we decide that the system is in fact unchanged if no significant increase of the covariance can be observed. Note that, thanks to the similarity between this case and the open-loop case [4, 5], the proposed change detection scheme is quite similar to that presented in these references.

Now, suppose
\[
\varepsilon_1, \ldots, \varepsilon_M \sim N(\mu, \sigma^2 I), \text{ i.i.d.} \quad (21)
\]
The change detection scheme tests between the two following hypotheses:
\[
\{ H_0 : \theta \in S_0 := \{(0, \sigma_0^2)\}, \quad H_1 : \theta \in S_1 := \{\mu, \sigma^2\} \mid \sigma^2 > \sigma_0^2, \mu \text{ is arbitrary.}\}
\]
where, according to (21), the parameter \( \theta \) is defined as \( \theta = (\mu, \sigma^2) \). Note that both the mean and covariance are known under \( H_0 \) while they are unknown under \( H_1 \). The logarithm of the likelihood ratio for the sequence \( \{\varepsilon_k\}_{k=1}^M \) is given by
\[
lmax = \ln \max_{\theta \in S_1} \frac{1}{\sigma_0} \exp \left\{ -\frac{1}{2\sigma_0^2} \sum_{k=1}^M (\varepsilon_k - \mu)^T (\varepsilon_k - \mu) \right\}
\]
\[
= \frac{Ml\hat{\sigma}^2}{2\sigma_0^2} - \frac{ML}{2} \left( 1 + \ln \frac{\hat{\sigma}^2 - \hat{\mu}^T \hat{\mu}}{\sigma_0^2} \right), \quad (22)
\]
where
\[
\sigma^2 := \frac{1}{M} \sum_{k=1}^M \varepsilon_k^T \varepsilon_k, \quad \hat{\mu} := \frac{1}{M} \sum_{k=1}^M \varepsilon_k. \quad (23)
\]

Note that the maximization in (22) is accomplished when \( \mu = \hat{\mu}, \sigma^2 = \frac{1}{M} \sum_{k=1}^M (\varepsilon_k - \hat{\mu})^T (\varepsilon_k - \hat{\mu}) \). Using an appropriate threshold \( c > 0 \), the decision rule is given by
\[
lmax > c \implies \text{reject } H_0. \quad (24)
\]
We now derive the threshold \( c \). From (23), the conditional probability under the null hypothesis, \( P_{(0,\sigma_0^2)}(lmax > c) \), is given by
\[
P_{(0,\sigma_0^2)}(lmax > c) = P_{(0,\sigma_0^2)} \left( \frac{Ml\hat{\sigma}^2}{2\sigma_0^2} - \frac{ML}{2} \ln \frac{\hat{\sigma}^2 - \hat{\mu}^T \hat{\mu}}{\sigma_0^2} > c \right)
\]
\[
\approx P_{(0,\sigma_0^2)} \left( \frac{Ml\hat{\sigma}^2}{2\sigma_0^2} - \frac{ML}{2} > c \right)
\]
\[
= P_{(0,\sigma_0^2)} \left( \frac{Ml\hat{\sigma}^2}{2\sigma_0^2} > 2 \left( c + \frac{ML}{2} \right) \right). \quad (25)
\]
Note that \( Ml\hat{\sigma}^2/\sigma_0^2 \) has a \( \chi^2 \) distribution with \( (ML - 1) \) degrees of freedom. Therefore, if \( \chi_0^2(ML - 1) \) denotes the 100(1 - \( \alpha \))-th percentile of the \( \chi^2 \) distribution with \( (ML - 1) \) degrees of freedom, from
\[
2 \left( c + \frac{ML}{2} \right) = \chi_0^2(ML - 1),
\]
the threshold at the \( \alpha \) significance level is given by
\[
c = \frac{1}{2} \left( \chi_0^2(ML - 1) - ML \right). \quad (26)
\]

In summary, from (22) and (25), the likelihood ratio test with respect to \( H_0 \) vs. \( H_1 \) is obtained by
\[
\frac{\sigma^2}{\sigma_0^2} - \ln \frac{\sigma^2 - \hat{\mu}^T \hat{\mu}}{\sigma_0^2} > \frac{\chi_0^2(ML - 1)}{ML} \implies \text{reject } H_0. \quad (27)
\]

4.4 Online change detection

In this subsection, an online change detection method is presented. GMA-based recursive algorithms for updating \( \hat{\sigma}^2 \) and \( \hat{\mu} \) in (23) are derived and they are applied to the likelihood ratio test (26). GMA is the abbreviation for Geometric Moving Average ([13]).

Replace \( \hat{\sigma}^2 \) and \( \hat{\mu} \) in (26), respectively, by
\[
s_N = \frac{1}{l} \sum_{i=0}^{\infty} \gamma_i \varepsilon_{N-i}^T \varepsilon_{N-i}, \quad m_N = \sum_{i=0}^{\infty} \gamma_i \varepsilon_{N-i}, \quad (28)
\]
where \( \gamma_i := (1 - \lambda)^i \), and \( 0 < \lambda \leq 1 \) is an exponential forgetting factor. Note that effective data-length can be approximated by \( M \approx 1/\lambda \). Then, recursive algorithms for updating \( s_N \) and \( m_N \) in (27) are, respectively, given by
\[
s_{N+1} = (1 - \lambda)s_N + \lambda \varepsilon_{N+1}^T \varepsilon_{N+1}, \quad (29)
\]
\[
m_{N+1} = (1 - \lambda)m_N + \lambda \varepsilon_{N+1}. \quad (30)
\]

The proposed online change detection method can be summarized as follows ([4]):

Proposition 2 (Online change detection method)

Suppose \( s_N \) and \( m_N \) in (28) are given. Then, \( s_{N+1} \) and \( m_{N+1} \) are computed according to (28) and (29), respectively, and then, the following likelihood ratio test is carried out:
\[
\frac{s_{N+1}}{\sigma_0^2} - \ln \frac{m_{N+1}^T m_{N+1}}{\sigma_0^2} > \frac{\chi_0^2(\lambda-1)}{\upsilon_K} \implies \text{reject } H_0. \quad (30)
\]
The left hand side of inequality (30) is called the decision function.

5 Numerical simulation

Consider a closed-loop system composed of a system \( P(s) \) to be observed and a stabilizing feedback controller \( K(s) \), depicted in Fig. 1, where
\[
P(s) = \frac{0.05}{s^2 + cs + 0.48}, \quad K(s) = \frac{3}{s + 2}. \quad (31)
\]
Assume that, at an unknown time instant, the parameter $c$ changes abruptly as follows:

$$c = \begin{cases} 
0.13 & \text{before the change,} \\
0.21 & \text{after the change.}
\end{cases}$$

Note that the closed-loop system is internally stable both before and after the change. The closed-loop system has three external inputs, namely, the persistently exciting signal $r$, the disturbance $v$ and the reference. The reference is set to 0. The signals $u$ and $y$ are available for change detection. Assume that the disturbance $v$ is an unmeasurable colored signal.

In the numerical simulation, $P(s)$ and $K(s)$ are discretized using the bilinear transform with the unit sampling period. Fig. 2 illustrates the Bode diagrams of $P(s)$ before and after the abrupt change. Let the abrupt change occur at the sampling instant $k_e = 2000$. Note that $k_e$ is not known nor available for change detection.

The persistently exciting signal $r$ is a zero mean Gaussian signal with the variance 0.16. The disturbance $v$ is generated as an output of the filter

$$1 - 1.56z^{-1} + 1.045z^{-2} - 0.3338z^{-3} \over 1 + 0.85z^{-1} - 0.31z^{-2} - 0.6675z^{-3},$$

whose input is a zero mean Gaussian signal with the variance 0.025². The profiles of $(u, r)$ and $(y, b)$ are depicted respectively in Figs. 3 and 4. Note that the signal $y$ to noise $v$ ratio (SNR) is about 0.4dB.

The problem considered in this simulation is to detect the abrupt change in the system $P(s)$, which occurs at $k_e = 2000$, as early as possible using the proposed online change detection method with observations of $u$ and $y$. The block size of block Hankel matrices, $s$, is set at 10. Using the first 500 samples, the initial values of the recursive algorithm in Proposition 1 is calculated. The recursive algorithm runs from the sampling instant $k = 501$ to generate the signal for the hypothesis testing $c$. The exponential forgetting factor $\lambda$ in (28) and (29) is set to 0.2. The level of significance is set to $\alpha = 0.1\%$.

The result of the numerical simulation is illustrated in Fig. 5. The proposed change detection method makes an alarm at $k = 2005$, which is 5 steps after the instance where the change occurs. The proposed method achieves early detection of the abrupt change in the system under surveillance. As illustrated in Fig. 4, note that SNR, $y$ to $v$, is about 0.4dB, that is, the noise level is almost comparable to the output signal level.

6 Conclusion

In this paper, we have developed an online change detection scheme based on whiteness testing, which is used for detection of changes in a system under surveillance in closed-loop. A recursive algorithm of closed loop subspace model identification has been presented, which is based on the matrix inversion lemma. The relation between the proposed recursive algorithm and PBSID has been investigated from the standpoint of a matrix-valued least squares problem. A test signal online generated by the recursive algorithm is studied and its asymptotic whiteness has explicitly been proved. The introduction of statistical hypotheses testing has brought us a statistically solid decision rule based on a significant level to be designed.

References


Fig. 3: The input $u$ and the external persistently exciting signal $r$.

Fig. 4: The output $y$ and the disturbance $v$. The signal to noise ratio, $y$ to $v$, is 0.4dB.

Fig. 5: Decision function (solid line) and a threshold (dashed line) in (30).


For three tail matrices, $F$, $G$ and $H$, the oblique projection of $F$ onto the space generated by the rows of $G$ along the space generated by the rows of $H$, denoted by $\hat{E}_{[\parallel H][F|G]}$, is defined as follows:

$$\hat{E}_{[\parallel H][F|G]} = \hat{\Sigma}_{FG|H} \hat{\Sigma}_{GG|H}^{-1} G,$$

where $\hat{\Sigma}_{FG} = \frac{1}{d} FG^T$, and $\hat{\Sigma}_{FH}$ and $\hat{\Sigma}_{GH}$ are defined in the same manner as $\hat{\Sigma}_{FG}$ is defined. We define

$$\hat{\Sigma}_{FG|H} = \hat{\Sigma}_{FG} - \hat{\Sigma}_{FH} \hat{\Sigma}_{HH}^{-1} \hat{\Sigma}_{HG},$$
$$\hat{\Sigma}_{GG|H} = \hat{\Sigma}_{GG} - \hat{\Sigma}_{GH} \hat{\Sigma}_{HH}^{-1} \hat{\Sigma}_{HG}.$$