Nonlinear Fault Detection and Isolation in a Three-Tank Heating System

Raffaella Mattone and Alessandro De Luca, *Senior Member, IEEE*

***Abstract*—We consider the fault detection and isolation (FDI) problem for a nonlinear dynamic plant (the IFATIS Heating System Benchmark) affected by actuator and/or sensor faults. A general procedure is proposed for modeling faults of sensors that measure the state of a nonlinear system, so that each sensor fault is typically associated to a set of (always concurrent) fault inputs and the resulting dynamic equations are affine in the introduced fault inputs. This allows the application of recently developed nonlinear FDI techniques, slightly extended to cover the considered model structure. For the presented case study, assuming nonconcurrency of the faults in the hardware components, we describe in detail the modeling procedure, the synthesis of residual generators, and the design of a combinatorial logics that recovers perfect isolation. Simulation results are reported in the presence of input and measurement noise.**

***Index Terms*—Fault detection and isolation (FDI), nonlinear systems, thermo-hydraulic systems, sensor and actuator faults.**

I. Introduction

In SUPERVISING the correct operation of dynamic plants, one is interested in the early detection and isolation of faults in the control hardware components [1], [2]. In a fault detection and isolation (FDI) system, the detection phase consists in the generation of diagnostic signals (*residuals*) triggered by a deviation of the plant from the expected behavior, based on the processing of the available commanded inputs and measured outputs and, possibly, the use of a dynamic model. The isolation phase consists in the discrimination of the occurrence of a specific fault out of a set of potential faults, which can be used for control system reconfiguration.

When a model of the nominal (faultless) plant is available, it is natural to model the presence of faults as additional inputs affecting the system dynamics. For plants with linear dynamics, several approaches to the model-based design of FDI systems have been proposed in the literature, e.g., based on Kalman filters, Luenberger observers, parity space or parameter estimation techniques (see the survey in [3]). Examples of applications include hard disk drive control systems [4], rotor/magnetic bearing systems [5], valve leakages in blast furnace processes [6], industrial gas turbines [7], and heat exchangers [8].

For systems with inherently nonlinear dynamics, the usual approach resorts to linear approximations in the FDI design, which however, limits the validity of the scheme to a chosen operating point and introduces endogenous disturbances in the dynamics of the residuals. Nonlinear dynamics has been directly considered in signal-based FDI methods using nonlinear ARMAX models (e.g., in [9] for an internal combustion engine) or neuro-fuzzy systems (see [10] for a well-stirred tank reactor), or by exploiting the linear-in-the-parameter nature of the model (as in [11] for a distillation column). Robust fault detection for general nonlinear systems has been considered in [12].

More recently, FDI techniques have been explicitly developed for nonlinear systems of the form (see [13] and [14])

$$\begin{aligned}\dot{x} &= g_0(x) + \sum_{i=1}^{m} g_i(x)u_i + \sum_{k=1}^{s} l_k(x)f_k + \sum_{j=1}^{d} n_j(x)w_j \\ y &= h(x)\end{aligned} \tag{1}$$

i.e., where the dynamics of the state $x \in I\!R^n$ (as well as the expression of the output $y \in I\!R^p$) are nonlinear in x, but *affine* in the control inputs $u_i, i = 1, \ldots, m$, in the fault inputs $f_k, k = 1, \ldots, s$, and in the disturbances $w_j, j = 1, \ldots, d$. Applications in this class include robot manipulators [15]–[17] and an industrial furnace [18]. The system vector fields $g_0, g_1, \ldots, g_m$ in (1) are assumed to be perfectly known. Furthermore, fault inputs and disturbances appear only in the state dynamics and not in the output equation. However, no assumption on the form and/or parameters of the fault time behavior is required, in general. As a consequence, FDI methods based on a model of the faulted process in the form (1) are useful to deal with failures of hardware components (e.g., actuators and/or sensors) of any type and time behavior, but not affecting the structure of the system dynamics (i.e., *system faults* are not easily treated). This is to be compared with the approach in [12], which is capable to detect and isolate system faults under model uncertainties, provided that state measurements are reliable, and that the structural change in the system dynamics can be parameterized by a class of functions out of a given set of finite cardinality.

For nonlinear systems in the form (1), differential-geometric conditions have been given in [14], that are necessary for the solution of the FDI problem with possibly concurrent faults. These conditions, however, are violated in many situations of practical interest, notably whenever the total number of fault inputs exceeds the dimension of the state space. In [19], we have proposed several ways to relax the FDI problem for system (1), when it is not solvable in the original formulation of [14]. One possibility is to introduce the additional assumption of nonconcurrency of faults, which results in much weaker necessary conditions for obtaining fault detection and isolation.

In this brief we present, through a complete case study, all steps involved in the nonlinear design of a fault detection and isolation scheme for multiple nonconcurrent actuator and state sensor faults. The process under consideration is a typical thermo-hydraulic system, also used as benchmark in the EU project *IFATIS* [20] and already taken as application problem for other approaches based on linear [21] or bilinear [22] approximations of the system dynamics. Here, we shall take into account a full nonlinear model for the faultless system. The considered faults may affect all input actuators (hydraulic

Manuscript received October 26, 2005; revised May 24, 2006. Manuscript received in final form June 6, 2006. Recommended by Associate Editor A. T. Vemuri. This work was supported by the EU project EU-IST-2001-32122 *IFATIS*.

The authors are with the Dipartimento di Informatica e Sistemistica, Università di Roma "La Sapienza," 00184 Roma, Italy (e-mail: mattone@dis.uniroma1.it; deluca@dis.uniroma1.it).

Digital Object Identifier 10.1109/TCST.2006.880221

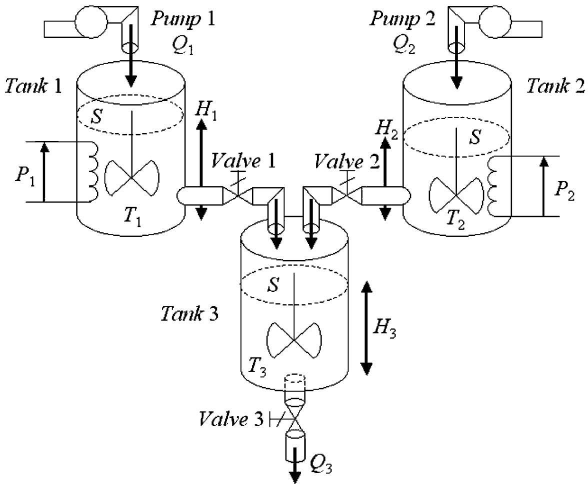

Fig. 1. Schematic diagram of the three-tank heating system.

pumps and electrical resistors) and all available state measurement sensors (level transducers and thermometers), that we assume available for all state variables.

In a preliminary modeling stage we derive, for the faulted plant, dynamic equations having the structure (1). While actuator faults of any type (power loss, bias, saturation, etc.) can be directly modeled in this form, this is not immediate for faults affecting the state sensors, due to the nonlinear dependence of the system model on the state variables. Therefore, we propose for this class of faults a new modeling procedure, where each physical sensor fault is modeled by a suitable set of (*always concurrent*) fault inputs. Correspondingly, we extend to this more general situation the necessary (and, under full state availability, also sufficient) conditions for nonconcurrent FDI, which are given in [19] for the standard case when each physical fault is modeled by just one fault input. Based on the fulfillment of these conditions of geometric nature, we illustrate the design of residual generators and of a combinatorial logics allowing perfect isolation of faults. FDI performance of the resulting hybrid scheme is demonstrated by numerical simulations, where the presence of unmodeled input and measurement noise is taken into account.

II. Three-Tank Heating System

The process used as a case study in this brief, is composed of three cylindrical tanks. According to the scheme of Fig. 1, Tanks 1 and 2 are used for preheating the fluids supplied by Pump 1 and Pump 2. The fluid temperature in these tanks can be adjusted by means of two electrical resistors. The third tank allows mixing the fluids coming from the two preheating tanks. Measures of the fluid level and temperature are available for all tanks. The control objective is to regulate the fluid level and temperature in Tank 3.

Writing the mass and energy balance equations for each of the three tanks results in a nonlinear, affine in the inputs, model of the form

$$\dot{x} = g_0(x) + \sum_{i=1}^{4} g_i(x)\, u_i. \tag{2}$$

In the absence of sensor faults, the output y coincides with the state vector $x = [H_1\, H_2\, H_3\, T_1\, T_2\, T_3]^T$, with H_i and T_i, the fluid level and, respectively, temperature in the i-th tank, $i = 1, \ldots, 3$. In (2), $u = [Q_1\, Q_2\, P_1\, P_2]^T$ is the *actually applied* input vector (which may be different from the *commanded* input u_c in case of fault), with Q_i and P_i the flow-rate and, respectively, heating power delivered to the i-th tank, $i = 1, 2$. The expression of vector fields $g_i(x), i = 0, \ldots, 4$, is given by

$$g_0 = \frac{1}{S}\begin{bmatrix} -\alpha_1\sqrt{x_1} \\ -\alpha_2\sqrt{x_2} \\ \alpha_1\sqrt{x_1} + \alpha_2\sqrt{x_2} - \alpha_3\sqrt{x_3} \\ 0 \\ 0 \\ -\frac{\alpha_1}{x_3}\sqrt{x_1}(x_6 - x_4) - \frac{\alpha_2}{x_3}\sqrt{x_2}(x_6 - x_5) \end{bmatrix}$$

$$g_1 = \frac{1}{S}\begin{bmatrix} 1 \\ 0 \\ 0 \\ -\frac{(x_4 - T_{1i})}{x_1} \\ 0 \\ 0 \end{bmatrix}, \quad g_2 = \frac{1}{S}\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -\frac{(x_5 - T_{2i})}{x_2} \\ 0 \end{bmatrix}$$

$$g_3 = \frac{1}{S\mu c}\begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{x_1} \\ 0 \\ 0 \end{bmatrix}, \quad g_4 = \frac{1}{S\mu c}\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{x_2} \\ 0 \end{bmatrix}.$$

In the above expressions, S is the section of the tanks, c the specific heat of the fluid, μ its density, T_{1i} and T_{2i}, the fluid temperature at the input of tanks 1, 2, and α_i the inverse of the (constant) throttling of the i-th output valve, $i = 1, \ldots, 3$. Note that the system vector fields $g_i(x), i = 0, \ldots, 4$, are smooth in the state-space region of interest $(x_i > 0, i = 1, \ldots, 6)$.

III. Fault Modeling

We assume that the system can be affected by faults of the actuators delivering the input commands $u_i, i = 1, \ldots, 4$, and of the sensors providing measures of the state variables $x_1, \ldots, x_6$. Hence, the total number of faults possibly affecting the system is $s = 10$. Since this number is larger than the dimension of the state space $(n = 6)$, the necessary conditions for FDI given in [14] for possibly concurrent faults are certainly (and structurally) violated. Therefore, following the approach in [19], we shall relax the FDI problem by assuming that at most one fault can affect the system at any time (*nonconcurrency* of faults).

A. Actuator Faults

Being model (2) affine in the control inputs, actuator faults can simply be modeled through the fault input vector f_u defined as

$$f_u = u - u_c \tag{3}$$

where u_c is the commanded control input. Replacing expression (3) in the system state dynamics (2), we get

$$\dot{x} = g_0(x) + \sum_{i=1}^{4} g_i(x)\, u_{c,i} + \sum_{i=1}^{4} l_{u_i}(x)\, f_{u_i} \tag{4}$$

with $l_{u_i}(x) = g_i(x)$. Equation (4) includes the effects of actuator faults and is still affine in the (control and fault) inputs, thus, allowing the direct application of the nonlinear FDI techniques in [14] and [19].

B. Sensor Faults

We focus now on the modeling of faults of the sensors providing measures of the state variables $x_1, \ldots, x_6$. The most natural way to take into account the possible occurrence of these faults would be defining the k-th measurement fault, $k = 1, \ldots, n$, as

$$F_{x_k} = x_k - y_k \tag{5}$$

i.e., as the difference between real and measured value of the k-th state variable. However, this modeling would lead either to the appearance of fault inputs in the output equation ($y = x - F_x$) or to a model that is nonlinear in the sensor fault inputs (when the state x is replaced by the expression $y + F_x$ in (2)).

We propose then a different procedure for modeling this class of faults, which results in the desired structure (1) for the system equations. As we will show, this is obtained by introducing a suitable *set* of $\nu_k \geq 1$ fault inputs $f_{x_k,i} (i = 1, \ldots, \nu_k)$, in place of the *natural* fault quantity F_{x_k} of (5). This implies that the one-to-one correspondence of (5) between physical event (fault of a sensing device) and mathematical representation (sensor fault input), is lost. For this reason, in the rest of this brief, we will often need to distinguish between the *physical* sensor fault, denoted for simplicity by F_{x_k} as in (5), and the corresponding *mathematical* fault inputs $f_{x_k,i}$. Whenever a physical sensor failure occurs (i.e., when $F_{x_k} \neq 0$), all associated fault inputs $f_{x_k,i}, i = 1, \ldots, \nu_k$ will become generically nonzero, although with completely different time behaviors and, in general, without a direct physical interpretation. According to this modeling, in order to detect and isolate a single physical fault of the k-th state sensor, it will be sufficient to recognize the occurrence of *any* (one or more) of the mathematical fault inputs $f_{x_k,i}, i = 1, \ldots, \nu_k$. This is in the spirit of isolating a given *set* of faults from the remaining ones, as formalized in [19].

For the generic k-th sensor fault ($k = 1, \ldots, n$), we propose the following modeling procedure.

1) Look in the system model for all different (and, in general, nonlinear) expressions $\varphi_{k,i}(x, u)$, involving x_k and such that the model is *affine* in $\varphi_{k,i}(x, u)$. For variable x_1, for example, it is[1]

$$\varphi_{1,1}(x_1) = \frac{\alpha_1}{S}\sqrt{x_1}$$
$$\varphi_{1,2}(x, u) = \left(-\frac{x_4 - T_{1i}}{S}u_1 + \frac{u_3}{S\mu c}\right)\frac{1}{x_1}.$$

2) For each expression $\varphi_{k,i}(x, u)$ found at step 1, define the fault input $f_{x_k,i} = \varphi_{k,i}(x, u) - \varphi_{k,i}(x, u)|_{x_k = y_k}$, i.e., the error induced in the computation of $\varphi_{k,i}(x, u)$ by the use of the measured value y_k in place of the real value x_k, and compute the corresponding fault vector field $l_{x_k,i}(x)$. Let us denote the number of faults introduced in this way by $\nu_k - 1$. Note that $f_{x_k,i}$ is, by definition, only affected by a fault of the k-th state sensor (which is consistent with the assumption of nonconcurrency[2]), and is zero whenever $F_{x_k} \equiv 0$, i.e., when $x_k = y_k$. As a result of this modeling step, any occurrence of the expression $\varphi_{k,i}(x, u)$ in the system model can be replaced by $\varphi_{k,i}(x, u)|_{x_k = y_k} + f_{x_k,i}$, and the model is certainly affine in the fault input $f_{x_k,i}$. Note that the right-hand side of (2) will now be only dependent on the variable y_k and not on x_k.
3) Define the further fault input $f_{x_k,\nu_k} = \dot{x}_k - \dot{y}_k$. The introduction of this additional fault input in the model allows writing also the left-hand side of the k-th system equation in terms of the new variable y_k (with dynamics $\dot{y}_k = \dot{x}_k - f_{x_k,\nu_k}$). The fault vector field associated to f_{x_k,ν_k} is, thus, $l_{x_k,\nu_k} = -I_k$ (I_k is the k-th column of the $n \times n$ identity matrix).
4) If, for two indices i, j, we can write $l_{x_k,i} = \alpha(x) l_{x_k,j}$ for some real function $\alpha(x)$, then set $f_{x_k,j} = f_{x_k,j} + \alpha(x) f_{x_k,i}$ and eliminate $f_{x_k,i}$ (vector field $l_{x_k,j}$ clearly remains the same). With a slight abuse of notation, we still use the symbol ν_k to indicate the final number of mathematical fault inputs corresponding to the k-th state sensor fault.

As a result of this general modeling procedure, in the case study under consideration, we introduce the following fault inputs.

- For sensor fault 1 ($\nu_1 = 3$)

$$f_{x_1,1} = \frac{\alpha_1}{S}(\sqrt{x_1} - \sqrt{y_1})$$
$$f_{x_1,2} = \left(-\frac{x_4 - T_{1i}}{S}u_1 + \frac{u_3}{S\mu c}\right)\left(\frac{1}{x_1} - \frac{1}{y_1}\right)$$
$$f_{x_1,3} = \dot{x}_1 - \dot{y}_1.$$

- For sensor fault 2 ($\nu_2 = 3$)

$$f_{x_2,1} = \frac{\alpha_2}{S}(\sqrt{x_2} - \sqrt{y_2})$$
$$f_{x_2,2} = \left(-\frac{x_5 - T_{2i}}{S}u_2 + \frac{u_4}{S\mu c}\right)\left(\frac{1}{x_2} - \frac{1}{y_2}\right)$$
$$f_{x_2,3} = \dot{x}_2 - \dot{y}_2.$$

- For sensor fault 3 ($\nu_3 = 2$)

$$f_{x_3,1} = -\frac{\alpha_3}{S}(\sqrt{x_3} - \sqrt{y_3}) - (\dot{x}_3 - \dot{y}_3)$$
$$f_{x_3,2} = -\frac{1}{S}\left(\alpha_1\sqrt{x_1}(x_6 - x_4) + \alpha_2\sqrt{x_2}(x_6 - x_5)\right) \cdot \left(\frac{1}{x_3} - \frac{1}{y_3}\right).$$

- For sensor fault 4 ($\nu_4 = 2$)

$$f_{x_4,1} = -\frac{u_1}{Sx_1}(x_4 - y_4) - (\dot{x}_4 - \dot{y}_4)$$
$$f_{x_4,2} = -\frac{\alpha_1}{Sx_3}\sqrt{x_1}(x_4 - y_4).$$

- For sensor fault 5 ($\nu_5 = 2$)

$$f_{x_5,1} = -\frac{u_2}{Sx_2}(x_5 - y_5) - (\dot{x}_5 - \dot{y}_5)$$
$$f_{x_5,2} = -\frac{\alpha_2}{Sx_3}\sqrt{x_2}(x_5 - y_5).$$

[1]In this case, the quantities $\varphi_{k,i}(x, u)$ can be easily given a physical interpretation. In particular, $\varphi_{1,1}(x_1)$ is the output flow from Tank 1 (through Valve 1), while $\varphi_{1,2}(x, u)$ is the net power heating Tank 1.

[2]In the absence of this assumption, we should define a different fault input for each combination of faulty/faultless devices.

- For sensor fault 6 ($\nu_6 = 1$)

$$f_{x_6,1} = -\frac{1}{Sx_3}\left(\alpha_1\sqrt{x_1} + \alpha_2\sqrt{x_2}\right)(x_6 - y_6) - (\dot{x}_6 - \dot{y}_6).$$

At this point, when the outputs $y_1, \ldots, y_6$ are taken as new state variables for the system dynamics, the general structure (1) is recovered. The final model, including the effect of all (nonconcurrent) faults of actuators and state sensors, is then

$$\dot{y} = g_0(y) + \sum_{i=1}^{4} g_i(y)\, u_{c,i} + \sum_{i=1}^{4} l_{u_i}(y)\, f_{u_i} + \sum_{k=1}^{6}\sum_{j=1}^{\nu_k} l_{x_k,j}(y) f_{x_k,j} \tag{6}$$

with the trivial output equation omitted. Equation (6), which models the faulted system, is expressed in terms of the available commanded inputs u_c and measured outputs y, and is affine in all control and (unknown) fault inputs, as requested. In other words, any discrepancy between faultless and faulted system dynamics is fully summarized within the introduced fault inputs. Model (6) may be reordered and more compactly rewritten as

$$\dot{y} = g_0(y) + \sum_{i=1}^{m} g_i(y)\, u_{c,i} + \sum_{k=1}^{s}\sum_{j=1}^{\mu_k} l_{k,j}(y) f_{k,j} \tag{7}$$

where $m = 4, s = 10$, and with $f_{k,1} = f_{u_k}$ and $\mu_k = 1$ for $k = 1, \ldots, 4$, and $f_{4+i,j} = f_{x_i,j}$ and $\mu_{4+i} = \nu_i$ for $i = 1, \ldots, 6$. The expression of fault vector fields $l_{k,j}, k = 1, \ldots, 10, j = 1, \ldots, \mu_k$, is given by $l_{k,1} = l_{u_k} = g_k, k = 1, \ldots, 4$, and

$$\begin{aligned}
l_{5,1} &= l_{x_1,1} = \begin{bmatrix} -1 & 0 & 1 & 0 & 0 & -\frac{y_6 - y_4}{y_3} \end{bmatrix}^T \\
l_{5,2} &= l_{x_1,2} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^T \\
l_{5,3} &= l_{x_1,3} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T \\
l_{6,1} &= l_{x_2,1} = \begin{bmatrix} 0 & -1 & 1 & 0 & 0 & -\frac{y_6 - y_5}{y_3} \end{bmatrix}^T \\
l_{6,2} &= l_{x_2,2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}^T \\
l_{6,3} &= l_{x_2,3} = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix}^T \\
l_{7,1} &= l_{x_3,1} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}^T \\
l_{7,2} &= l_{x_3,2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \\
l_{8,1} &= l_{x_4,1} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^T \\
l_{8,2} &= l_{x_4,2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \\
l_{9,1} &= l_{x_5,1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}^T \\
l_{9,2} &= l_{x_5,2} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \\
l_{10,1} &= l_{x_6,1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T.
\end{aligned}$$

Note that input and measurement noise are not explicitly included in model (7) used for FDI design. Indeed, it is rather intuitive (and can be formally proven) that these disturbances cannot be exactly distinguished from faults of the input actuators and state sensors. On the other hand, when these unmodeled disturbances are "small enough," relevant faults can still be correctly isolated, in practice, by suitably thresholding residuals (we propose a simple but effective method in Section VI).

IV. Necessary and Sufficient Conditions for FDI

Introducing the additional assumption of nonconcurrency of faults, results into much weaker conditions for FDI than those given in [14]. In particular, when each physical fault F_k is modeled by just one fault input with associated vector field l_k, we have shown in [19], that the necessary and sufficient condition for nonconcurrent FDI (under full state availability and absence of disturbances) is

$$\mathrm{span}\{l_i\} \not\subseteq \mathrm{span}\{l_k\}, \qquad \forall i, k, \qquad i \neq k. \tag{8}$$

Condition (8) guarantees that, for each couple of faults F_i, F_k, a dynamic system (*residual generator*) can be found, whose output r (*residual*) is affected by just one of the two faults and not by the other.

When each physical fault F_k is modeled by a set of *always concurrent* fault inputs, as in the case of (7), condition (8) can be easily extended by observing that a residual i) is affected by F_k, if it is affected by *at least one* of the associated fault inputs $f_{k,1}, \ldots, f_{k,\mu_k}$ and ii) is decoupled from F_k, if it is not affected by *any* of the fault inputs $f_{k,1}, \ldots, f_{k,\mu_k}$. This leads to the following necessary and sufficient condition:

$$\begin{aligned}
\forall k, \forall i \neq k, \exists j \in \{1, \ldots, \mu_k\} &: \mathrm{span}\{l_{k,j}\} \not\subseteq \bar{P}_i \\
\text{OR } \exists h \in \{1, \ldots, \mu_i\} &: \mathrm{span}\{l_{i,h}\} \not\subseteq \bar{P}_k
\end{aligned} \tag{9}$$

where $P_i = \mathrm{span}\{l_{i,1}, \ldots, l_{i,\mu_i}\}$ and $\bar{P}_i$ denotes the involutive closure of P_i, i.e., the closure of P_i under the Lie bracket operator.[3] Note that, differently from (8), which is symmetric with respect to i and k, the two conditions in the left- and right-hand sides of the "OR" operator in (9) may not hold at the same time. Thus, it may happen that a residual generator exists, that is affected by F_i and not by F_k, but that any residual affected by F_k is necessarily also affected by F_i.

In the case under consideration, where $(F_1, \ldots, F_4) = (f_{u_1}, \ldots, f_{u_4})$ and $(F_5, \ldots, F_{10}) = (F_{x_1}, \ldots, F_{x_6})$, it can be readily verified that all distributions P_i are involutive (thus, $\bar{P}_i = P_i$) and that (9) holds for any $k, i = 1, \ldots, 10, i \neq k$.

V. Residual Generator Design

The fulfillment of condition (9) implies that i) for any pair of faults F_i and F_k, a residual exists that is affected by F_i or by F_k, but not by both and ii) there exists a set R of residuals such that each fault affects a different, nonempty subset of diagnostic signals within R.

Under the assumption of full state measurability, the design of residuals at point 1) follows directly from (9), using the fault vector fields of model (7), as shown in detail in the Appendix. Essentially, condition (9) guarantees the existence, for any pair (i, k), of a suitable *decoupling* output function $Y = \psi(y)$, whose dynamics is affected by at least one of the fault inputs associated to, e.g., F_i, and decoupled from all fault inputs corresponding to F_k. Then, a residual generator can be designed as a standard nonlinear observer with linear error dynamics [23]. In particular, the dynamics of the state ξ of this observer/residual generator is a copy of the nominal (faultless) dynamics of Y, plus a correction term $K(Y - \xi)$ that makes the observation error/residual $r = Y - \xi$ asymptotically converge to zero in the absence of faults.

[3]The Lie bracket of two vector fields, $v_1(x)$ and $v_2(x)$, is defined as $[v_1, v_2] = (\partial v_2/\partial x)v_1 - (\partial v_1/\partial x)v_2$ [23].

TABLE I

Residual generator RG_1 $(Y_1 = y_1)$
$\dot{\xi}_1 = -\frac{\alpha_1}{S}\sqrt{y_1} + \frac{1}{S}u_{c1} + K_1\,(Y_1 - \xi_1)$ $r_1 = K_1\,(Y_1 - \xi_1)$
Dynamics of residual r_1
$\dot{r}_1 = -K_1\,r_1 + \frac{K_1}{S}f_{u_1} - K_1\,f_{x_1,1} - K_1\,f_{x_1,3}$

TABLE II

Residual generator RG_2 $(Y_2 = y_2)$
$\dot{\xi}_2 = -\frac{\alpha_2}{S}\sqrt{y_2} + \frac{1}{S}u_{c2} + K_2\,(Y_2 - \xi_2)$ $r_2 = K_2\,(Y_2 - \xi_2)$
Dynamics of residual r_2
$\dot{r}_2 = -K_2\,r_2 + \frac{K_2}{S}f_{u_2} - K_2\,f_{x_2,1} - K_2\,f_{x_2,3}$

TABLE III

Residual generator RG_3 $(Y_3 = y_1(y_4 - T_{1i}))$
$\dot{\xi}_3 = -\frac{\alpha_1}{S}\,(y_4 - T_{1i})\,\sqrt{y_1} + \frac{1}{S\mu c}u_{c3} + K_3\,(Y_3 - \xi_3)$ $r_3 = K_3\,(Y_3 - \xi_3)$
Dynamics of residual r_3
$\dot{r}_3 = -K_3\,r_3 + \frac{K_3}{S\mu c}\,f_{u_3} - K_3\,(y_4 - T_{1i})\,(f_{x_1,1} + f_{x_1,3})$ $+K_3\,y_1\,(f_{x_1,2} + f_{x_4,1})$

TABLE IV

Residual generator RG_4 $(Y_4 = y_2(y_5 - T_{2i}))$
$\dot{\xi}_4 = -\frac{\alpha_2}{S}\,(y_5 - T_{2i})\,\sqrt{y_2} + \frac{1}{S\mu c}u_{c4} + K_4\,(Y_4 - \xi_4)$ $r_4 = K_4\,(Y_4 - \xi_4)$
Dynamics of residual r_4
$\dot{r}_4 = -K_4\,r_4 + \frac{K_4}{S\mu c}f_{u_4} - K_4\,(y_5 - T_{2i})\,(f_{x_2,1} + f_{x_2,3})$ $+K_4\,y_2\,(f_{x_2,2} + f_{x_5,1})$

In order to build the solution set of residuals at point 2), it is convenient to work with residual matrices. Recall that the $s \times \sigma$ *residual matrix* RM associated to the set of residuals $R = \{r_1, \ldots, r_\sigma\}$ is the binary matrix whose element $\mathrm{RM}(i,j)$ is nonzero if and only if the physical fault F_i (or, equivalently, at least one of the associated fault inputs $f_{i,1}, \ldots, f_{i,\mu_i}$) affects residual $r_j \in R$. Therefore, a set R of residuals solves the nonconcurrent FDI problem if and only if

$$\mathrm{RM}(i,:) \neq 0, \qquad \mathrm{RM}(k,:) \neq \mathrm{RM}(i,:) \qquad \forall i,k, \qquad i \neq k \tag{10}$$

where $\mathrm{RM}(i,:)$ denotes the ith row of RM.

The following procedure incrementally builds the columns of a residual matrix RM satisfying (10).

TABLE V

Residual generator RG_5 $(Y_5 = y_3)$
$\dot{\xi}_5 = \frac{1}{S}\,(\alpha_1\sqrt{y_1} + \alpha_2\sqrt{y_2} - \alpha_3\sqrt{y_3}) + K_5\,(Y_5 - \xi_5)$ $r_5 = K_5\,(Y_5 - \xi_5)$
Dynamics of residual r_5
$\dot{r}_5 = -K_5\,r_5 + K_5\,f_{x_1,1} + K_5\,f_{x_2,1} + K_5\,f_{x_3,1}$

TABLE VI

Residual generator RG_6 $(Y_6 = y_3 y_6)$
$\dot{\xi}_6 = \frac{\alpha_1}{S}\sqrt{y_1}y_4 + \frac{\alpha_2}{S}\sqrt{y_2}y_5 + \frac{\alpha_3}{S}\sqrt{y_3}y_6 + K_6\,(Y_6 - \xi_6)$ $r_6 = K_6\,(Y_6 - \xi_6)$
Dynamics of residual r_6
$\dot{r}_6 = -K_6\,r_6 + K_6\,y_4 f_{x_1,1} + K_6\,y_5 f_{x_2,1} + K_6\,y_6 f_{x_3,1}$ $+K_6\,y_3\,(f_{x_3,2} + f_{x_4,2} + f_{x_5,2} + f_{x_6,1})$

TABLE VII
RESIDUAL MATRIX/ISOLATION LOGICS FROM PHYSICAL FAULTS $F_1, \ldots, F_{10}$ TO RESIDUALS $r_1, \ldots, r_6$ OF TABLES I–VI

residual / fault	r_1	r_2	r_3	r_4	r_5	r_6
$F_1 = f_{u_1}$	1	0	0	0	0	0
$F_2 = f_{u_2}$	0	1	0	0	0	0
$F_3 = f_{u_3}$	0	0	1	0	0	0
$F_4 = f_{u_4}$	0	0	0	1	0	0
$F_5 = F_{x_1}$	1	0	1	0	1	1
$F_6 = F_{x_2}$	0	1	0	1	1	1
$F_7 = F_{x_3}$	0	0	0	0	1	1
$F_8 = F_{x_4}$	0	0	1	0	0	1
$F_9 = F_{x_5}$	0	0	0	1	0	1
$F_{10} = F_{x_6}$	0	0	0	0	0	1

- Start by designing the first element r_1 of the residual set R so that it discriminates fault F_1 from F_2, and build the corresponding one-column residual matrix RM. After this first step, the first two rows of RM are different, but one of them is zero.
- Design residual r_2 so that the first two rows of the resulting two-column RM become nonzero.[4] From now on, the first two rows of RM will certainly be different and nonzero for whatever selection of the remaining residuals.
- After the generic k-th step of the procedure, the first k rows of the current RM matrix are different and nonzero. Look then at rows $k+1$ to s of RM and find (if it exists) the first row being either zero, or equal to one of the rows above (and, necessarily, only to one). Then, the residual r_{k+1} to be appended to the set R can be designed so as to introduce either a "1" element in the null row, or the needed difference between the two equal rows.

[4]This means finding a residual that is affected by that physical fault, between F_1 and F_2, that does not affect r_1. This is always possible if all considered faults affect the system.

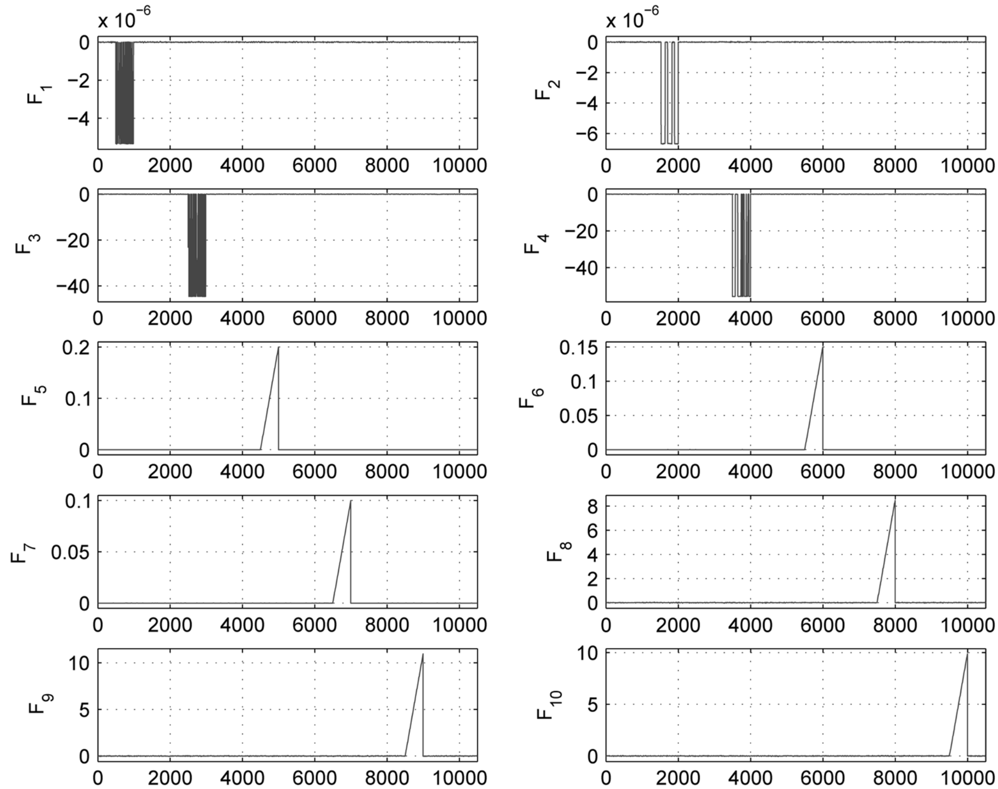

Fig. 2. Behavior of physical faults $F_1, \ldots, F_{10}$

- In a finite number σ of steps, with $\log_2 s \le \sigma \le s$, all s rows of RM will certainly be different and nonzero, and condition (10) will be satisfied by the σ designed residuals.

We have applied the described procedure to the case study under consideration. Tables I–VI contain the residuals designed according to the differential-geometric method described in the Appendix, while Table VII reports the corresponding residual matrix. For each residual r_i, we also report in the corresponding table the expression of the decoupling output $Y_i = \psi_i(y)$ and that of the resulting residual dynamics. A number $\sigma = 6 < 10 = s$ of residuals were sufficient to make all rows of the residual matrix RM in Table VII nonzero and mutually different. Finally, note that matrix RM can also be used to set up a combinatorial *isolation logics* mapping the excited/not excited residuals $r_i, i = 1, \ldots, 6$, to the boolean signals $r_{u_i} (i = 1, \ldots, 4)$ and $r_{x_j} (j = 1, \ldots, 6)$ that are one-to-one related to the physical faults of the actuators and of the state sensors.

VI. Numerical Simulations

In order to verify the performance of the proposed FDI scheme on the considered benchmark system, simulations have been performed with the following physical data:

$$\begin{aligned}
\alpha_1 &= 2.108 \cdot 10^{-5}\,\mathrm{m}^{\frac{5}{2}}/\mathrm{s}\\
\alpha_2 &= 3.043 \cdot 10^{-5}\,\mathrm{m}^{\frac{5}{2}}/\mathrm{s}\\
\alpha_3 &= 6.708 \cdot 10^{-5}\,\mathrm{m}^{\frac{5}{2}}/\mathrm{s}\\
S &= 0.0154\,\mathrm{m}^2\\
c &= 4.18 \cdot 10^3\,\mathrm{J/kg}\\
\mu &= 1.0 \cdot 10^3\,\mathrm{kg/m}^3\\
T_{1i} &= 15\,^\circ\mathrm{C}\\
T_{2i} &= 20\,^\circ\mathrm{C}.
\end{aligned}$$

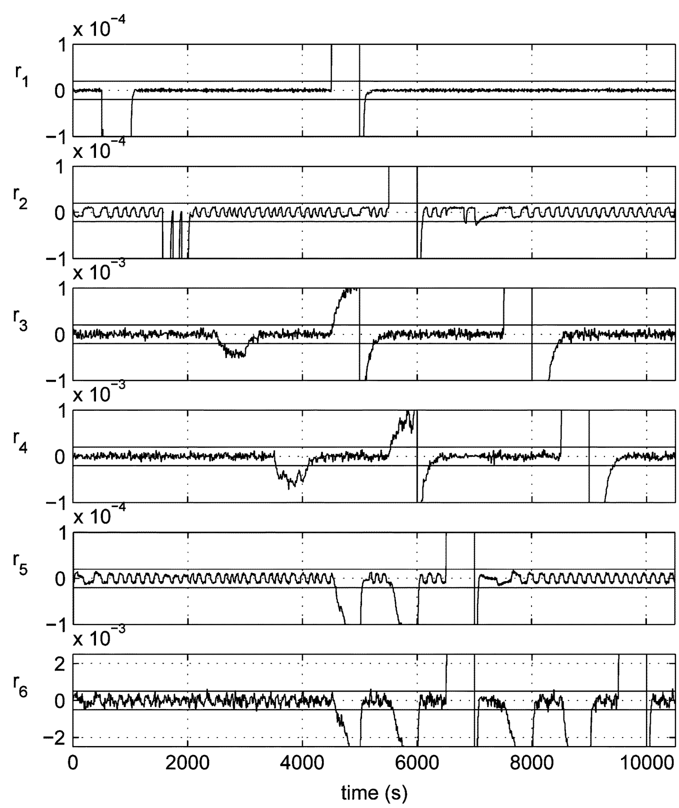

Fig. 3. Behavior of residuals $r_1, \ldots, r_6$ (vertical scale has been truncated and detection thresholds are included).

The system is controlled by a standard nonlinear state feedback law that exactly linearizes the system dynamics and decouples

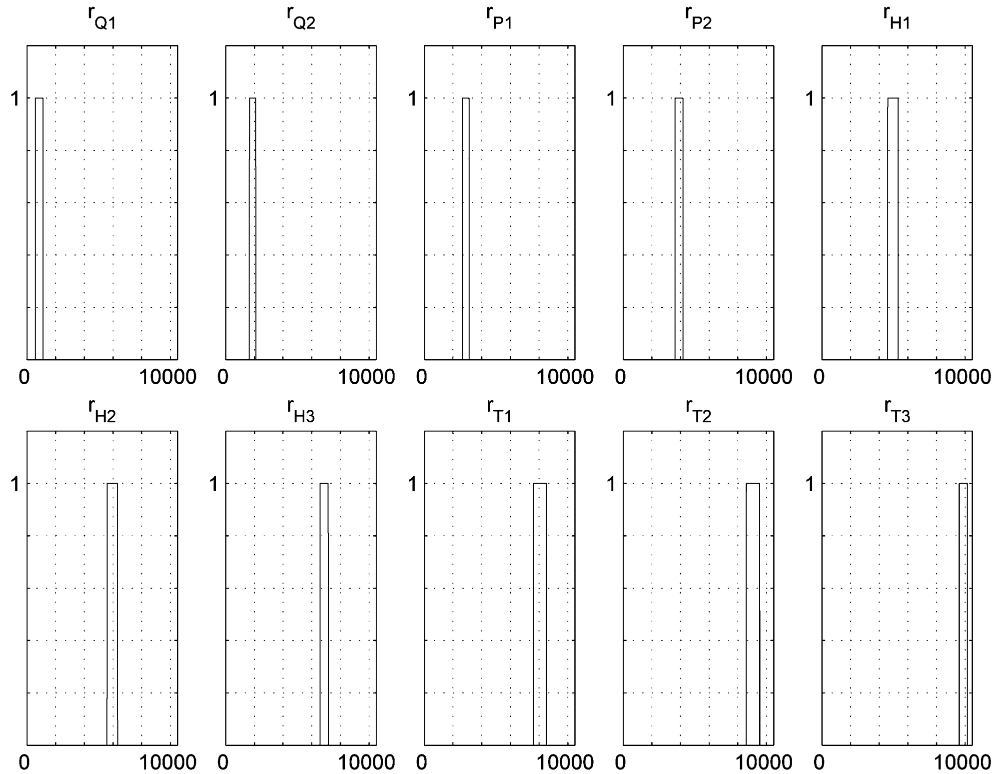

Fig. 4. FDI residuals after dynamic thresholding and isolation logics.

the input-output behavior, having set as controlled outputs z the fluid levels and temperatures in Tank 1 and Tank 3 [i.e., $z = (x_1, x_3, x_4, x_6)$]. On the linear side, regulation is obtained by four scalar P or PD controllers. The constant reference values for the outputs are $H_{10} = 0.4$ m, $H_{30} = 0.2$ m, $T_{10} = 17$ °C and $T_{30} = 19.78$ °C, respectively. According to model (2), the corresponding steady state values for x_2 and x_5 are $H_{20} = 0.3$ m and $T_{20} = 22$ °C, respectively. For each control input $u_i, i = 1, \dots, 4$, the range of commanded values has been limited to the interval $[0, 2u_{i0}]$, being u_{i0} the nominal value of u_i at the desired equilibrium configuration

$$\begin{aligned} u_{10} &= 1.333 \cdot 10^{-5}\,\mathrm{m}^3/\mathrm{s} \\ u_{20} &= 1.667 \cdot 10^{-5}\,\mathrm{m}^3/\mathrm{s} \\ u_{30} &= 111.467\,\mathrm{W} \\ u_{40} &= 139.333\,\mathrm{W}. \end{aligned}$$

The system is initialized far from the desired equilibrium.

The considered faults are power losses of 20% for the input actuators and linear drifts of $x_k^0 \cdot 10^{-3}$ units/s on the measures y_k of the state variables $x_k, k = 1, \dots, 6$, being x_k^0 the desired steady-state value of x_k. Each physical fault $F_i, i = 1, \dots 10$, is active for $t \in [i - 0.5, i] \cdot 1000$ s, so that the assumption of nonconcurrency holds (see Fig. 2). Simulation were conducted in the presence of realistic levels of input and measurement white noise (similar to those experimentally validated for a simpler plant within the *IFATIS* project [20]). The obtained numerical results are reported in Figs. 3–4.

From the residuals $r_1, \dots, r_6$, shown in Fig. 3, it can be observed that, in agreement with Table VII, each residual is affected by more than one physical fault, actuator faults excite one and only one of the residuals $r_1, \dots, r_4$, while sensor faults excite multiple residuals. Note that, since unmodeled input and measurement noise affect the system, the designed residuals are not perfectly zero in the absence of faults. In order to filter out this "background noise" from the diagnostic signals so as to allow a reliable detection of faults, we have implemented a *dynamic thresholding* mechanism: each residual is treated as zero unless it has been above a fixed threshold for at least a given time T_{set}. When this happens, the occurrence of a fault is recognized. Similarly, after a fault diagnosis, residuals are considered "unaffected" again (signalizing the fault end) only after they have been below their respective thresholds for a given time T_{reset}. At the cost of a small delay in the diagnosis, this mechanism allows the use of relatively low detection thresholds (corresponding to a good sensitivity to faults) without getting unacceptable false-alarm rates. This is confirmed by the reported isolation results (see Fig. 4), obtained for $T_{\mathrm{set}} = T_s$ and $T_{\mathrm{reset}} = 2T_s$, being $T_s = 10$ s the sampling time in the computation of residuals. The thresholds used for residuals $r_1, \dots, r_6$ are $(0.2, 0.2, 2, 2, 0.2, 5) \cdot 10^{-4}$, also displayed in Fig. 3. This choice depends, in general, on the level of noise affecting the system, but also on the observer gains K_i in the residual generators. In the reported simulations, we have used the values $(0.05, 0.05, 0.01, 0.01, 0.05, 0.05)$ for the constants $K_i, i = 1, \dots, 6$ in Tables I–VI. Guidelines for an optimal choice of these parameters may be provided by the theory of Kalman filtering [7].

VII. Conclusion

We have presented a complete case study for the design of a nonlinear FDI system, using the *IFATIS* Heating System Benchmark in the presence of nonconcurrent faults of all actuators and sensors.

A method has been proposed for modeling faults of the sensors measuring the state of a nonlinear system, so as to obtain state equations that are affine in all (control and fault) inputs, and output equations that are independent of fault inputs—a convenient form for the application of differential-geometric methods. As a result, for any physical sensor fault, a set of always concurrent fault inputs is typically introduced in the model used for FDI design. This modeling approach covers situations where performance degradation can be attributed to faults of hardware components (sensors and/or actuators) and not to changes/uncertainties in the structure of the system dynamics.

The problem of detecting and isolating any single sensor and actuator fault has been formulated and solved extending recent results on *fault set* detection and isolation [19], without any restricting assumption on the type and time behavior of the faults. Residual generators have been designed as a bank of dynamic systems with the structure of nonlinear observers. The obtained diagnostic signals are processed by a suitable combinatorial logics, which allows perfect isolation of faults (in the absence of noise). The obtained FDI system is, thus, a hybrid analog/digital system.

The presence of unmodeled input and measurement noise has been handled though a dynamic thresholding mechanism. Realistic simulations showed good performance in terms of false positive/negative alarms. In general, the few design parameters of the proposed FDI scheme can be simply tuned according to the actual signal–to–noise ratio so to attain the desired tradeoff between detection reliability and fault sensitivity.

APPENDIX A
DIFFERENTIAL-GEOMETRIC FDI DESIGN

With reference to system (7), we briefly recall here how to design, under the assumption of full state measurability, a residual generator that discriminates the occurrence of fault F_i from F_k, i.e., whose output is certainly affected by F_i or by F_k, but not by both. This constitutes the elementary step in the design procedure of Section V. Assume that condition (9) is fulfilled and, in particular

$$\exists h \in \{1,\ldots,\mu_i\} : \operatorname{span}\{l_{i,h}\} \not\subseteq \bar{P}_k. \tag{11}$$

Since $\bar{P}_k$ is involutive by construction, setting $d = \dim(\bar{P}_k)$, $n-d$ independent functions $\psi_{k,1}(y),\ldots,\psi_{k,n-d}(y)$ certainly exist, such that their differentials $d\psi_{k,1}(y),\ldots, d\psi_{k,n-d}(y)$ span the annihilating codistribution $\bar{P}_k^{\perp}$ (see, e.g., [23]), i.e.,

$$\frac{\partial \psi_{k,j}}{\partial y}\, v = 0 \qquad \forall j = 1,\ldots,n-d,\ \forall v \in \bar{P}_k. \tag{12}$$

Furthermore, from (11) it follows:

$$\exists \ell \in \{1,\ldots,n-d\} : \frac{\partial \psi_{k,\ell}}{\partial y}\, l_{i,h} \neq 0. \tag{13}$$

Defining the scalar output $Y = \psi_{k,\ell}(y)$, its evolution is

$$\dot{Y} = \frac{\partial \psi_{k,\ell}}{\partial y}\left(g_0(y) + \sum_{j=1}^{m} g_j(y)\, u_j + \sum_{p=1}^{s}\sum_{q=1}^{\mu_p} l_{p,q}(y) f_{p,q}\right)$$

and, thus, is affected by the physical fault F_i (in particular, by the associated fault input $f_{i,h}$), and not affected by F_k (since (12) holds, in particular, for $v = l_{k,q}, q = 1,\ldots,\mu_k$). Then, a residual generator that discriminates F_i from F_k is

$$\begin{aligned}\dot{\xi} &= \frac{\partial \psi_{k,\ell}}{\partial y}\left(g_0(y) + \sum_{j=1}^{m} g_j(y)\, u_j\right) + K(Y-\xi)\\ r &= K(Y-\xi)\end{aligned} \tag{14}$$

with $K > 0$, that has the structure of a nonlinear observer with scalar state ξ and linear error dynamics [23]. The associated dynamics of the residual r is

$$\dot{r} = -Kr + K\sum_{p=1}^{s}\sum_{q=1}^{\mu_p} \frac{\partial \psi_{k,\ell}}{\partial y}\, l_{p,q}(y) f_{p,q}$$

i.e., that of a linear, exponentially stable system driven by the set of all fault inputs $f_{p,q}$ for which it holds

$$\frac{\partial \psi_{k,\ell}}{\partial y}\, l_{p,q}(y) \neq 0.$$

By construction, the set of these fault inputs certainly includes $f_{i,h}$ and does not include any $f_{k,j}, j = 1,\ldots,\mu_k$. Note, finally, that all other faults *may* or *may not* affect this residual r.

REFERENCES

[1] S. Simani, C. Fantuzzi, and R. Patton, *Model-Based Fault Diagnosis in Dynamic Systems Using Identification Techniques*. London, U.K.: Springer-Verlag, 2002.

[2] F. Caccavale and L. Villani, Eds., "Fault Diagnosis and Fault Tolerance for Mechatronics Systems," in *STAR*. Berlin, Germany: Springer-Verlag, 2003, vol. 1.

[3] P. Frank, "Diagnosis in dynamic systems using analytical and knowledge-based redundancy—A survey," *Automatica*, vol. 26, pp. 459–474, 1990.

[4] D.-S. Hwang, S.-C. Peng, and P.-L. Hsu, "An integrated control/diagnostic system for a hard disk drive," *IEEE Trans. Contr. Syst. Technol.*, vol. 2, no. 4, pp. 318–326, Dec. 1994.

[5] I. Cade, P. Keogh, and M. Sahinkaya, "Fault identification in rotor/magnetic bearing systems using discrete time wavelet coefficients," *IEEE/ASME Trans. Mechatronics*, vol. 10, no. 6, pp. 648–657, Dec. 2005.

[6] A. Johansson and A. Medvedev, "Model-based leakage detection in a pulverized coal injection vessel," *IEEE Trans. Contr. Syst. Technol.*, vol. 7, no. 6, pp. 675–682, Nov. 1999.

[7] S. Simani, C. Fantuzzi, and S. Beghelli, "Diagnosis techniques for sensor faults of industrial processes," *IEEE Trans. Contr. Syst. Technol.*, vol. 8, no. 5, pp. 848–855, Sep. 2000.

[8] Y. Peng, A. Youssouf, P. Arte, and M. Kinnaert, "A complete procedure for residual generation and evaluation with application to a heat exchanger," *IEEE Trans. Contr. Syst. Technol.*, vol. 5, no. 6, pp. 542–555, Nov. 1997.

[9] E. Laukonen, K. Passino, V. Krishnaswami, G.-C. Luh, and G. Rizzoni, "Fault detection and isolation for an experimental internal combustion engine via fuzzy identification," *IEEE Trans. Contr. Syst. Technol.*, vol. 3, no. 3, pp. 347–355, Jun. 1995.

[10] Y. Maki and K. Loparo, "A neural-network approach to fault detection and diagnosis in industrial processes," *IEEE Trans. Contr. Syst. Technol.*, vol. 5, no. 6, pp. 529–541, Nov. 1997.

[11] B. Huang, "Detection of abrupt changes of total least squares models and application in fault detection," *IEEE Trans. Contr. Syst. Technol.*, vol. 9, no. 2, pp. 357–367, Mar. 2001.

[12] A. Vemuri and M. Polycarpou, "On-line approximation based methods for robust fault detection," in *Proc. 13th IFAC World Congress*, 1996, pp. 319–324.

[13] H. Hammouri, M. Kinnaert, and E. El Yaagoubi, "Observer-based approach to fault detection and isolation for nonlinear systems," *IEEE Trans. Autom. Contr.*, vol. 44, no. 10, pp. 1878–1884, Oct. 1999.

[14] C. De Persis and A. Isidori, "A geometric approach to nonlinear fault detection and isolation," *IEEE Trans. Autom. Contr.*, vol. 46, no. 6, pp. 853–865, Jun. 2001.

[15] W. Dixon, I. Walker, D. Dawson, and J. Hartranft, "Fault detection for robot manipulators with parametric uncertainty: A prediction-error-based approach," *IEEE Trans. Robot. Autom.*, vol. 16, no. 6, pp. 689–699, Dec. 2000.

[16] A. De Luca and R. Mattone, "Actuator fault detection and isolation using generalized momenta," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2003, pp. 634–639.

[17] M. McIntyre, W. Dixon, D. Dawson, and I. Walker, "Fault identification for robot manipulators," *IEEE Trans. Robot.*, vol. 21, no. 5, pp. 1028–1034, Oct. 2005.

[18] D.-L. Yu, "Diagnosing simulated faults for an industrial furnace based on bilinear model," *IEEE Trans. Contr. Syst. Technol.*, vol. 8, no. 3, pp. 435–442, Jun. 2000.

[19] R. Mattone and A. De Luca, "Relaxed fault detection and isolation: An application to a nonlinear case study," *Automatica*, vol. 42, no. 1, pp. 109–116, 2006.

[20] D. Sauter, F. Hamelin, and S. Lèger, Heating system benchmark for fault detection and isolation and fault tolerant control IFATIS, University of Nancy, Nancy, France, Rep. IFAN012R01, 2003.

[21] S. Lèger, F. Hamelin, and D. Sauter, "Fault detection and isolation in dynamic systems using principal component analysis—application to an heating system benchmark," presented at the 5th IFAC Symp. Fault Detection, Supervision Safety Tech. Process., Washington, DC, 2003.

[22] L. El Bahir and M. Kinnaert, "Fault detection and isolation for a three tank system based on a bilinear model of the supervised process," in *Proc. UKACC Int. Conf. Contr.*, 1998, pp. 1486–1491.

[23] A. Isidori, *Nonlinear Control Systems*, 3rd ed. London, U.K.: Springer-Verlag, 1995.