Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

A Direct Approach to Fault Detection in Non-uniformly Sampled Systems

Iman Izadi * **Sirish L. Shah** * **Tongwen Chen** **

* *Department of Chemical & Materials Engineering, University of Alberta, Edmonton, Alberta, Canada, T6G 2G6 (e-mail: iman@ualberta.ca, sirish.shah@ualberta.ca)*
** *Department of Electrical & Computer Engineering, University of Alberta, Edmonton, Alberta, Canada, T6G 2V4 (e-mail: tchen@ece.ualberta.ca)*

Abstract: Non-uniformly sampled systems are widely found in industry. In these systems the process output is sampled and the control input is generated at non-uniformly distributed time instants. In this paper, an optimal residual generator is developed for fault detection in non-uniformly sampled systems. In the direct approach used here, the intersample behavior of fault and disturbance is captured by introducing operators that map continuous-time signals to discrete-time signals. No periodicity assumption is made for the sampling instants.

Keywords: Non-uniformly sampled systems; Sampled-data systems; Fault detection and diagnosis; Parity space; Optimal design.

1. INTRODUCTION

Modern industrial control systems are widely exposed to faults which can cause undesirable performance, instability, total failure of the system and even dangerous situations. In order to maintain quality, reliability and safety, faults should be promptly detected and identified so that appropriate remedies can be applied. The problem of fault detection and isolation (FDI) has been widely studied in the past decades and numerous design methods are available [Chen et al., 1999, Isermann, 2006].

Sampled-data systems, on the other hand, are extensively used and accepted in industry, due to numerous advantages of digital technology. In a sampled-data system, the actual process which is often continuous-time, is connected to the computer network through analog-to-digital (A/D) and digital-to-analog (D/A) converters. Control and fault detection algorithms are then implemented by the computer. Thus a sampled-data system utilizes both continuous-time and discrete-time systems/signals. A typical sampled-data process with digitally implemented controller and FDI system is illustrated in Fig. 1.

In conventional sampled-data systems, it is assumed that each process variable is sampled at a constant rate and each control signal is generated at a constant rate. The sampling rates of different A/D and D/A converters may be equal (single-rate system) or different (multirate systems). However, in many practical situations, for instance in chemical processes, this is not often the case. Frequently, process outputs are sampled at non-uniformly spaced time instants. Control inputs may also be generated at non-uniformly spaced times. This could happen due to a number of reasons, including unpredictable delays in sensors and laboratory analysis and the nature of the network that connects the elements of the control system. Also in many

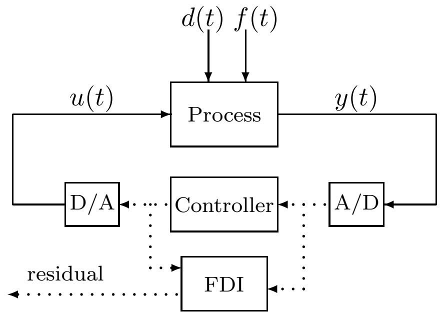

Fig. 1. FDI in a sampled-data scheme

typical applications, the control algorithm is implemented on the same distributed computer system that monitors the process and manages other aspects of the plant. In such task-sharing situations, it is more reasonable and cost-effective to allow non-uniform sampling. Moreover, it has been shown that non-uniform sampling can introduce some advantages in controlling the process [Kreisselmeier, 1999, Sheng et al., 2002].

In this paper, we develop an optimal FDI methodology for non-uniformly sampled systems based on the parity space approach. The method presented in this paper is distinctive from previous works on non-uniformly sampled systems in two ways:

- there is no need for sampling instants to follow a periodic pattern (no periodicity assumption); and
- the fault and disturbance signals can vary arbitrarily over time (no piecewise constant assumption).

A number of results are available on control, identification and fault detection of non-uniformly sampled systems [Albertos et al., 1999, Li et al., 2006, 2008, Sheng et al., 2002]. In all of these works, a non-uniform yet periodic sampling pattern was considered (i.e., the sampling instants are non-uniformly distributed in a window of time, and this window is periodically repeated). This assumption restricts the applications of the proposed methods. In this paper, we assume that the output sampling and control updating times can be arbitrarily distributed over time. Due to this non-periodicity assumption, the lifting technique which is usually used for multirate and non-uniformly sampled systems, can not be used. Here, we use a direct time-varying formulation to approach non-uniformly sampled systems.

In addition, to design the optimal FDI scheme in this paper, we use the direct approach [Chen et al., 1995, Izadi et al., 2005, 2007, Zhang et al., 2001]. In the direct approach, it is assumed that the fault and disturbance inputs can take any value at any instant of time. As a result, operators should be used to capture the effect of continuous-time fault and disturbance on discrete-time residual, and the optimization problem is stated in terms of operator norms. On the contrary, in the indirect approach, for instance in Li et al. [2006, 2008], usually the invalid but convenient assumption is made that fault and disturbance signals are constant over the sampling intervals. This assumption is restrictive and will result in an approximate residual and likely later fault detection, specifically when the sampling intervals are relatively large.

2. PRELIMINARIES

2.1 Parity-space Approach

The parity space approach was originally introduced by Chow and Willsky [1984] for discrete-time systems. Consider the following system

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) + Ed(k) + Ff(k) \\ y(k) = Cx(k) \end{cases}$$

where $x(k) \in \mathbb{R}^{n_x}$ is the state vector, $u(k) \in \mathbb{R}^{n_u}$ the vector of control input, $y(k) \in \mathbb{R}^{n_y}$ the vector of process output, $d(k) \in \mathbb{R}^{n_d}$ the vector of unknown inputs (e.g., disturbance, noise, model mismatch, etc.) and $f(k) \in \mathbb{R}^{n_f}$ the vector of faults to be detected. A, B, C, E, and F are known matrices of appropriate dimensions.

For a fixed number s, referred to as the order of parity relation, define $y_s(k)$ as

$$y_s(k) = \begin{bmatrix} y(k-s) \\ y(k-s+1) \\ \vdots \\ y(k) \end{bmatrix}_{(s+1)n_y \times 1}.$$

$u_s(k)$, $d_s(k)$ and $f_s(k)$ are also defined similarly. It can be easily shown that $y_s(k)$, $u_s(k)$, $d_s(k)$ and $f_s(k)$ are related through the following expression

$$y_s(k) = H_o x(k-s) + H_u u_s(k) + H_d d_s(k) + H_f f_s(k), \quad (1)$$

where

$$H_o = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^s \end{bmatrix}, H_u = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ CB & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ CA^{s-1}B & CA^{s-2}B & \cdots & CB & 0 \end{bmatrix}.$$

H_d and H_f are defined similar to H_u. Based on (1), a parity space residual generator can be formulated as

$$r(k) = v_s\big(y_s(k) - H_u u_s(k)\big),$$

where $r(k) \in \mathbb{R}$ is the residual. The parity vector $v_s \in \mathbb{R}^{1\times(s+1)n_y}$ is the design parameter and belongs to the parity space P_s defined by

$$P_s = \{v_s | v_s H_o = 0\}.$$

Dynamics of the residual generator is then expressed by

$$r(k) = v_s\big(H_d d_s(k) + H_f f_s(k)\big), \quad v_s \in P_s.$$

If the residual $r(k)$ can not be perfectly decoupled from the unknown input $d(k)$, the effect of $d(k)$ on $r(k)$ will be minimized by optimizing a performance index. A common choice of performance index for optimization is [Chen et al., 1999]

$$J = \frac{\|v_s H_d\|_2^2}{\|v_s H_f\|_2^2} = \frac{v_s H_d H_d^T v_s^T}{v_s H_f H_f^T v_s^T}.$$

The numerator and denominator of J reflect the effect of unknown input $d(k)$ and fault $f(k)$ on the residual. Therefore, by minimizing J a compromise is made between sensitivity to the fault and robustness to the disturbance. Solution of this optimization problem is well-known in the literature [Chen et al., 1999].

2.2 Operator Norm and Adjoint Operator

Consider Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$ with inner products $\langle x_1, x_2\rangle_{\mathcal{X}}$, $x_1, x_2 \in \mathcal{X}$ and $\langle y_1, y_2\rangle_{\mathcal{Y}}$, $y_1, y_2 \in \mathcal{Y}$, respectively. $\mathcal{X}$ and $\mathcal{Y}$ are not necessarily the same space, and even if they are, the inner products can be different. The norms of members of $\mathcal{X}$ and $\mathcal{Y}$ are defined using the corresponding inner products as $\|x\|_{\mathcal{X}}^2 = \langle x, x\rangle_{\mathcal{X}}$, $x \in \mathcal{X}$ and $\|y\|_{\mathcal{Y}}^2 = \langle y, y\rangle_{\mathcal{Y}}$, $y \in \mathcal{Y}$. Also assume that $T : \mathcal{X} \to \mathcal{Y}$ is a bounded operator that maps $\mathcal{X}$ to $\mathcal{Y}$. The adjoint of T, denoted by T^*, is the unique bounded operator mapping $\mathcal{Y}$ to $\mathcal{X}$ that satisfies [Chen et al., 1995]

$$\langle Tx, y\rangle_{\mathcal{Y}} = \langle x, T^* y\rangle_{\mathcal{X}}, \quad x \in \mathcal{X}, \ y \in \mathcal{Y}.$$

It can be easily shown that the adjoint of a constant matrix is its transpose.

The induced norm of the operator T is defined by

$$\|T\| = \sup_{\|x\|_{\mathcal{X}} \le 1} \|Tx\|_{\mathcal{Y}}.$$

It is a well known fact that [Chen et al., 1995]

$$\|T\|^2 = \|T^*\|^2 = \|T^*T\| = \|TT^*\|. \quad (2)$$

2.3 Process Description

Consider an LTI, strictly proper, continuous-time process with the following state-space representation

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + Ed(t) + Ff(t) \\ y(t) = Cx(t) \end{cases} \quad (3)$$

where $x(t) \in \mathbb{R}^{n_x}$ is the state vector, $u(t) \in \mathbb{R}^{n_u}$ the known vector of control input, $y(t) \in \mathbb{R}^{n_y}$ the vector of process output, $d(t) \in \mathbb{R}^{n_d}$ the vector of unknown input (to represent disturbance, noise, model mismatch and other uncertainties) and $f(t) \in \mathbb{R}^{n_f}$ the vector of fault to be detected. A, B, C, E and F are known matrices of appropriate dimensions. The assumption of strictly properness is standard in the sampled-data literature and necessary for boundedness of the sampling operator. In practice, because of antialiasing filters that are used before sampling, the systems are always strictly proper. Notice that here, again due to antialiasing filters, $f(t)$ can represent both actuator and sensor faults.

In general, different input/output channels of this process can be generated/sampled at non-uniformly spaced time instants (non-uniformly sampled multirate systems). However, for simplicity we assume that all the input/output channels are generated/sampled synchronously at the same time instants (the approach can be applied to the general multirate case with little modification). Let $T = \{t_0, t_1, t_2, \cdots\}$ be the set of time instants when the output is sampled (or the input is updated). Let $\ell_T(\mathbb{Z})$ be the vector space of all discrete-time signals corresponding to the time instants in T. Notice that the discrete-time signals in $\ell_T(\mathbb{Z})$ have no practical meaning unless the corresponding time instants, given by T, are known. Let $\mathcal{L}(\mathbb{R})$ be the vector space of all continuous-time signals.

The non-uniform digital-to-analog (D/A) converter is modeled by non-uniform (zero-order) hold operator $H_T : \ell_T(\mathbb{Z}) \to \mathcal{L}(\mathbb{R})$ defined as

$$u(t) = H_T \upsilon_T(k) = \upsilon_T(k), \quad t_k \le t < t_{k+1}.$$

The non-uniform analog-to-digital (A/D) converter is also modeled by non-uniform sampling operator $S_T : \mathcal{L}(\mathbb{R}) \to \ell_T(\mathbb{Z})$ defined as $\psi_T(k) = S_T y(t) = y(t_k)$. Here $\upsilon_T(k)$ and $\psi_T(k)$ represent the (irregular) discrete-time input and output, respectively.

The control signal $u(t)$ is the output of a hold operator, and therefore is constant over the sampling interval (i.e., it is piecewise constant). The disturbance $d(t)$ and the fault $f(t)$, on the other hand, can have arbitrary values at any time (notice that in indirect design, $d(t)$ and $f(t)$ are also assumed to be piecewise constant, which is obviously non-realistic).

3. INPUT-OUTPUT RELATION

The parity space based residual generator, discussed in Section 2, is obtained based on (1). This equation expresses how the output of the system within an interval of time ($(s+1)h$ units of time, where h is the sampling period) is related to the state of the system at the beginning of the interval and the inputs of the system (including controlled input, disturbance and fault) during the interval. Likewise, the first step in constructing a residual generator for non-uniformly sampled systems is to derive an expression similar to (1). For this purpose, at each sampling instant t_k, we select a time frame that contains $s+1$ samples of the output $\big(\psi_T(k-s)$ to $\psi_T(k)\big)$, hence the time frame is $[t_{k-s}, t_k)$. Notice that due to the non-uniform sampling pattern, the actual length of the time frame is different at every instant. Define

$$\psi_{T,s}(k) = \begin{bmatrix} \psi_T(k-s) \\ \psi_T(k-s+1) \\ \vdots \\ \psi_T(k) \end{bmatrix}_{(s+1)n_y \times 1}$$

and similarly $\upsilon_{T,s}(k)$. The objective is to express $\psi_{T,s}(k)$ in terms of the state of the system at the beginning of the time frame $\big(x(t_{k-s})\big)$, the controlled input within the time frame $\big(\upsilon_{T,s}(k)\big)$ and the uncontrolled inputs within the time frame $\big(d(t)$ and $f(t)$ for $t_{k-s} \le t < t_k\big)$.

3.1 Case (i): Controlled Input

In the first case, assume that there is no uncontrolled input in the system, i.e., $d(t) = 0$ and $f(t) = 0$. It can be shown that the input-output relation in the selected time frame is given by

$$\psi_{T,s}(k) = H_{o,T}(k)x(t_{k-s}) + H_T(k)H_{B_d,T}(k)\upsilon_{T,s}(k),$$

where $H_{o,T}(k) : (s+1)n_y \times n_x$, $H_T(k) : (s+1)n_y \times (s+1)n_x$ and $H_{B_d,T}(k) : (s+1)n_x \times (s+1)n_u$ are given by

$$H_{o,T}(k) = \begin{bmatrix} C \\ CA_d(t_{k-s}, t_{k-s+1}) \\ CA_d(t_{k-s}, t_{k-s+2}) \\ \vdots \\ CA_d(t_{k-s}, t_k) \end{bmatrix}, \tag{4}$$

$$H_T(k) = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ C & \cdots & 0 & 0 \\ CA_d(t_{k-s+1}, t_{k-s+2}) & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ CA_d(t_{k-s+1}, t_k) & \cdots & C & 0 \end{bmatrix}, \tag{5}$$

$$H_{B_d,T}(k) = \begin{bmatrix} B_d(t_{k-s}, t_{k-s+1}) & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & B_d(t_k, t_{k+1}) \end{bmatrix}.$$

Here, $A_d(\tau_1, \tau_2)$ and $B_d(\tau_1, \tau_2)$ for $\tau_1 \le \tau_2$ are defined as

$$A_d(\tau_1, \tau_2) = e^{(\tau_2 - \tau_1)A},$$

$$B_d(\tau_1, \tau_2) = \int_{\tau_1}^{\tau_2} e^{(\tau_2 - \tau)A} d\tau B = \int_0^{\tau_2 - \tau_1} e^{\tau A} d\tau B.$$

3.2 Case (ii): Uncontrolled Input

In the second case, assume that the system is only driven by uncontrolled inputs, i.e., $u(t) = 0$. Also for simplicity assume that $f(t) = 0$. It is well known that, for any two times $t_1 \le t_2$,

$$x(t_2) = e^{(t_2 - t_1)A}x(t_1) + \int_{t_1}^{t_2} e^{(t_2 - \tau)A} E d(\tau) d\tau.$$

By substituting $t_1 = t_{k-s}$ and $t_2 = t_{k-s+i}$, $i = 0, 1, \cdots s$ we get

$$x(t_{k-s+i}) = e^{(t_{k-s+i} - t_{k-s})A}x(t_{k-s}) + \int_{t_{k-s}}^{t_{k-s+i}} e^{(t_{k-s+i} - \tau)A} E d(\tau) d\tau. \tag{6}$$

Now we can rewrite the last term as

$$\int_{t_{k-s}}^{t_{k-s+i}} e^{(t_{k-s+i} - \tau)A} E d(\tau) d\tau = \sum_{m=1}^{i} \int_{t_{k-s+m-1}}^{t_{k-s+m}} e^{(t_{k-s+i} - \tau)A} E d(\tau) d\tau = \sum_{m=1}^{i} e^{(t_{k-s+i} - t_{k-s+m})A} \times \int_{t_{k-s+m-1}}^{t_{k-s+m}} e^{(t_{k-s+m} - \tau)A} E d(\tau) d\tau.$$

Define

$$\bar{\delta}_T(k) = \int_{t_k}^{t_{k+1}} e^{(t_{k+1} - \tau)A} E d(\tau) d\tau.$$

Then (6) can be simplified to

$$x(t_{k-s+i}) = A_d(t_{k-s}, t_{k-s+i})x(t_{k-s}) + \sum_{m=1}^{i} A_d(t_{k-s+m}, t_{k-s+i})\bar{\delta}_T(k-s+m-1).$$

Using the output equation in (3) we have $\psi_T(k-s+i) = Cx(t_{k-s+i})$ and therefore

$$\psi_T(k-s+i) = CA_d(t_{k-s}, t_{k-s+i})x(t_{k-s}) + \sum_{m=1}^{i} CA_d(t_{k-s+m}, t_{k-s+i})\bar{\delta}_T(k-s+m-1)$$

By changing i from 0 to s, and stacking all the equations, the input-output relation becomes

$$\psi_{T,s}(k) = H_{o,T}(k)x(t_{k-s}) + H_T(k)\bar{\delta}_{T,s}(k),$$

where

$$\bar{\delta}_{T,s}(k) = \begin{bmatrix} \bar{\delta}_T(k-s) \\ \bar{\delta}_T(k-s+1) \\ \vdots \\ \bar{\delta}_T(k) \end{bmatrix}_{(s+1)n_x \times 1}.$$

$H_{o,T}(k) : (s+1)n_y \times n_x$ and $H_T(k) : (s+1)n_y \times (s+1)n_x$ are the same as those obtained for controlled input and are given in (4) and (5) respectively.

Now define the operator $\Gamma_{E,T} : \mathcal{K}^{n_d} \to \mathbb{R}^{(s+1)n_x}$ as

$$\Gamma_{E,T}d = \begin{bmatrix} \int_{t_{k-s}}^{t_{k-s+1}} e^{(t_{k-s+1}-\tau)A} Ed_0(\tau)\, d\tau \\ \int_{t_{k-s+1}}^{t_{k-s+2}} e^{(t_{k-s+2}-\tau)A} Ed_1(\tau)\, d\tau \\ \vdots \\ \int_{t_k}^{t_{k+1}} e^{(t_{k+1}-\tau)A} Ed_s(\tau)\, d\tau \end{bmatrix},$$

where for $i = 0, 1, \cdots, s$,

$$d_i(t) = \begin{cases} d(t) & t_{k-s+i} \le t < t_{k-s+i+1}; \\ \text{undefined} & \text{elsewhere.} \end{cases}$$

$\mathcal{K}^{n_d}$ is the space of all vector valued continuous-time signals in the interval $[t_{k-s}, t_{k+1})$ with finite norm. In other words $\mathcal{K}^{n_d} = \mathcal{L}_2([t_k, t_{k+1}), \mathbb{R}^{n_d})$. For each k, the operator $\Gamma_{E,T}$ maps a continuous-time signal in $[t_{k-s}, t_{k+1})$ to a vector that is interpreted as a discrete-time signal. It can be easily observed that

$$\bar{\delta}_{T,s}(k) = \Gamma_{E,T}d(t).$$

Therefore, the input-output relation can be written as

$$\psi_{T,s}(k) = H_{o,T}(k)x(t_{k-s}) + H_T(k)\Gamma_{E,T}d(t).$$

This equation shows how continuous-time input $d(t)$ during a specific frame of time affects the discrete-time output $\psi_{T,s}(k)$.

4. OPTIMAL RESIDUAL GENERATION

Consider the LTI continuous-time process in (3). Based on the results of Section 3, when both the controlled ($u(t)$) and uncontrolled ($d(t)$ and $f(t)$) inputs are driving the process, the input-output relation in time frame $[t_{k-s}, t_{k+1}]$ is given by

$$\psi_{T,s}(k) = H_{o,T}(k)x(t_{k-s}) + H_T(k)H_{B_d,T}(k)\upsilon_{T,s}(k) + H_T(k)\Gamma_{E,T}d(t) + H_T(k)\Gamma_{F,T}f(t). \quad (7)$$

The operator $\Gamma_{F,T} : \mathcal{K}^{n_f} \to \mathbb{R}^{(s+1)n_x}$ maps continuous-time signal $f(t)$ to discrete-time signals and is defined similar to $\Gamma_{E,T}$.

Based on (7), a parity space residual generator for the non-uniformly sampled system is formulated as

$$r(k) = v_s(k)\big(\psi_{T,s}(k) - H_T(k)H_{B_d,T}(k)\upsilon_{T,s}(k)\big). \quad (8)$$

Here, $r(k) \in \mathbb{R}$ is the residual and s is the order of parity relation. The parity vector $v_s(k) \in \mathbb{R}^{1\times(s+1)n_y}$ is the design parameter. Since the non-uniformly sampled system described above is inherently time-varying, the residual generator should also be time-varying. That's why the parity vector is a function of k and should be calculated at each iteration. The parity vector $v_s(k)$ belongs to the parity space $P_s(k)$ given by

$$P_s(k) = \{v_s(k) | v_s(k)H_{o,T}(k) = 0\}.$$

Dynamics of the discrete-time residual with respect to continuous-time inputs $d(t)$ and $f(t)$ is then expressed by

$$r(k) = v_s(k)H_T(k)\big(\Gamma_{E,T}d(t) + \Gamma_{F,T}f(t)\big).$$

The parity vector $v_s(k)$ is designed to ensure robustness of the residual generator to the unknown input $d(t)$, while keeping it sensitive with respect to the fault $f(t)$.

If there exists a parity vector $v_s(k) \in P_s(k)$ such that

$$v_s(k)H_T(k)\Gamma_{E,T} \equiv 0,$$
$$v_s(k)H_T(k)\Gamma_{F,T} \ne 0,$$

then the unknown input $d(t)$ has no effect on the residual and perfect disturbance decoupling is achieved. Otherwise, the parity vector is designed by optimizing a performance index to minimize the effect of $d(t)$ on $r(k)$. Inspired by the LTI case, a common choice of performance index for optimization is

$$J_J(k) = \frac{\|v_s(k)H_T(k)\Gamma_{E,T}\|^2}{\|v_s(k)H_T(k)\Gamma_{F,T}\|^2}.$$

The norms here are induced operator norms.

To minimize this objective function, the first step is to calculate the norm of the operators, which in turn requires calculating the adjoint operators. Using the norm relationship in (2) and the fact that $v_s(k)$ and $H_T(k)$ are real matrices (hence their adjoints are their transposes), the performance index is simplified to

$$J_J(k) = \frac{\left\|v_s(k)H_T(k)\Gamma_{E,T}\Gamma^*_{E,T}H_T^T(k)v_s^T(k)\right\|}{\left\|v_s(k)H_T(k)\Gamma_{F,T}\Gamma^*_{F,T}H_T^T(k)v_s^T(k)\right\|} \quad (9)$$

Consider the operator $\Gamma_{E,T} : \mathcal{K}^{n_d} \to \mathbb{R}^{(s+1)n_x}$. The inner product in $\mathcal{K}^{n_d} = \mathcal{L}_2([t_{k-s}, t_{k+1}), \mathbb{R}^{n_d})$ is

$$\langle x, y\rangle_{\mathcal{K}^{n_d}} = \int_{t_{k-s}}^{t_{k+1}} x^T(\tau)y(\tau)\, d\tau = \sum_{i=0}^{s} \int_{t_{k-s+i}}^{t_{k-s+i+1}} x_i^T(\tau)y_i(\tau)\, d\tau, \quad x, y \in \mathcal{K}^{n_d}.$$

The inner product in $\mathbb{R}^{(s+1)n_x}$ is defined as usual: $\langle a_1, a_2\rangle_{\mathbb{R}^{(s+1)n_x}} = a_1^T a_2,\ a_1, a_2 \in \mathbb{R}^{(s+1)n_x}$. The adjoint operator $\Gamma^*_{E,T} : \mathbb{R}^{(s+1)n_x} \to \mathcal{K}^{n_d}$ is uniquely determined by

$$\langle \Gamma_{E,T}x, a\rangle_{\mathbb{R}^{(s+1)n_x}} = \langle x, \Gamma^*_{E,T}a\rangle_{\mathcal{K}^{n_d}}, \quad (10)$$

where $x \in \mathcal{K}^{n_d}$ and $a \in \mathbb{R}^{(s+1)n_x}$. Partition $a \in \mathbb{R}^{(s+1)n_x}$ into $s+1$ blocks as

$$a = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_s \end{bmatrix}.$$

Then, the left hand side of (10) becomes

$$\begin{aligned}\langle \Gamma_{E,T} x, a \rangle_{\mathbb{R}^{(s+1)n_x}} &= (\Gamma_{E,T} x)^T a \\ &= \begin{bmatrix} \int_{t_{k-s}}^{t_{k-s+1}} e^{(t_{k-s+1}-\tau)A} E x_0(\tau)\, d\tau \\ \int_{t_{k-s+1}}^{t_{k-s+2}} e^{(t_{k-s+2}-\tau)A} E x_1(\tau)\, d\tau \\ \vdots \\ \int_{t_k}^{t_{k+1}} e^{(t_{k+1}-\tau)A} E x_s(\tau)\, d\tau \end{bmatrix}^T \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_s \end{bmatrix} \\ &= \sum_{i=0}^{s} \int_{t_{k-s+i}}^{t_{k-s+i+1}} x_i^T(\tau) E^T e^{(t_{k-s+i+1}-\tau)A^T} a_i \, d\tau.\end{aligned}$$

On the other hand, the right hand side of (10) is

$$\left\langle x, \Gamma_{E,T}^* a \right\rangle_{\mathcal{K}^{n_d}} = \sum_{i=0}^{s} \int_{t_{k-s+i}}^{t_{k-s+i+1}} x_i^T(\tau) (\Gamma_{E,T}^* a)_i(\tau)\, d\tau.$$

Comparing the two equations we get

$$(\Gamma_{E,T}^* a)_i(t) = E^T e^{(t_{k-s+i+1}-t)A^T} a_i, \quad i = 0, 1, \cdots, s.$$

$(\Gamma_{E,T}^* a)(t)$ is now obtained by concatenating $(\Gamma_{E,T}^* a)_i(t)$'s.

$\Gamma_{E,T}\Gamma_{E,T}^*$ maps $\mathbb{R}^{(s+1)n_x}$ onto itself and hence is (equivalent to) an $(s+1)n_x \times (s+1)n_x$ matrix. To find this matrix we have

$$\begin{aligned}\Gamma_{E,T}\Gamma_{E,T}^* a &= \Gamma_{E,T} \left[E^T e^{(t_{k-s+i+1}-t)A^T} a_i \right]_{i=0,\cdots,s} \\ &= \begin{bmatrix} \int_{t_{k-s}}^{t_{k-s+1}} e^{(t_{k-s+1}-\tau)A} E E^T e^{(t_{k-s+1}-\tau)A^T} a_0 \, d\tau \\ \int_{t_{k-s+1}}^{t_{k-s+2}} e^{(t_{k-s+2}-\tau)A} E E^T e^{(t_{k-s+2}-\tau)A^T} a_1 \, d\tau \\ \vdots \\ \int_{t_k}^{t_{k+1}} e^{(t_{k+1}-\tau)A} E E^T e^{(t_{k+1}-\tau)A^T} a_s \, d\tau \end{bmatrix}.\end{aligned}$$

Now define $E_J(\tau_1, \tau_2)$ as a matrix of smallest dimensions that satisfies

$$\begin{aligned}E_J(\tau_1, \tau_2) E_J^T(\tau_1, \tau_2) &= \int_{\tau_1}^{\tau_2} e^{(\tau_2-\tau)A} E E^T e^{(\tau_2-\tau)A^T} \, d\tau \\ &= \int_0^{\tau_2-\tau_1} e^{\tau A} E E^T e^{\tau A^T} \, d\tau.\end{aligned}$$

Then we have,

$$\Gamma_{E,T}\Gamma_{E,T}^* a = \begin{bmatrix} E_J(t_{k-s}, t_{k-s+1}) E_J^T(t_{k-s}, t_{k-s+1}) a_0 \\ E_J(t_{k-s+1}, t_{k-s+2}) E_J^T(t_{k-s+1}, t_{k-s+2}) a_1 \\ \vdots \\ E_J(t_k, t_{k+1}) E_J^T(t_k, t_{k+1}) a_s \end{bmatrix}$$

Now define $H_{E_J,T}(k)$ as

$$H_{E_J,T}(k) = \begin{bmatrix} E_J(t_{k-s}, t_{k-s+1}) & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & E_J(t_k, t_{k+1}) \end{bmatrix}.$$

Then

$$\Gamma_{E,T}\Gamma_{E,T}^* a = H_{E_J,T}(k) H_{E_J,T}^T(k) a,$$

which implies that

$$\Gamma_{E,T}\Gamma_{E,T}^* = H_{E_J,T}(k) H_{E_J,T}^T(k).$$

Using this expression, the performance index in (9) can be simplified to

$$\begin{aligned}J_J(k) &= \frac{v_s(k) H_T(k) H_{E_J,T}(k) H_{E_J,T}^T(k) H_T^T(k) v_s^T(k)}{v_s(k) H_T(k) H_{F_J,T}(k) H_{F_J,T}^T(k) H_T^T(k) v_s^T(k)} \\ &= \frac{\| v_s(k) H_T(k) H_{E_J,T}(k) \|_2^2}{\| v_s(k) H_T(k) H_{F_J,T}(k) \|_2^2}.\end{aligned} \tag{11}$$

$H_{F_J,T}(k)$ is defined similar to $H_{E_J,T}(k)$. The performance index in (11) is now expressed in terms of regular matrix norms.

The parity vector $v_s(k)$ is then designed by solving the optimization problem

$$\min_{v_s(k) \in P_s(k)} J_J(k).$$

Let $N_B(k)$ be the basis vector for parity space $P_s(k)$. Also let $\lambda_{\min}(k)$ and $p_{s,\min}(k)$ be the minimum generalized eigenvalue and the corresponding generalized eigenvector, satisfying

$$\begin{aligned}&p_{s,\min}(k) N_B(k) H_T(k) \Big(H_{E_J,T}(k) H_{E_J,T}^T(k) \\ &\quad - \lambda_{\min}(k) H_{F_J,T}(k) H_{F_J,T}^T(k) \Big) H_T^T(k) N_B^T(k) = 0.\end{aligned}$$

Then, similar to the LTI case [Chen et al., 1999], $v_s^*(k) = p_{s,\min}(k) N_B(k)$ is the optimal solution and $J^*(k) = \lambda_{\min}(k)$ is the optimal performance. Once the optimal parity vector $v_s^*(k)$ is designed, the residual generator in (8) can be implemented.

Notice that the residual generator (8) updates the residual at time instants $t_k,\ k = s, s+1, \cdots$. These are the instants of time when the output is measured and sampled. So, as soon as new information from the process becomes available (through measurements) the residual will be updated. Therefore, the fault can be detected at the earliest time possible.

As mentioned before, the residual generator designed for the non-uniformly sampled system is time-varying. Therefore, the related matrices $H_{o,T}(k)$, $H_T(k)$, $H_{B_d,T}(k)$, $H_{E_J,T}(k)$ and $H_{F_J,T}(k)$ should be recalculated and the optimization problem re-solved at every iteration. This is a result of the non-uniformly sampled system being inherently time-varying and unpredictable. However, if the non-uniform sampling follows a certain pattern (for instance a periodic pattern as in Albertos et al. [1999], Li et al. [2006] and Sheng et al. [2002]), then the matrices can be calculated before hand and the parity vector computed off-line. In any case, the calculations at each step, mainly simple matrix computations and an eigenvector problem, are not numerically complex. Notice that, although operator norms are used in the derivation of the solution, the obtained performance index (11) is stated in terms of regular matrix norms.

5. EXAMPLE

Consider the LTI continuous-time process in (3) with

$$A = \begin{bmatrix} -0.2 & 3 \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad E = \begin{bmatrix} 0.1 \\ 1 \end{bmatrix},$$
$$F = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [\,1\ 1\,].$$

The output is non-uniformly sampled at the time instants given by (in seconds)

$$T = \{0, 0.7, 1.4, 2.1, 3.3, 3.8, 4.4, 5.8, 6.4, 6.8, 7.7, 8.8, 9.4, 10.4, 11.0, 11.7, 12.3, 13.1, 13.9, 14.9\}$$

The control signal is also updated according to T with random numbers between -3 and 3. The disturbance $d(t)$ is white noise with variance 1 (updated every 0.1 sec) and the fault $f(t)$ is a step function, changing from 0 to 1 at 8 sec. The input and output of the system are shown in Fig. 2.

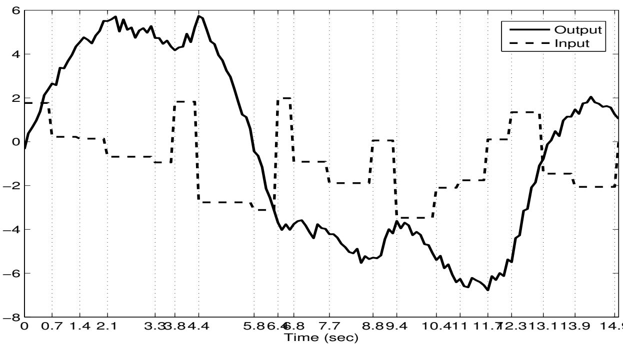

Fig. 2. System input and output

The first time that the output is sampled after the fault occurrence is at 8.8 sec, and this is the first time that fault information is available to the control/monitoring algorithm. Therefore, a well-designed residual generator should be able to reflect the fault at 8.8 sec. Choosing $s = 3$, a residual generator was designed for this non-uniformly sampled system, with the threshold set to be at 1. The result of simulation is shown in Fig. 3. As it can be seen, the proposed residual generator was able to detect the fault at the earliest possible time (8.8 sec).

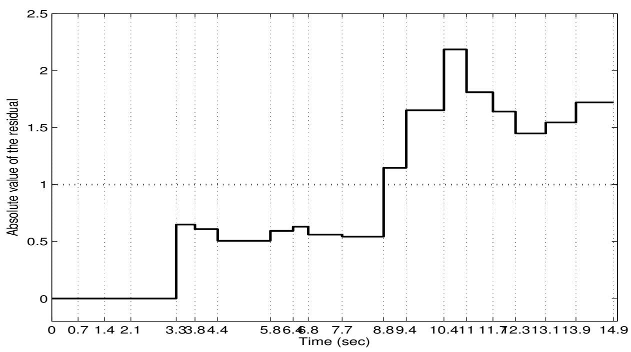

Fig. 3. Residual signal for fault detection

6. CONCLUSIONS

In this paper, we have presented a direct method to design an optimal residual generator for non-uniformly sampled systems. In this direct design, in order to avoid approximations, no assumption is made on fault and disturbance inputs and they can vary arbitrarily over time. As a result, the relationship between continuous-time fault/disturbance and discrete-time residual is expressed in terms of an operator rather than a matrix. However, it was shown that the norm of the operator is equal to the norm of a certain matrix. Therefore, the optimization problem can be converted to a regular matrix problem whose solution is known.

In the development of the residual generator, no *a priori* information is required regarding the output sampling and input updating times. In particular, there is no need for the sampling and updating times to follow a periodic pattern. The method can therefore be applied to general non-uniformly sampled systems. In the proposed method, as soon as a new measurement from the process becomes available, the residual can be updated. So any unnecessary delay in fault detection is avoided.

REFERENCES

P. Albertos and A. Crespo. Real-time control of non-uniformly sampled systems. *Control Engineering Practice*, vol. 7, pp. 445–458, 1999.

T. Chen and B. Francis, *Optimal Sampled-data Control Systems*, Springer, New York, 1995.

J. Chen and R. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Kluwer Academic Publishers, Boston, 1999.

E.Y. Chow and A.S. Willsky. Analytical redundancy and the design of robust failure detection systems. *IEEE Transactions on Automatic Control*, vol. 29, pp. 603–614, 1984.

R. Isermann, *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*, Springer, Berlin, 2006.

I. Izadi, T. Chen and Q. Zhao. Norm invariant discretization for sampled-data fault detection. *Automatica*, vol. 41, pp. 1633–1637, 2005.

I. Izadi, Q. Zhao and T. Chen. Analysis of performance criteria in sampled-data fault detection. *Systems & Control Letters*, vol. 45, pp. 320–325, 2007.

G. Kreisselmeier. On sampling without loss of observability controllability. *IEEE Transactions on Automatic Control*, vol. 44, pp. 1021-1025, 1999.

W. Li, Z. Han and S.L. Shah. Subspace identification for FDI in systems with non-uniformly sampled multirate data. *Automatica*, vol. 42, pp 619–62, 2006.

W. Li, S.L. Shah and D. Xiao. Kalman filters in non-uniformly sampled multirate systems: for FDI and beyond. *Automatica*, vol. 44, pp 199–208, 2008.

J. Sheng, T. Chen and S.L. Shah. Generalized predictive control for non-uniformly sampled systems. *Journal of Process Control*, vol. 12, pp. 875–885, 2002.

P. Zhang, S.X. Ding, G.Z. Wang and D.H. Zhou. An FDI approach for sampled-data systems. *Proceedings of the American Control Conference*, pp. 2702–2707, 2001.