

# Fault Detection and Isolation in Nonlinear Dynamic Systems: A Combined Input-Output and Local Approach

Qinghua Zhang, Michèle Basseville, Albert Benveniste

► **To cite this version:**

Qinghua Zhang, Michèle Basseville, Albert Benveniste. Fault Detection and Isolation in Nonlinear Dynamic Systems: A Combined Input-Output and Local Approach. [Research Report] RR-3120, INRIA. 1997. inria-00073570

**HAL Id: inria-00073570**

**<https://hal.inria.fr/inria-00073570>**

Submitted on 24 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Fault Detection and Isolation in Nonlinear Dynamic  
Systems: A Combined Input-Output and Local Approach***

Qinghua Zhang, Michèle Basseville and Albert Benveniste

**N° 3120**

————— THÈME 4 —————



***rapport  
de recherche***



# Fault Detection and Isolation in Nonlinear Dynamic Systems: A Combined Input-Output and Local Approach

Qinghua Zhang, Michèle Basseville and Albert Benveniste

Thème 4 — Simulation et optimisation  
de systèmes complexes  
Projet Sigma2

Rapport de recherche n° 3120—Fevrier1997—27pages

**Abstract:** The problem of fault detection and isolation (FDI) is addressed for nonlinear dynamic systems modeled by polynomial differential-algebraic equations. Ritt's algorithm is used to obtain an input-output representation of the monitored system by eliminating unknown variables. A local approach to change detection is then applied to the input-output representation for FDI residual generation and evaluation. Simulation results on a microbial growth process are provided, which illustrate the relevance of the proposed FDI method.

**Key-words:** fault detection and isolation, nonlinear dynamic systems, differential-algebraic equation, residual generation and evaluation, local approach.

*(Résumé : tsvp)*

# Détection et diagnostic de pannes dans des systèmes dynamiques non linéaires: représentation entrées-sorties et approche locale

**Résumé :** Nous considérons le problème de la détection et du diagnostic de pannes dans des systèmes dynamiques non linéaires modélisés par des équations différentielles algébriques polynômiales. L'algorithme de Ritt est utilisé pour obtenir une représentation entrée-sortie du système à surveiller par élimination des variables non observées. Une méthode locale pour la détection de changement est appliquée à la représentation entrée-sortie pour générer et évaluer des résidus. Des résultats de simulation sur un procédé micro-biologique sont présentés pour illustrer la pertinence de la méthode proposée.

**Mots-clé :** détection et diagnostic de pannes, système dynamique non linéaire, équation différentielle algébrique, génération et évaluation de résidu, méthode locale.

# 1 Introduction

The problem of fault detection and isolation (FDI) is a crucial issue for industrial system monitoring and has been theoretically and experimentally investigated with different types of approaches, as can be seen from the survey papers (Willsky, 1976; Gertler, 1988; Basseville, 1988; Frank, 1990; Isermann, 1993) and the books (Patton et al., 1989; Basseville and Nikiforov, 1993) among other references. It has been widely acknowledged that the FDI problem can be split into two steps : *generation of FDI residuals*, which are transformations of raw data with desired properties; and *residual evaluation*, namely *decision making* based on these residuals.

When restricted to the case of *additive faults* occurring on Gaussian dynamic controlled processes modeled by *linear* time invariant (LTI) *state-space systems*, the FDI problem has been addressed in the literature using either deterministic or stochastic approaches. The main deterministic methods are parity checks or analytical redundancy relations, detection filters, fault observers. The main stochastic methods are statistical tests based on likelihood ratio and minmax techniques for dealing with nuisance parameters. Thanks to the assumed additivity of the fault, likelihood methods rely upon *innovation* computations. The links between both types of approaches are discussed in (Basseville and Nikiforov, 1993; Basseville, 1997; Nikiforov et al., 1996).

As explained in (Benveniste et al., 1987; Basseville and Nikiforov, 1993), even in linear state-space systems, the case of *non additive faults* (e.g. changes in the state transition matrix) is more tricky. Monitoring innovations or observer errors is not sufficient, likelihood ratio tests may be computationally complex, and approximations such as local tests relying upon efficient scores have been shown to be relevant, including for practical applications (Basseville et al., 1987; Basseville et al., 1993). The same is true for the case of *static nonlinear systems* in which local tests have also proved useful, as shown in (Zhang et al., 1994; Basseville et al., 1994). We point out that the local test proposed in (Basseville et al., 1987) relies upon an input-output representation of the considered linear state-space system.

In this paper, we address the FDI problem in *nonlinear dynamic systems*, which, of course, has been investigated to a lesser extent; we refer to (Wünnenberg, 1990; Seliger and Frank, 1991) for the use of observer-based techniques. As an alternative, because of the relevance of the approach proposed in the linear case (Basseville et al., 1987) and thanks to recent results on input-output representation of nonlinear systems (Ljung and Glad, 1994), we propose a new approach to the monitoring of nonlinear dynamic systems. It is based on both an input-output representation and a local approach to change detection. The input-output representation is obtained through elimination of unknown variables. The local approach is a statistic tool allowing to transform very general FDI problems into an asymptotically equivalent simple problem, namely the detection of changes in the mean of a Gaussian random vector. It is based on the assumption that the faults are “small” and is thus particularly suited for the detection and isolation of incipient faults.

The paper is organized as follows. In Section 2 we formulate the FDI problem considered in this paper. In Section 3, we recall the Ljung-Glad method for deriving input-output representations of polynomial differential-algebraic systems. In Section 4, we explain how to design local tests for FDI. Then in Section 5, we provide numerical results which illustrate the relevance of the proposed method when applied to a microbial growth process. Finally, Section 6 is devoted to conclusions and questions open for future research.

## 2 Problem statement

In this section we formulate the FDI problem through a model-based approach. The considered model class and the problem of data sampling are also discussed.

### 2.1 System models

State-space models of dynamic systems are often considered in the FDI literature, as well as in the control theory literature. This is because physical laws governing engineering systems are often expressed by differential equations. For instance, continuous time state-space model of nonlinear dynamic systems can generally be written as:

$$\begin{cases} \dot{x} &= f(x, u, \theta) \\ y &= g(x, u, \theta) \end{cases} \quad (1)$$

where  $f$  and  $g$  are nonlinear functions;  $x$  is the state variable;  $u$  and  $y$  are the input and output signals;  $\dot{x}$  is the time derivative of  $x$ ;  $\theta$  collects the model parameters. All these variables and functions are possibly vector-valued. It is usually assumed that the input  $u$  and output  $y$  are known or measured (through actuators and sensors) and the state  $x$  is *not* directly measured, though this assumption is not always true in practice.

In this paper we consider systems modeled by differential-algebraic equations (DAEs). Specifically, a system is modeled by a set of differential-algebraic equations of the form

$$f_i(x, u, y, \theta, p) = 0, \quad i = 1, 2, \dots, r \quad (2)$$

where  $u$  and  $y$  are input and output signals,  $x$  represents (unknown) auxiliary variables,  $p$  is the time derivative operator,  $f_i$ 's are polynomials in  $u, y, x, p$  and parameterized by  $\theta$ . Such functions  $f_i$ 's are called *differential polynomials* in  $u, y, x, \theta$  (i.e., polynomials in  $u, y, x, \theta$  and their time derivatives). The quantities  $u, y, x$  and  $\theta$  can all be vector-valued. We assume that the input  $u$  and output  $y$  are known or measured (otherwise the unknown ones could be put into  $x$ ).

State-space model (1) is a particular instance of DAE model. In this sense the latter is more general than the former and allows to model a larger class of dynamic systems. Restricting the nonlinear functions  $f_i$ 's to polynomials makes it possible to use mathematic results in *differential algebra*. This nonlinearity is, however, not as restrictive as it seems, because many equations with *non* polynomial nonlinearities can be transformed into equivalent polynomial DAEs. For example, the equation

$$y = \sin x$$

can be transformed into an equivalent polynomial DAE

$$\dot{y}^2 + y^2 \dot{x}^2 = \dot{x}^2$$

where  $\dot{x}$  and  $\dot{y}$  are the time derivatives of  $x$  and  $y$ , respectively<sup>1</sup>. There are many other examples of this kind, in particular, all equations with polynomial fractions can be trivially transformed into polynomial DAEs.

---

<sup>1</sup>Strictly speaking, these two equations are equivalent except for some singular solutions.

## 2.2 Fault detection and isolation

The problem of fault detection (FD) consists in making the decision on the presence or absence of faults in a monitored system. When no fault is present in the system, the system is said in its *safe mode*, otherwise it is in a *faulty mode*. The problem of fault isolation (FI) consists in deciding the present faulty mode among a number of possible modes. As for all model-based approaches, we make these decisions based on available signals  $u$ ,  $y$  and a mathematical model. In particular, we assume the availability of a mathematical model in the form of (2) resulting from physical knowledges. Black-box models<sup>2</sup> could also be considered, but they are out of the scope of this paper. The interested reader may see (Zhang, 1996) for some related work.

After having established a mathematical model of the system, we need to express system faults in terms of the model in order to detect and isolate them through a model-based approach. In the FDI literature, faults are often modeled as additional terms in a state-space model. This is a reasonable approach for most sensor and actuator faults, but it is not the case for faults in the system itself. An alternative approach consists in modeling faults as changes in system parameters. This more versatile approach is adopted in this paper. Thus FD consists in deciding if any system parameter has changed, and FI consists in deciding which system parameters have changed.

Notice that in the above discussion we informally used the term “system parameter”. This assumes that the monitored system is characterized by some physical parameters. The model parameter vector  $\theta$  is in fact an image of system parameters. For the sake of simplicity, in this paper we only consider the case where the monitored system is in the considered model set. The extension to the case where the system is not in the model set is out of the scope of this paper. The reader is referred to (Zhang et al., 1994) for some related work. We thus assume that the system has the same parameterization as its mathematic model.

For a rigorous problem formulation, the system parameter should be distinguished from the model parameter. For the sake of presentation simplicity, however, we do not explicitly make the difference between the system and model parameters in this paper. We denote each of them by  $\theta$ . This is acceptable, as long as the monitored system is in the considered model set.

The value  $\theta = \theta_0$  corresponding to the safe mode of the monitored system is called the *nominal parameter value* and is usually obtained through system identification. Now we can formally state the FDI problems as follows.

**Problem 1 (FD)** *The fault detection problem is to decide between the two hypotheses*

$$\mathcal{H}_0 : \theta = \theta_0 \quad \text{and} \quad \mathcal{H}_0^c : \theta \neq \theta_0$$

**Problem 2 (FI)** *Let  $\underline{\theta}$  be any sub-vector of  $\theta$  and  $\underline{\theta}_0$  be the corresponding sub-vector of the nominal parameter value  $\theta_0$ . The fault isolation problem is to decide, for each sub-vector of  $\theta$ , between the two hypotheses*

$$\mathcal{H}_0(\underline{\theta}) : \underline{\theta} = \underline{\theta}_0 \quad \text{and} \quad \mathcal{H}_0^c(\underline{\theta}) : \underline{\theta} \neq \underline{\theta}_0$$

**Comment.** The FI problem is stated in terms of sub-vectors of  $\theta$ . This assumes the components of  $\theta$  are physically meaningful and their changes correspond to faults to be detected and isolated. The system model, in particular its parameterization, should be suitably chosen so

---

<sup>2</sup>Black-box models are identified from measured data without physical knowledge. See for example (Sjöberg et al., 1995; Juditsky et al., 1995).



that the faults to be detected and isolated can be modeled in this manner. More specifically, each faulty mode should correspond to changes in a sub-vector of  $\theta$ . Otherwise, the model should be modified (e.g., by changing its parameterization) so that the above fault modeling is possible.

### 2.3 Unknown variable estimation or elimination

One of the main difficulties in FDI of dynamic systems is due to the presence of unknown and unmeasured variables  $x$ . Two approaches are commonly used to deal with them: estimation and elimination. The estimation of  $x$  is usually performed with *observers*, also called *filters* for stochastic systems. The elimination of  $x$  directly explores the analytical redundancy in the mathematic model.

Observers or filters for linear systems have been extensively studied as well as their applications in FDI (see, e.g., (Patton et al., 1989; Frank, 1990; Patton, 1994)). They are much less studied for nonlinear systems, especially for nonlinear systems modeled by DAEs. It is known as a difficult problem and very few results are available. We can nevertheless cite (Nikoukhah, 1995; Krener and Duarte, 1996) and references therein for some recent theoretic results on nonlinear systems observer design.

By eliminating  $x$ , the original mathematical model is transformed into an equivalent input-output representation of the system. For linear systems, this approach is well known under the name of *parity check*. In this paper, we consider the elimination of  $x$  in the class of nonlinear systems modeled by DAEs (2).

The relation between observer and parity check based approaches for FDI has been discussed by several authors (Viswanadham and Srichander, 1987; Magni and Mouyon, 1994; Patton, 1994). It is known that parity equations correspond to some particular observers. Here we make one more remark on the relation between these two approaches. As shown later in this paper, eliminating  $x$  in a DAE model involves time derivatives of  $u$  and  $y$  in the resulting input-output model. Using such a model for FDI requires the estimation of these time derivatives. In some sense, the estimation of these derivatives is not very different from that of  $x$ , performed by observers or filters. However, if the derivatives of  $u$  and  $y$  are estimated using available signals only (not the system model), then the FDI algorithms are based on signals and estimates which are independent of the system model. In contrast, the estimate of  $x$  is necessarily based on the system model (in particular the nominal parameter value  $\theta = \theta_0$ ). Since the system parameter  $\theta$  is subject to unknown changes related to system faults, the design of FDI algorithms based on estimated  $x$  is more tricky<sup>3</sup>.

### 2.4 Sampled data and continuous time model

The use of continuous time models, in particular DAEs, is motivated by two reasons of very different nature: many physical laws are expressed by differential equations on one hand, and powerful mathematic tools in differential algebra are available for manipulating DAEs on the other hand. In practice, available signals are usually sampled at discrete time instants. Eliminating  $x$  in a DAE model in order to obtain an equivalent input-output representation concerns only the analytical equations, not the data. The computation of FDI residuals involves both the resulting input-output model and the data.

---

<sup>3</sup>An exception is additive faults in linear systems for which the innovation is a sufficient statistic.

Assume that an equivalent input-output representation of the DAE model (2) is

$$g(u, y, \theta, p) = 0 \quad (3)$$

where  $g$  is a vector of differential polynomials in  $u$ ,  $y$  and  $\theta$ . Let  $\tau$  be the sampling period, the available data is the pair of sequences

$$u_k = u(k\tau), \quad y_k = y(k\tau), \quad k = 0, 1, 2, \dots \quad (4)$$

The computation of FDI residuals will be based on the sampled data  $u_k, y_k$  and their time derivatives at the sampling instants. Usually these time derivatives are not directly measured, they must be estimated from the sampled data. Replacing time derivatives of  $u$  and  $y$  by their estimates in the DAE (3) and considering only the sampling time instants correspond to a discretization of the DAE. We formally write this discretized model as

$$g(u_k, y_k, \theta, \delta) = 0 \quad (5)$$

where  $\delta$  is the corresponding discrete approximation of the differentiation operator  $p$ . We refer the reader to (Goodwin et al., 1986) and references therein for discussions about  $\delta$  operator.

Before going to the next section we want to make the remark that the models presented in this section are all deterministic. Neither perturbation nor noise is introduced in these models. The elimination algorithm introduced in the next section is based on the deterministic DAE model (2). Appropriate stochastic assumption must be made for the design of FDI residuals based on the local approach, however. This assumption is stated in Section 4 where the FDI residual design is presented.

### 3 Input-output representation of DAE systems

The elimination of unknown variables in a dynamic system with general nonlinearities is a difficult problem for which no general algorithm is available. When polynomial DAEs (2) are considered, a general method has been proposed in (Ljung and Glad, 1994) for checking its identifiability. As mentioned in Section 2.1, many nonlinear equations can be transformed into polynomial DAEs, this method can thus deal with a large class of nonlinear dynamic systems. One by-product of this method is an input-output representation of the considered system. In this section, we give an informal summary of this method.

Assume the monitored system is modeled by a set of  $r$  DAEs as in (2). Let us re-write it for a better readability:

$$f_i(x, u, y, \theta, p) = 0, \quad i = 1, 2, \dots, r \quad (6)$$

where  $u$  and  $y$  are input and output signals,  $x$  represents (unknown) auxiliary variables,  $p$  is the time derivative operator,  $f_i$ 's are polynomials in  $u, y, x, p$  and parameterized by  $\theta$ .  $u, y, x$  and  $\theta$  can all be vector-valued.

The functions  $f_i$ 's in (6) are *differential polynomials* in  $x, u, y$  and  $\theta$ . Adding several  $f_i$ 's, differentiating a  $f_i$  in time, or multiplying a  $f_i$  by any polynomial in  $x, u, y, \theta$  results in another differential polynomial in the same variables. An infinite number of differential polynomials can be obtained by repeatedly performing the above operations. This infinite set of differential polynomials is called the *differential ideal* formed by the  $f_i$ 's. Let us denote by  $F$  this differential ideal.

Each differential polynomial  $\phi$  in  $F$  corresponds to a DAE  $\phi = 0$ . Due to this one-to-one correspondence between differential polynomials and polynomial DAEs, sometimes in our presentation we do not distinguish them. For instance, we may talk about the solution of a differential polynomial which is in fact the solution of the corresponding DAE.

If there exist  $x(t)$  and  $\theta$  such that signals  $u(t)$  and  $y(t)$  satisfy a set of polynomial DAEs, then we say that the pair  $u(t)$  and  $y(t)$  is a *solution* of the corresponding differential polynomials (which is in fact a solution of the set of DAEs). If signals  $u(t)$  and  $y(t)$  are a solution of the original set of DAEs (6), it is obvious that they are also a solution of the differential ideal  $F$ .

It turns out that, though  $F$  is an infinite set, it is sufficient to select a finite subset of  $F$  to specify its solutions. Such a subset is similar to a basis in linear algebra and is called a *characteristic set* of the differential ideal  $F$ . There exist infinitely many characteristic sets of  $F$ . A set of DAEs corresponding to any characteristic set of  $F$  is in some sense an equivalent representation of the original model (6), since they have the same solutions in  $u$  and  $y$ .

For the purpose of FDI, as well as for identifiability investigation, the characteristic sets of  $F$  not involving the variable  $x$  are of particular interest. As a matter of fact, they correspond to input-output representations of the system modeled by the original DAEs (6). Now the question is how to find such characteristic sets.

As explained in (Ljung and Glad, 1994), Ritt's algorithm (Ritt, 1950) is convenient for searching characteristic sets of a differential ideal with some desired properties. It can be thought of as the equivalent, for differential polynomials, of Gram-Schmidt or Gauss elimination algorithms in linear algebra. It is of practical importance to point out that computer implementations of Ritt's algorithm are available (Ljung and Glad, 1994).

The theorem on global identifiability presented in (Ljung and Glad, 1994) can be informally stated as follows. If there exists a characteristic set of  $F$  including differential polynomials of the form

$$P_j(u, y, p) \theta_{(j)} - Q_j(u, y, p) \quad (7)$$

for all  $j = 1, 2, \dots, n$  ( $n$  is the dimension of  $\theta$  and  $\theta_{(j)}$  is its  $j$ -th component), where  $P_j(u, y, p)$  and  $Q_j(u, y, p)$  are differential polynomials in  $u$  and  $y$ , then  $\theta$  is globally identifiable, i.e.,  $\theta$  can be uniquely determined in  $\mathbf{C}^n$  under the condition of persistent excitation. Otherwise  $\theta$  is not globally identifiable. See (Ljung and Glad, 1994) for more details.

A remarkable fact in (7) is that not only  $x$  has been eliminated, but also the parameter components  $\theta_{(j)}$ 's have been fully decoupled. The DAEs corresponding to these differential polynomials seem to provide us with a straightforward way to design residuals for fault isolation. However, this decoupling in  $\theta_{(j)}$ 's is usually obtained at the price of high order time derivatives of  $u$  and  $y$ , whose estimation is difficult in practice. For this reason, we prefer input-output representations less decoupled in  $\theta$ , but with time derivatives of lower order. Thus we search within  $F$  for differential polynomials of the form

$$g_i(u, y, \theta, p), \quad i = 1, 2, \dots$$

where only  $x$  is eliminated from the differential polynomials  $g_i$ 's. For notation compactness, we write the set of  $g_i$ 's in a vectorial form:

$$g(u, y, \theta, p) \quad (8)$$

where  $g$  is a vector of differential polynomials in  $u, y, \theta$ .

Ritt's algorithm can be used to search within  $F$  for particular  $g(u, y, \theta, p)$  which is linear in  $\theta$ , i.e.,

$$g(u, y, \theta, p) = P(u, y, p) \theta - Q(u, y, p) \quad (9)$$

where  $P(u, y, p)$  and  $Q(u, y, p)$  are polynomial matrices or vectors of appropriate dimensions.

The DAE corresponding to (8), namely

$$g(u, y, \theta, p) = 0 \quad (10)$$

or in its particular case

$$P(u, y, p) \theta - Q(u, y, p) = 0 \quad (11)$$

is an input-output representation of the original system.

The existence of (9) in  $F$  is a sufficient and necessary condition for  $\theta$  to be globally identifiable (Ljung and Glad, 1994). If  $\theta$  is not globally, but locally identifiable (in particular at the nominal value  $\theta_0$ ), only differential polynomials  $g(u, y, \theta, p)$  which are nonlinear in  $\theta$  can be found. In both cases, (10) and (11) are valid input-output representations of the original system for the purpose of FDI. Moreover, a scalar equation (10) (with a single polynomial  $g(u, y, \theta, p)$  instead of a vector of polynomials) is sufficient for our FDI design. Using more such equations usually involves higher order time derivatives of  $u$  and  $y$ . Anyway for the sake of generality, we assume that (10) and (11) are vectors.

**Example.** Let us illustrate variable elimination with an example borrowed from (Ljung and Glad, 1994). Consider the model

$$\dot{x}_1 = \theta x_2^2 \quad (12)$$

$$\dot{x}_2 = u \quad (13)$$

$$y = x_1 \quad (14)$$

The elimination of  $x_1$  is trivial through equation (14). Plugging (14) into (12) results in

$$\dot{y} = \theta x_2^2 \quad (15)$$

Differentiating (15) gives

$$\ddot{y} = 2\theta x_2 \dot{x}_2 \quad (16)$$

Plugging (13) into (16) leads to

$$\ddot{y} = 2\theta x_2 u \quad (17)$$

Squaring (17) gives

$$\dot{y}^2 = 4\theta^2 x_2^2 u^2 \quad (18)$$

Finally, combining (18) and (15) results in

$$\ddot{y}^2 - 4\theta \dot{y} u^2 = 0 \quad (19)$$

We have eliminated  $x_1, x_2$  “by hand”. Ritt’s algorithm does not perform exactly the same steps, but results in the same solution. As we find an equation in the form of (11),  $\theta$  is globally identifiable.

## 4 FDI design based on the local approach

In this section we present the local approach to designing FDI residuals based on the input-output representations introduced in the previous section. As only sampled signals are available in practice, the residual design is based on discrete time approximations of DAEs. As explained in Section 2.4, DAE (10) is discretized into

$$g(u_k, y_k, \theta, \delta) = 0 \quad (20)$$

where  $\delta$  is a discrete approximation of the differentiation operator.

### 4.1 Stochastic modeling and primary residual

So far only deterministic models have been considered. Due to modeling uncertainty and measurement errors, some stochastic assumption should be introduced in order to take them into consideration. For nonlinear systems, such assumption should be made based on physical knowledge. For the sake of simplicity, we make the following assumption:

$$g(u_k, y_k, \theta, \delta) = \varepsilon_k \quad (21)$$

where the noise  $\varepsilon_k$  is zero mean and independent of  $u_k, y_k$ . Notice that this assumption on statistic properties of  $\varepsilon_k$  is very weak.

Now we are ready to define the *primary residual* for FDI.

**Definition 1** *Given an input-output model (21), a function  $H(u_k, y_k, \theta, \delta)$  is a valid primary residual for monitoring  $\theta$  if it is differentiable in  $\theta$  and there exists a neighborhood  $\kappa(\theta_0)$  of the nominal parameter value  $\theta_0$  such that the following properties hold:*

$$\mathbf{E}_\theta (H(u_k, y_k, \theta_0, \delta)) = 0 \quad \text{if } \theta = \theta_0 \quad (22)$$

$$\mathbf{E}_\theta (H(u_k, y_k, \theta_0, \delta)) \neq 0 \quad \text{if } \theta \neq \theta_0 \text{ and } \theta \in \kappa(\theta_0) \quad (23)$$

where  $\mathbf{E}_\theta$  is the expectation when the actual system parameter value is  $\theta$ .

The following remarks are in order about this definition.

1. The expectations in (22) and (23) are conditioned by the system parameter  $\theta$  since the measurement sample  $(u_k, y_k)$  is governed by  $\theta$ .
2. In (22) and (23) the primary residual  $H(u_k, y_k, \theta_0, \delta)$  is evaluated at the nominal parameter value  $\theta_0$ . The value  $\theta_0$  is usually estimated by system identification when the system is in its safe mode, whereas the actual value of the system parameter  $\theta$  (governing the collected data) is unknown (we do *not* re-identify the parameter).
3. The primary residual  $H(u_k, y_k, \theta_0, \delta)$  is usually a vector-valued function. Its dimension should be greater than or equal to the dimension of  $\theta$  for the two properties (22) and (23) to hold.

A typical example of the primary residual is

$$\begin{aligned} H(u_k, y_k, \theta, \delta) &= -\frac{1}{2} \frac{\partial}{\partial \theta} \left( g^T(u_k, y_k, \theta, \delta) g(u_k, y_k, \theta, \delta) \right) \\ &= - \left( \frac{\partial g(u_k, y_k, \theta, \delta)}{\partial \theta} \right)^T g(u_k, y_k, \theta, \delta) \end{aligned} \quad (24)$$

As a matter of fact, if  $\theta$  is locally identifiable<sup>4</sup> at the nominal value  $\theta = \theta_0$ , then the two required properties (22) and (23) hold.

As an interesting particular instance, assume that the noise  $\varepsilon_k$  in (21) is i.i.d. with a Gaussian distribution and with an identity covariance matrix, then  $H(u_k, y_k, \theta, \delta)$  is nothing but an *efficient score*<sup>5</sup> for  $\theta$  in model (21).

Actually, the above example of primary residual is strongly related to some algorithms for the identification of  $\theta_0$ . Alternative primary residuals can be derived from other identification algorithms. See, for example, (Benveniste et al., 1987; Basseville and Nikiforov, 1993).

Properties (22) and (23) are the very basic ones usually required for FDI residuals. In order to perform the evaluation of a residual for FDI, it is important to know the statistical properties of the residual. Unfortunately, even if the noise  $\varepsilon_k$  in (21) has a Gaussian distribution, we cannot say much about the distribution of the primary residual  $H(u_k, y_k, \theta, \delta)$ . For this reason, we will derive an improved residual from the primary one based on a local approach. This improved residual will have well normalized asymptotic statistic properties. We first introduce it for the purpose of FD, then it is also used for FI.

## 4.2 Fault detection

As mentioned in the introduction, detecting non additive faults is a difficult problem even for linear dynamic systems. Let us consider the local approach to FD which is based on a particular assumption: the faults are “small”. Formally speaking, we consider the following problem instead of Problem 1.

**Problem 3 (Local FD)** *Let  $\theta$  be the system parameter with nominal value  $\theta_0$ . Given a  $N$ -size sample of input-output data*

$$\{(u_k, y_k) : k = 1, 2, \dots, N\} \quad (25)$$

*decide between the two hypotheses*

$$\mathcal{H}_0 : \theta = \theta_0 \quad \text{and} \quad \mathcal{H}_1 : \theta = \theta_0 + \frac{\eta}{\sqrt{N}}$$

*where  $\eta \neq 0$  is an unknown vector with the same dimension as  $\theta$ .*

The hypothesis  $\mathcal{H}_1$  is asymptotically local in the sense that, for fixed  $\eta$ , the deviation of  $\theta$  from  $\theta_0$  tends to zero when  $N$  becomes infinitely large. In practice, of course, parameter changes are independent of the sample size.  $\mathcal{H}_1$  is only a working hypothesis. See (Zhang et al., 1994, Section 1.1) for more explanations.

<sup>4</sup>See (Ljung and Glad, 1994) for exact conditions on local identifiability.

<sup>5</sup>The efficient score is the gradient, with respect to the model parameters, of the log-likelihood function of a given observation sample.

### 4.2.1 Normalized residual based on the local approach

Now we are ready to derive the *normalized residual* from a primary residual.

**Definition 2** Given a primary residual function  $H(u_k, y_k, \theta, \delta)$  as defined in Definition 1 and a  $N$ -size data sample as in (25), the corresponding normalized residual is defined as

$$\zeta_N(\theta) = \frac{1}{\sqrt{N}} \sum_{k=1}^N H(u_k, y_k, \theta, \delta) \quad (26)$$

In order to state the properties of this normalized residual, we need the following notations:

$$M(\theta_0) = - \left. \frac{\partial}{\partial \theta} \mathbf{E}_{\theta_0} (H(u_k, y_k, \theta, \delta)) \right|_{\theta=\theta_0} \quad (27)$$

$$\Sigma(\theta_0) = \sum_{k=-\infty}^{+\infty} \text{cov}_{\theta_0} (H(u_0, y_0, \theta_0, \delta), H(u_k, y_k, \theta_0, \delta)) \quad (28)$$

where  $\mathbf{E}_{\theta_0}$  and  $\text{cov}_{\theta_0}$  are respectively the expectation and the covariance when the actual system parameter  $\theta = \theta_0$ .

**Theorem.** Given a primary residual  $H(u_k, y_k, \theta_0, \delta)$  as defined in Definition 1. If  $\Sigma(\theta_0)$  defined in (28) exists and is positive definite, then under both  $\mathcal{H}_0$  and  $\mathcal{H}_1$  as defined in Problem 3, the normalized residual  $\zeta_N(\theta_0)$  defined in Definition 2 is Gaussian distributed when  $N$  goes to infinity. Specifically, when  $N \rightarrow \infty$ ,

$$\begin{aligned} \zeta_N(\theta_0) &\sim \mathcal{N}(0, \Sigma(\theta_0)) && \text{under } \mathcal{H}_0 \\ \zeta_N(\theta_0) &\sim \mathcal{N}(M(\theta_0)\eta, \Sigma(\theta_0)) && \text{under } \mathcal{H}_1 \end{aligned} \quad (29)$$

See (Benveniste et al., 1987; Zhang et al., 1994) for a proof of this theorem.

The estimation of  $M(\theta_0)$  from a data sample is straightforward by replacing the expectation in (27) by sample average. The estimation of  $\Sigma(\theta_0)$  is numerically more tricky. It can be estimated either as a truncated version of (28) or as an empirical covariance matrix of the residual  $\zeta_N(\theta_0)$ . See (Zhang et al., 1994) for more hints on the estimation of  $\Sigma(\theta_0)$ .

As an example, let us consider the particular case where

$$g(u_k, y_k, \theta, \delta) = P(u_k, y_k, \delta) \theta - Q(u_k, y_k, \delta)$$

with polynomial matrices (or vectors)  $P$  and  $Q$  of appropriate dimensions. Notice that  $g(u_k, y_k, \theta, \delta)$  is linear in  $\theta$  in this case. The existence of such a DAE is in fact the sufficient and necessary condition for  $\theta$  to be globally identifiable (Ljung and Glad, 1994). The corresponding stochastic model is

$$P(u_k, y_k, \delta) \theta - Q(u_k, y_k, \delta) = \varepsilon_k \quad (30)$$

Then the primary residual defined in (24) writes

$$H(u_k, y_k, \theta, \delta) = - \left( P^T(u_k, y_k, \delta) P(u_k, y_k, \delta) \right) \theta + \left( P^T(u_k, y_k, \delta) Q(u_k, y_k, \delta) \right)$$

and

$$M(\theta_0) = \mathbf{E}_{\theta_0} \left( P^T(u_k, y_k, \delta) P(u_k, y_k, \delta) \right)$$

which is an invertible matrix if the system is persistently excited.

For the same example, from model (30) we have

$$H(u_k, y_k, \theta, \delta) = -P^T(u_k, y_k, \delta) \varepsilon_k$$

Therefore,

$$\Sigma(\theta_0) = \sum_{k=-\infty}^{+\infty} \text{cov}_{\theta_0} \left( P^T(u_0, y_0, \delta) \varepsilon_0, P^T(u_k, y_k, \delta) \varepsilon_k \right)$$

And if  $\varepsilon_k$  is an independent sequence with variance  $\Sigma_\varepsilon$ , then

$$\Sigma(\theta_0) = \mathbf{E}_{\theta_0} \left( P^T(u_k, y_k, \delta) \Sigma_\varepsilon P(u_k, y_k, \delta) \right)$$

#### 4.2.2 Residual evaluation for FD

For the sake of notation simplicity in the following presentation, we define the abbreviated notations:

$$\zeta = \zeta_N(\theta_0), \quad M = M(\theta_0), \quad \Sigma = \Sigma(\theta_0)$$

For evaluating the normalized residual  $\zeta$ , we need an extra assumption: the matrix  $M$  defined in (27) has full column rank.

Thanks to the properties (29), the detection of changes in the system parameter  $\theta$  is asymptotically equivalent to the detection of changes in the mean of a Gaussian vector. It is well known that the *generalized likelihood ratio (GLR) test* detecting unknown changes in the mean of a Gaussian vector is a  $\chi^2$ -test. See, e.g., (Benveniste et al., 1987; Basseville and Nikiforov, 1993) and references therein for further explanations. Specifically, assume that  $N$  is sufficiently large so that the Gaussian distribution of  $\zeta$  is verified, the GLR test of  $\mathcal{H}_1$  against  $\mathcal{H}_0$  as stated in Problem 3 is

$$t_\eta = \zeta^T \Sigma^{-1} M \left( M^T \Sigma^{-1} M \right)^{-1} M^T \Sigma^{-1} \zeta$$

This quantity, named as the *global test*, has a central  $\chi^2$ -distribution under  $\mathcal{H}_0$ , and a non-central  $\chi^2$ -distribution under  $\mathcal{H}_1$ . Its number of degrees of freedom is equal to the dimension of  $\theta$ . The non-centrality parameter of  $t_\eta$  under  $\mathcal{H}_1$  is

$$\gamma = \eta^T M^T \Sigma^{-1} M \eta$$

It is known (see e.g. (Basseville, 1997)) that the detection power of  $t_\eta$  (the probability of successful detection for a given probability of false alarm) is an increasing function of  $\gamma$ .

The statistic properties of  $t_\eta$  provide a useful guideline for the choice of a threshold based on the  $\chi^2$ -table and the desired false alarm probability. Let  $\lambda$  be such a threshold, the FD decision is performed by the rule:

$$\begin{aligned} t_\eta \leq \lambda &\longrightarrow \text{no fault} \\ t_\eta > \lambda &\longrightarrow \text{fault} \end{aligned}$$

If  $M$  is an invertible matrix, then the global test  $t_\eta$  reduces to

$$t_\eta = \zeta^T \Sigma^{-1} \zeta \tag{31}$$



### 4.3 Fault isolation

We consider off-line FI after the detection of faults by the FD rule. More clearly, the FDI problem is addressed in two stages:

1. FD is performed as frequently as possible in order to achieve quick detection of faults. It is of course conditioned by the availability of measurements and the power of the used computer. The measured data can be processed in blocks or in a sequential and recursive manner. In this paper only the algorithm for block processing is presented for the sake of simplicity. When the presence of a fault is decided, we say that a FD alarm is delivered.
2. FI is performed only after the deliverance of a FD alarm. It should be performed over all the available data measured since the FD alarm.

So in our discussion on FI, we assume that a FD alarm is delivered and a data sample is collected based on which FI is performed. Moreover, we make the assumption that the sample size  $N$  is large enough so that the Gaussian distribution of  $\zeta$  is verified. More specifically, we assume that

$$\begin{aligned}\theta &= \theta_0 + \frac{\eta}{\sqrt{N}} \\ \zeta &\sim \mathcal{N}(M\eta, \Sigma)\end{aligned}$$

As explained in Section 2.2, FI is performed by deciding which components of the system parameter  $\theta$  have changed. This problem is asymptotically equivalent to deciding which components of  $\eta$  are non zero. Problem 2 is thus replaced by the asymptotically equivalent problem:

**Problem 4 (Local FI)** *For all sub-vector  $\eta_a$  of  $\eta$ , decide between the two hypotheses*

$$\mathcal{H}_0(\eta_a) : \eta_a = 0 \quad \text{and} \quad \mathcal{H}_0^c(\eta_a) : \eta_a \neq 0$$

Without loss of generality, assume that  $\eta_a$  consists of the first components of  $\eta$  (otherwise a permutation would be needed). Let  $\eta_b$  be the complementary vector of  $\eta_a$  in  $\eta$ , i.e.,

$$\eta = \begin{bmatrix} \eta_a \\ \eta_b \end{bmatrix} \tag{32}$$

where the dimensions of  $\eta_a$  and  $\eta_b$  are respectively  $n_a$  and  $n_b$  ( $n_a + n_b = n$ ).

Remark that nothing about the value of  $\eta_b$  is assumed in Problem 4. There are two natural ways to deal with this problem: assume  $\eta_b$  to be zero or equal to the least favorable value for making the decision between  $\mathcal{H}_0(\eta_a)$  and  $\mathcal{H}_0^c(\eta_a)$ . These two solutions result in the so-called *sensitivity test* and *minmax test*, respectively (Basseville and Nikiforov, 1993; Basseville, 1997). We summarize these two tests in the following.

#### 4.3.1 Sensitivity test

For a given partition of  $\eta$  as in (32), the sensitivity test consists in testing

$$\eta_a \neq 0, \eta_b = 0 \quad \text{against} \quad \eta_a = 0, \eta_b = 0 \tag{33}$$

Make the corresponding partition of  $M$ :

$$M = [M_a, M_b]$$

As shown in (Basseville and Nikiforov, 1993; Basseville, 1997), the GLR test for (33) results in the sensitivity test:

$$\tilde{t}_{\eta_a} = \zeta^T \Sigma^{-1} M_a \left( M_a^T \Sigma^{-1} M_a \right)^{-1} M_a^T \Sigma^{-1} \zeta$$

which is a  $\chi^2$ -test with  $n_a$  degrees of freedom.

### 4.3.2 Minmax test

Let

$$\mathbf{F} = M^T \Sigma^{-1} M$$

be the *Fisher information* on  $\eta$  contained in  $\zeta$ . Make the partition of  $\mathbf{F}$  corresponding to that of  $M$

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{aa} & \mathbf{F}_{ab} \\ \mathbf{F}_{ba} & \mathbf{F}_{bb} \end{bmatrix} = \begin{bmatrix} M_a^T \Sigma^{-1} M_a & M_a^T \Sigma^{-1} M_b \\ M_b^T \Sigma^{-1} M_a & M_b^T \Sigma^{-1} M_b \end{bmatrix}$$

and define

$$\mathbf{F}_a^* = \mathbf{F}_{aa} - \mathbf{F}_{ab} \mathbf{F}_{bb}^{-1} \mathbf{F}_{ba}$$

Perform a linear transformation of  $\zeta$ :

$$\zeta_a^* = \mathcal{G}_a^* \zeta$$

with

$$\mathcal{G}_a^* = \begin{bmatrix} I_a, & -\mathbf{F}_{ab} \mathbf{F}_{bb}^{-1} \end{bmatrix} M^T \Sigma^{-1}$$

where  $I_a$  is the  $n_a \times n_a$  identity matrix.

The *minmax test* (Basseville and Nikiforov, 1993; Basseville, 1997), which maximizes the detection power with the least favorable value of  $\eta_b$ , writes

$$t_{\eta_a}^* = \zeta_a^{*T} \mathbf{F}_a^{*-1} \zeta_a^*$$

which is also a  $\chi^2$ -test with  $n_a$  degrees of freedom.

### 4.3.3 Using sensitivity and minmax tests for FI

For each sub-vector  $\eta_a$  of  $\eta$  we can compute both a sensitivity test and a minmax test. How should we make FI decisions based on these tests? Should we choose a threshold for each  $\tilde{t}_{\eta_a}$  or  $t_{\eta_a}^*$ ?

Consider the sensitivity test. When  $\eta_a = 0$  but  $\eta_b \neq 0$ , we may expect  $\tilde{t}_{\eta_a}$  to be close to  $n_a$  (mean value of a central  $\chi^2(n_a)$ ) and  $\tilde{t}_{\eta_b}$  to be large. In fact, due to the coupling effect between  $\eta_a$  and  $\eta_b$ , the value of  $\tilde{t}_{\eta_a}$  is far from  $n_a$ . The correlation between  $\tilde{t}_{\eta_a}$  and  $\tilde{t}_{\eta_b}$  is related to the off-diagonal entries of the Fisher information matrix  $\mathbf{F} = M^T \Sigma^{-1} M$ . See (Basseville, 1997) for more explanations. It is thus difficult to choose a threshold for  $\tilde{t}_{\eta_a}$  in general.

If the minmax test is considered, at least in theory,  $t_{\eta_a}^*$  follows the central  $\chi^2$ -distribution when  $\eta_a = 0$ , it is thus possible to choose a threshold with the aid of the  $\chi^2$ -table. In this case, only the one-dimensional  $\eta_a$ 's (sub-vectors of  $\eta$ ) need to be considered and the non-zero

components of  $\eta$  are found one by one. This method works well if  $\zeta$  really follows the Gaussian distribution  $\mathcal{N}(M\eta, \Sigma)$ . In practice, however, due to approximations (in particular,  $N$  is never infinite), this assumption is not fully true, and it is difficult to choose a threshold for each  $t_{\eta_a}^*$ .

If  $n_a$  (the number of non-zero components of  $\eta$ ) is known, we strongly recommend the following approach which have proved efficient in our experiences. Compute a minmax test for each possible sub-vector  $\eta_a$  of dimension  $n_a$ . Among all these minmax tests, the one having the largest value indicates the non zero sub-vector  $\eta_a$ . An obvious advantage of this approach is that no threshold is needed. It is known that this strategy is optimal in some precise sense (Spjøtvoll, 1972).

It is interesting to notice the following relation:

$$t_\eta = t_{\eta_a}^* + \tilde{t}_{\eta_b} \quad (34)$$

and that this decomposition of  $t_\eta$  is orthogonal in the sense that  $t_{\eta_a}^*$  and  $\tilde{t}_{\eta_b}$  are statistically uncorrelated (Basseville, 1997). This provides an indirect way to compute the minmax test:

$$t_{\eta_a}^* = t_\eta - \tilde{t}_{\eta_b}$$

## 5 Numerical example

In this section, we present results of the proposed FDI method applied to a simulated microbial growth process.

### 5.1 System description

Recently, there has been an increasing interest in the application of control theory to microbial growth processes, such as productions by fermentation and waste water processing (Gauthier et al., 1992). In a microbial growth process, there are initially a certain number of microorganisms fed with a substrate (sugar, waste water, etc.). The number of microorganisms grows by consuming the substrate. Besides the automatic regulation of such processes, it may be useful to monitor changes in the biological reactions. For this purpose, let us consider a microbial growth process with a continuously fed substrate. It can be modeled by a Monod model (Monod, 1950)

$$\dot{x}_1 = \frac{ax_1x_2}{x_2 + b} - x_1u \quad (35)$$

$$\dot{x}_2 = -\frac{cax_1x_2}{x_2 + b} + (d - x_2)u \quad (36)$$

with the following notations

- $x_1$ : microbial concentration in the process,
- $x_2$ : substrate concentration in the process,
- $u$ : dilution rate,
- $a$ : maximum growth rate,
- $b$ : saturation parameter,
- $c$ : yield factor,
- $d$ : inlet substrate concentration.

$x_1, x_2$  are state variables;  $u$  is the system input;  $a, b, c, d$  are parameters. We assume that  $x_2$  is observed, i.e.,

$$y = x_2 . \quad (37)$$

## 5.2 Input-output representation

Before applying Ritt's algorithm, model (35)–(37) has to be transformed into polynomial DAEs:

$$(x_2 + b)\dot{x}_1 - ax_1x_2 + (x_2 + b)x_1u = 0 \quad (38)$$

$$(x_2 + b)\dot{x}_2 + cax_1x_2 - (x_2 + b)(d - x_2)u = 0 \quad (39)$$

$$y - x_2 = 0 \quad (40)$$

Glad and Ljung's implementation of Ritt's algorithm (Ljung and Glad, 1994) is applied in order to obtain an input-output representation of (38)–(40). The result shows that the parameter  $c$  is not identifiable. It can be intuitively explained as follows. Assume that  $c$  is multiplied by a factor, say  $\xi$ , then equation (39) would still be satisfied if  $x_1$  was divided by  $\xi$ . Moreover, it is obvious that (38) and (40) are not affected by any scale change of  $x_1$ . For this reason  $c$  can be neither identified nor monitored using the input-output measurements  $u$  and  $y$ .

The input-output representation given by Ritt's algorithm is

$$\begin{aligned} & a \left( -\dot{y}y^2 + duy^2 - y^3u \right) \\ & + b \left( d\dot{y}u + y^2\dot{u} - d\dot{y}y - (\dot{y})^2 + y^2u^2 + \dot{y}uy - du^2y + \ddot{y}y \right) \\ & - \left( dy^2\dot{u} - y^3\dot{u} - y^2\ddot{y} - y^3u^2 - 2y^2u\dot{y} + du^2y^2 \right) = 0 \end{aligned} \quad (41)$$

As the inlet substrate concentration  $d$  is directly measured and thus can be monitored separately, we only consider the monitoring of  $a$  and  $b$ . Equation (41) is in form (11) if we set

$$\theta = (a, b)^T$$

## 5.3 Simulation, filtering and identification

Equations (35), (36) and (37) are used to simulate the microbial growth process with the aid of SIMULINK. The nominal values of the model parameters are

$$a = 0.55, \quad b = 0.15, \quad c = 2, \quad d = 0.8 .$$

The input signal is

$$u(t) = 0.07 \sin(0.011t) + 0.38 . \quad (42)$$

The output measurement  $y$  is corrupted by a white noise

$$y_k = x_2(k\tau) + \epsilon_k$$

The standard deviation of the noise  $\epsilon_k$  is 0.02. The sampling period  $\tau = 0.5$  second.

Remark that the simulation of the measurement noise  $\epsilon_k$  does not conform to the stochastic model assumed in (21). Nevertheless, the FDI method presented in this paper seems to work in such situations.

The implementation of the  $\delta$  operator, or equivalently, the estimation of time derivatives of a signal from sampled data, is a difficult problem when the data is noise corrupted. Classical methods can be found, for example, in (Oppenheim et al., 1983, Section 8.4.1). As a particular implementation, the (measured) noise corrupted signal  $y$  is first filtered with a low pass filter (a 4th order Butterworth filter with cut-off frequency 0.02 Hz). The filtered  $y$  is then locally fitted with first and second degrees polynomials from which the derivatives  $\dot{y}$ ,  $\ddot{y}$  are estimated. The derivative  $\dot{u}$  is estimated in a similar way.

The values of  $c$  and  $d$  are assumed to be known. The nominal value of  $\theta = (a, b)^T$  (denoted as  $\theta_0$ ) is identified on a sample of  $N = 2000$  simulated measurements. The identification is simply performed by the least squares solution of (41) as follows. Rewrite (41) as

$$P_k \theta - Q_k = 0 \quad (43)$$

where  $k$  is the discrete time instant,  $\theta = (a, b)^T$  and

$$P_k = \left[ \left( -\dot{y}y^2 + du^2y^2 - y^3u \right), \left( d\dot{y}u + y^2\dot{u} - d\dot{u}y - (\dot{y})^2 + y^2u^2 + yuy - du^2y + \ddot{y}y \right) \right]_{t=k\tau} \quad (44)$$

$$Q_k = \left[ dy^2\dot{u} - y^3\dot{u} - y^2\ddot{y} - y^3u^2 - 2y^2u\dot{y} + du^2y^2 \right]_{t=k\tau} \quad (45)$$

where  $t = k\tau$  means the variables are sampled at time instant  $k$ . Let

$$\mathcal{P}_1^N = \begin{bmatrix} P_1 \\ \vdots \\ P_N \end{bmatrix} \quad \mathcal{Q}_1^N = \begin{bmatrix} Q_1 \\ \vdots \\ Q_N \end{bmatrix}$$

then

$$\hat{\theta}_0 = \left[ (\mathcal{P}_1^N)^T (\mathcal{P}_1^N) \right]^{-1} (\mathcal{P}_1^N)^T \mathcal{Q}_1^N$$

The identified value is  $\hat{\theta}_0 = (0.5499, 0.1497)^T$ .

Formula (24) gives

$$H(u_k, y_k, \theta, \delta) = -P_k^T P_k \theta + P_k^T Q_k$$

where  $P_k$  and  $Q_k$  are as in (44) and (45).

The matrix  $\Sigma$  is estimated by

$$\Sigma \approx \frac{1}{n_0} \sum_{k=1}^{n_0} H_k H_k^T + \sum_{i=1}^I \frac{1}{n_i} \sum_{k=1}^{n_i} \left( H_k H_{k+i}^T + H_{k+i} H_k^T \right)$$

where  $H_k = H(u_k, y_k, \hat{\theta}_0, \delta)$ , the values of  $n_0, n_1, \dots, n_I$  and  $I$  should be properly selected. See (Zhang et al., 1994) for more details on the estimation of  $\Sigma$ .

In figures 1 and 2, two samples of simulated output measurements  $y$  are depicted. The first one is simulated with the nominal parameter  $\theta = \theta_0 = (0.55, 0.15)^T$ , and the second one is simulated with  $\theta = (0.5418, 0.1397)^T$ . They are both simulated with the same input signal (42), but with different noise samples.

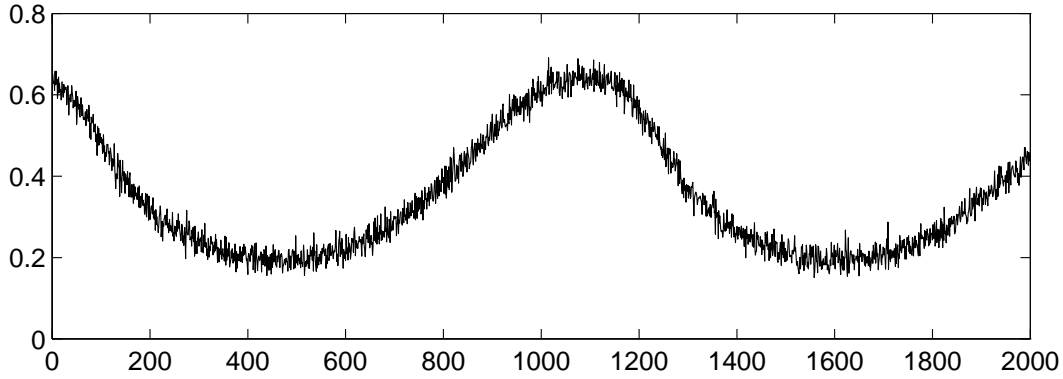


Figure 1: Output  $y_k$  simulated with the nominal parameter  $\theta = (0.55, 0.15)^T$ .

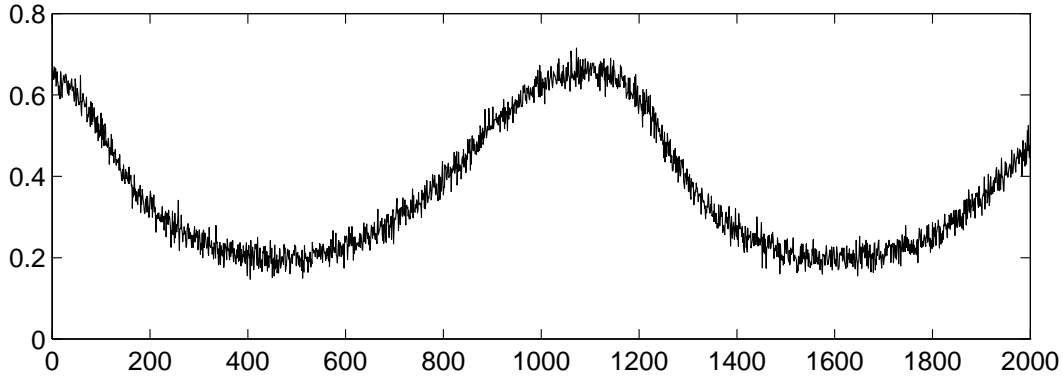


Figure 2: Output  $y_k$  simulated with a changed parameter  $\theta = (0.5465, 0.15)^T$ .

## 5.4 Fault detection

The global test (31) is applied to data simulated with the nominal model, and to data simulated with modified parameters  $a$  and  $b$ . For each experiment, the sample size is  $N = 2000$ . For each considered parameter value (nominal or changed), 100 experiments are performed with different samples of measurement noise.

The results of these tests are summarized in table 1, where the empirical means and standard deviations of the tests are compared with their theoretical values.

Actually, a  $\chi^2$  distribution, with  $n$  degrees of freedom and non-centrality parameter  $\gamma$ , has the mean value

$$\mu = n + \gamma$$

and the standard deviation

$$\sigma = \sqrt{2n + 4\gamma} = \sqrt{4\mu - 2n}$$

In our example,  $n = \dim(\theta) = 2$ . If no changes occurs in the system, then  $\gamma = 0$ , the  $\chi^2$ -distribution is said to be central and  $\mu = n$ .

When no fault occurs, the global test (31) should have a central  $\chi^2$ -distribution, then we know its theoretic mean  $\mu = n = 2$  and standard deviation  $\sigma = \sqrt{2n} = 2$ . In the presence of a fault, we know neither the theoretic mean nor the non-centrality parameter of the test. In order to have some idea how large the standard deviation should be, we compute a “semi-theoretic standard deviation” with the empirical mean  $\mu_{emp}$  and  $n$ :

$$\sigma_{semi} = \sqrt{4\mu_{emp} - 2n}$$

Table 1: Numerical values of test (31) for the microbial process. Each column corresponds to 100 experiments simulated with a given value of  $\theta$  (the values are noted as  $\theta_0, \theta_1, \theta_2, \theta_3$ ), and with different samples of the measurement noise. According to the  $\chi^2$ -table, applying the threshold  $\lambda = 9.21$  to this test corresponds (in theory) to the false alarm probability = 0.01.

	$\theta_0$ (nominal) (0.55, 0.15)	$\theta_1$ (0.5465, 0.15)	$\theta_2$ (0.55, 0.1465)	$\theta_3$ (0.5535, 0.1465)
empirical mean	3.22	327.68	340.36	1328.30
empirical st. deviation	3.04	36.93	44.93	85.94
theoretic mean	2	—	—	—
theoretic st. deviation	2	—	—	—
semi-theoretic st. deviation	2.98	36.15	36.84	72.86

It is shown in the last row of table 1.

It is obvious that the three type of changes can be successfully detected if we choose the threshold  $\lambda = 9.21$ . According to the  $\chi^2$ -table, this threshold corresponds to the false alarm probability of 0.01.

## 5.5 Fault isolation

The correctness of FI depends on the magnitude of the parameter changes, on the noise level and on the used sample size. We have chosen to show results with different magnitudes of parameter changes, but with fixed noise level and fixed sample size ( $N = 2000$ ).

### 5.5.1 Simulations with changes in parameter $a$

Simulations are performed with 10 different values of  $a$  to generate signals  $u$  and  $y$ . For each of the 10 values of  $a$  and the nominal value of  $b$ , 100 experiments are simulated with different random noise realizations. For each experiment, the global test  $t_\eta$ , the two minmax tests  $t_{\eta_a}^*$ ,  $t_{\eta_b}^*$ , and the two sensitivity tests  $\tilde{t}_{\eta_a}$ ,  $\tilde{t}_{\eta_b}$  are computed. Under the assumption that only one parameter has changed, the largest minmax test indicates the changed parameter.

Note that due to the relation (34), in the case of two parameters, the largest sensitivity test indicates the same parameter as the largest minmax test does.

Over 100 experiments for each of the 10 values of  $a$ , the rate of successful isolation, the mean value of the global test, the mean value of the minmax tests and the mean value of the sensitivity tests are computed. These results are shown in figure 3.

Notice that, in figure 3-(c), the values of the minmax tests  $t_{\eta_b}^*$  (represented by the crosses), though remain relatively small, increase with the change magnitude of  $a$ . This is because the local assumption becomes less correct when the the parameter change grows.

For this example the Fisher information matrix of the residual  $\zeta$  in regard to  $\theta = (a, b)^T$  is

$$\mathbf{F} = \begin{bmatrix} 13534 & -12618 \\ -12618 & 13373 \end{bmatrix}$$

As the off-diagonal entries are close (in absolute value) to the diagonal entries, the two parameters  $a, b$  are strongly coupled in the residual  $\zeta$ . This explains correlation between the two

sensitivity tests  $\tilde{t}_{\eta_a}$  and  $\tilde{t}_{\eta_b}$  as shown in figure 3-(b). Due to the same reason, it is difficult to determine the number of changed parameters.

### 5.5.2 Simulations with changes in parameter $b$

Similar simulations are performed with changes of  $b$ , like for the previous case. The results are similarly shown in figure 4.

### 5.5.3 Simulations with changes in both $a$ and $b$

Simulations are now performed with changes in both  $a$  and  $b$ . Ten pairs of values of  $(a, b)$  are used in these simulations. For each pair of  $(a, b)$ , 100 experiments are simulated with different random noise realizations. The results are illustrated in figure 5. Because both parameters change, no isolation rate is shown. The two minmax tests  $t_{\eta_a}^*$  and  $t_{\eta_b}^*$  have comparable values. The two sensitivity tests have similar behaviors.

## 6 Conclusion

We have proposed a new residual generation and evaluation method for FDI in nonlinear dynamic systems modeled by polynomial DAEs. The basic idea is to combine the equivalent input-output representation of DAE systems (Ljung and Glad, 1994) and the local approach to change detection (Benveniste et al., 1987; Basseville and Nikiforov, 1993).

Of course, the proposed method suffer from the limitation of the algorithm which computes the input-output representation. The computer implementation of this algorithm may have difficulties to work on systems with a large number of variables and parameters. Actually this limitation is shared by most nonlinear observer design methods (see discussions in (Nikoukhah, 1995) and references therein, among which (Walcott et al., 1987; Moraal and Grizzle, 1995)). Another practical difficulty of the proposed method is in the estimation of the time derivatives of (noise corrupted) observed variables. Nevertheless, we believe this combined method to be of potential interest for monitoring small parameter deviations in nonlinear dynamic systems.

## Acknowledgment

The authors would like to thank T. Glad and L. Ljung for having kindly provided their software described in (Ljung and Glad, 1994) and for useful comments.



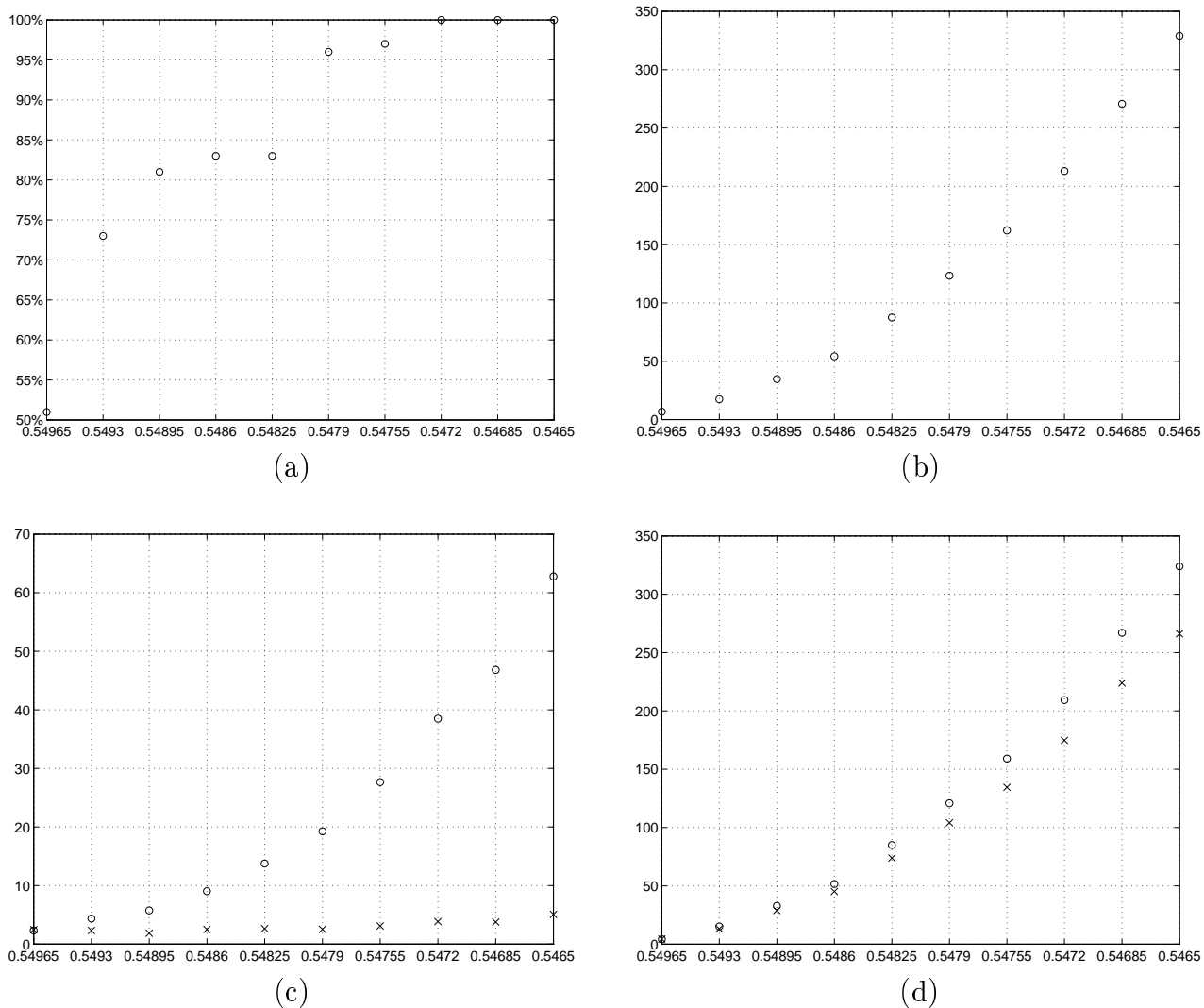


Figure 3: FI results with simulated changes in parameter  $a$ . 10 different values of  $a$  are used in the simulations. These values are shown on the abscissa of each plot. 100 experiments are performed for each of the 10 values of  $a$  with different random noise realizations. The nominal value of  $a$  is 0.55.  $b$  keeps its nominal value.

(a): The rate of successful isolation.

(b): The mean value of the global test  $t_\eta$ .

(c): The mean values of the minmax tests. The circles represent  $t_{\eta_a}^*$ , and the crosses represent  $t_{\eta_b}^*$ .

(d): The mean values of the sensitivity tests. The circles represent  $\tilde{t}_{\eta_a}$ , and the crosses represent  $\tilde{t}_{\eta_b}$ .

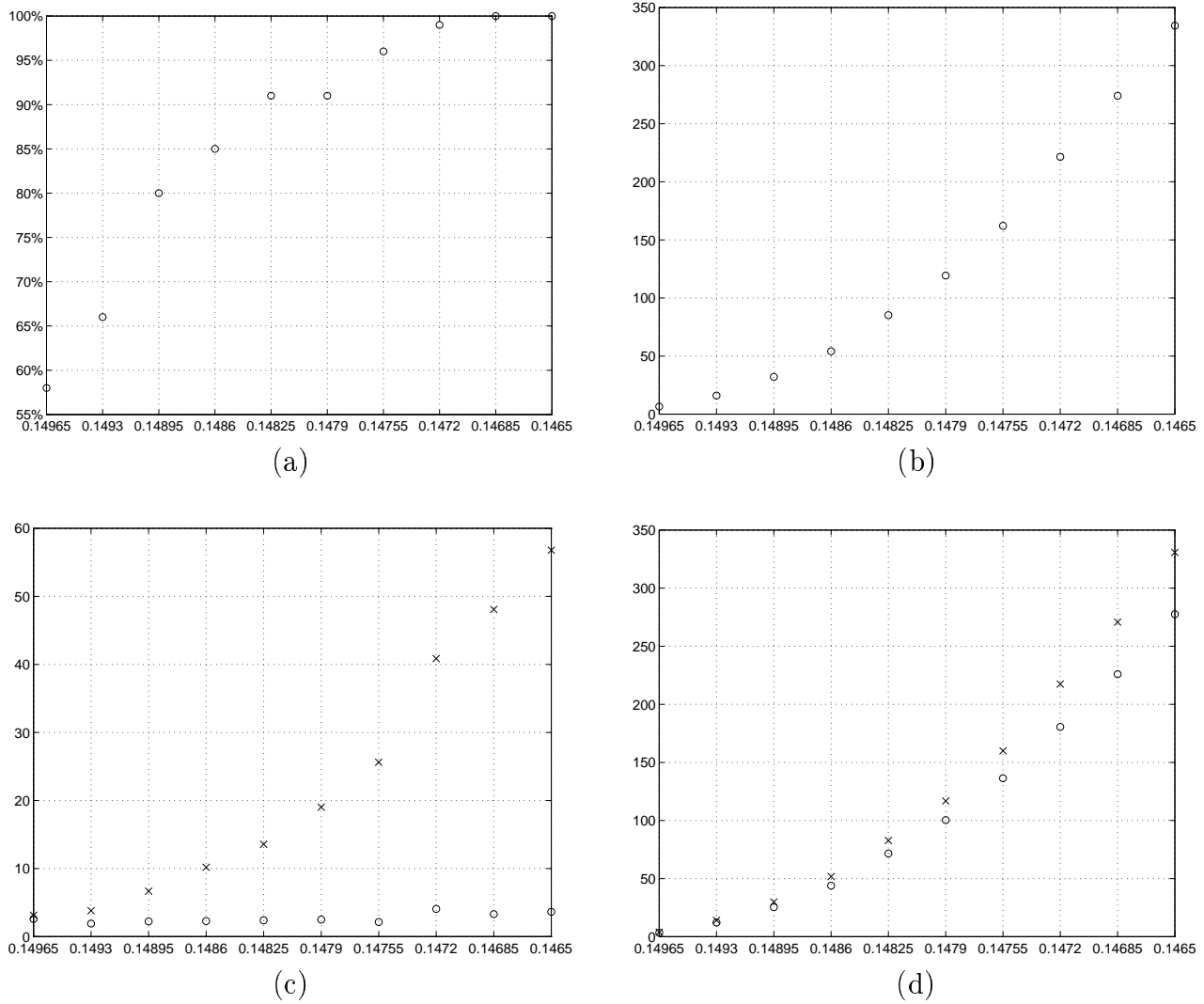


Figure 4: FI results with simulated changes in parameter  $b$ . 10 different values of  $b$  are used in the simulations. These values are shown on the abscissa of each plot. 100 experiments are performed for each of the 10 values of  $b$  with different random noise realizations. The nominal value of  $b$  is 0.15.  $a$  keeps its nominal value.

(a): The rate of successful isolation.

(b): The mean value of the global test  $t_\eta$ .

(c): The mean values of the minmax tests. The circles represent  $t_{\eta_a}^*$ , and the crosses represent  $t_{\eta_b}^*$ .

(d): The mean values of the sensitivity tests. The circles represent  $\tilde{t}_{\eta_a}$ , and the crosses represent  $\tilde{t}_{\eta_b}$ .

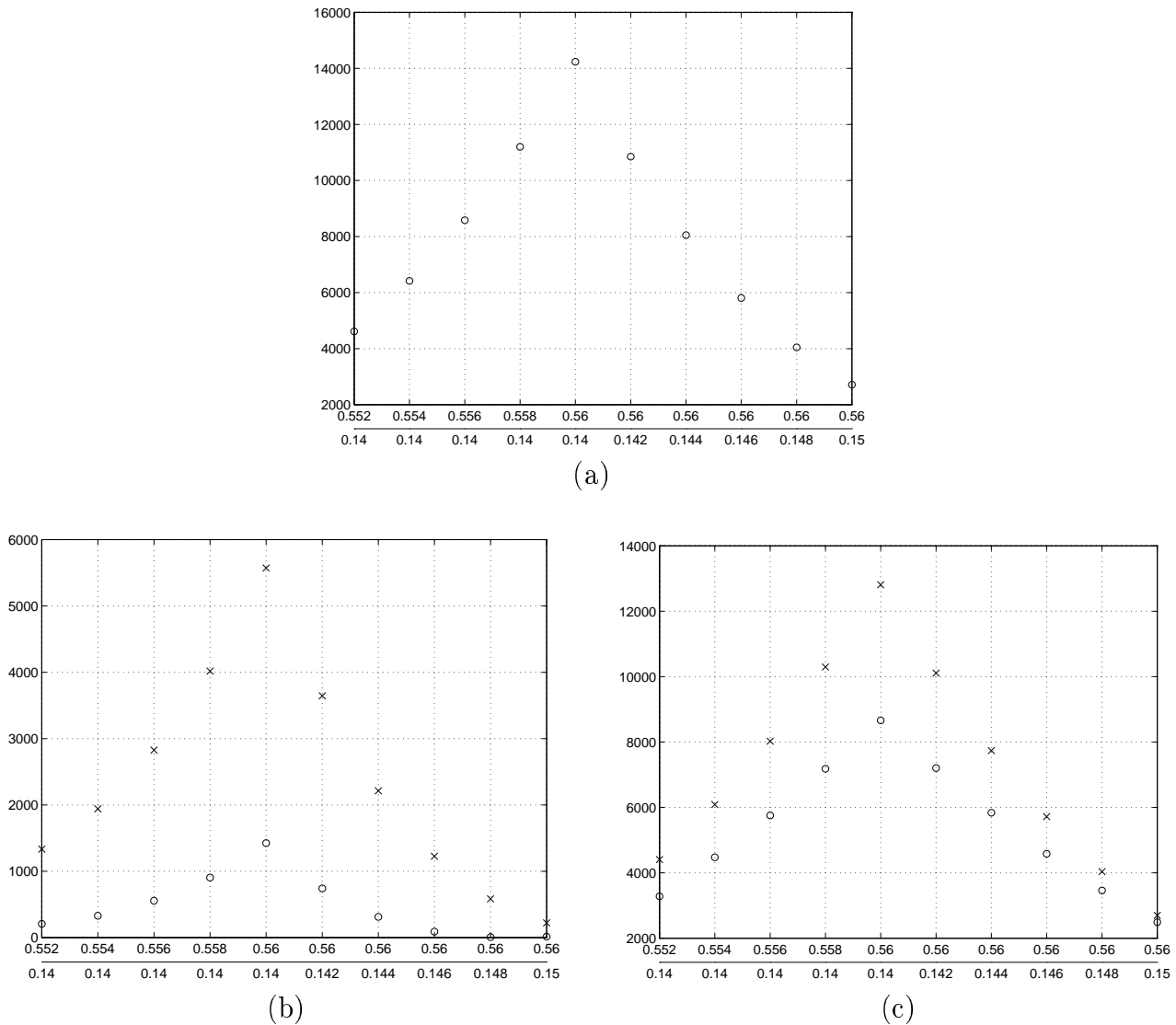


Figure 5: FI results with simulated changes in both parameters  $a$  and  $b$ . Ten pairs of values of  $(a, b)$  are used in the simulations. In each plot, the upper line on the abscissa shows the values of  $a$ , and the lower line shows the values of  $b$ . 100 experiments are performed for each of the 10 pairs of parameter values, with different random noise realizations. The nominal parameter values are  $(a, b) = (0.55, 0.15)$ .

(a): The mean value of the global test  $t_\eta$ .

(b): The mean values of the minmax tests. The circles represent  $t_{\eta_a}^*$ , and the crosses represent  $t_{\eta_b}^*$ .

(d): The mean values of the sensitivity tests. The circles represent  $\hat{t}_{\eta_a}$ , and the crosses represent  $\hat{t}_{\eta_b}$ .

## References

- Basseville, M. (1988). Detecting changes in signals and systems – a survey. *Automatica*, 24(3):309–326.
- Basseville, M. (1997). Information criteria for residual generation and fault detection and isolation. *Automatica*, 33(6).
- Basseville, M., Benveniste, A., Gach-Devauchelle, B., Goursat, M., Bonneau, D., Dorey, P., Prevosto, M., and Olagnon, M. (1993). Damage monitoring in vibration mechanics: issues in diagnostics and predictive maintenance. *Mechanical Systems and Signal Processing*, 7(5):401–423.
- Basseville, M., Benveniste, A., Mathis, G., and Zhang, Q. (1994). Monitoring the combustion set of a gas turbine. In *Proc. Safeprocess'94*, Helsinki, Finland.
- Basseville, M., Benveniste, A., Moustakides, G., and Rougée, A. (1987). Detection and diagnosis of changes in the eigenstructure of nonstationary multivariable systems. *Automatica*, 23(4):479–489.
- Basseville, M. and Nikiforov, I. (1993). *Detection of Abrupt Changes – Theory and Applications*. Information and System Sciences Serie. Prentice Hall, Englewood Cliffs, N.J.
- Benveniste, A., Basseville, M., and Moustakides, G. (1987). The asymptotic local approach to change detection and model validation. *IEEE Trans. on Automatic Control*, 32(7):583–592.
- Frank, P. M. (1990). Fault diagnosis in dynamic systems using analytical and knowledge based redundancy - a survey and new results. *Automatica*, 26:459–474.
- Gauthier, J.-P., Hammouri, H., and Othman, S. (1992). A simple observer for nonlinear systems, application to bioreactors. *IEEE Trans. on Automatic Control*, 37(6):875–880.
- Gertler, J. (1988). Survey of model-based failure detection and isolation in complex plants. *IEEE Control Systems Magazine*, 8(6):3–11.
- Goodwin, G. C., Lozano-Leal, R., Mayne, D. Q., and Middleton, R. H. (1986). Rapprochement between continuous and discrete model reference adaptive control. *Automatica*, 22(2):199–207.
- Isermann, R. (1993). Fault diagnosis of machines via parameter estimation and knowledge processing - tutorial paper. *Automatica*, 29(4):815–836.
- Juditsky, A., Hjalmarsson, H., Benveniste, A., Deylon, B., Ljung, L., Sjöberg, J., and Zhang, Q. (1995). Nonlinear black-box models in system identification: Mathematical foundations. *Automatica*, 31(12):1725–1750.
- Krener, A. J. and Duarte, A. (1996). A hybrid computational approach to nonlinear estimation. In *Proc. 35th IEEE Conference on Decision and Control (CDC'96)*, pages 1815–1819, Kobe, Japan.
- Ljung, L. and Glad, T. (1994). On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2):265–276.

- Magni, J. F. and Mouyon, P. (1994). On residual generation by observer and parity space approaches. *IEEE Trans. on Automatic Control*, 39(7):441–447.
- Monod, J. (1950). La technique de cultures continues: théorie et applications. *Ann. Inst. Pasteur*, 79(4). In Fench.
- Moraal, P. E. and Grizzle, J. W. (1995). Observer design for nonlinear systems with discrete-time measurements. *IEEE Trans. on Automatic Control*, 40(3):395–404.
- Nikiforov, I., Staroswiecki, M., and Vozel, B. (1996). Duality of analytical redundancy and statistical approach in fault diagnosis. In *Proc. IFAC'96 World Congress*, San Francisco, CA.
- Nikoukhah, R. (1995). A new methodology for observer design and implementation. Research Report 2677, Inria.
- Oppenheim, A. V., Willsky, A. S., and Young, I. T. (1983). *Signals and systems*. Signal Processing Series. Prentice Hall, London, Sydney.
- Patton, R., Frank, P., and Clark, R., editors (1989). *Fault Diagnosis in Dynamic Systems – Theory and Application*. International Series in Systems and Control Engineering. Prentice Hall International, London, UK.
- Patton, R. J. (1994). Robust model-based fault diagnosis: the state of the art. In *Proc. Safeprocess'94*, Helsinki, Finland.
- Ritt, J. (1950). *Differential Algebra*. American Mathematical Society, Providence, RI.
- Seliger, R. and Frank, P. (1991). Robust component fault detection and isolation in nonlinear dynamic systems using nonlinear unknown input observers. In *Proc. Safeprocess'91*, pages 313–318, Baden-Baden.
- Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Deylon, B., Glorennec, P.-Y., Hjalmarsson, H., and Juditsky, A. (1995). Non-linear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724.
- Spjøtvoll, E. (1972). On the optimality of some multiple comparison procedures. *Annals Statistics*, 21(3):1486–1521.
- Viswanadham, N. and Srichander, R. (1987). Fault detection using unknown input observers. *Control Theory and Advanced Technology*, 3(2):91–101.
- Walcott, B. L., Corless, M. J., and Zak, S. H. (1987). Comparative study of non-linear state observation techniques. *Int. J. Control*, 45(6):2109–2132.
- Willsky, A. S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, 12:601–611.
- Wünnenberg, J. (1990). *Observer-based Fault Detection in Dynamic Systems*. PhD thesis, Universität-Gesamthochschule-Duisburg, Germany.
- Zhang, Q. (1996). Using nonlinear black-box models in fault detection. In *Proc. 35th IEEE Conference on Decision and Control (CDC'96)*, Kobe, Japan.

Zhang, Q., Basseville, M., and Benveniste, A. (1994). Early warning of slight changes in systems. *Automatica*, 30(1):95–114. Special Issue on Statistical Methods in Signal Processing and Control.



---

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,  
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY  
Unité de recherche INRIA Rennes, Irista, Campus universitaire de Beaulieu, 35042 RENNES Cedex  
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN  
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex  
Unité de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

---

Éditeur  
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)  
<http://www.inria.fr>  
ISSN 0249-6399