International Journal of Automation and Computing 04(2), April 2007, 141-148
DOI: 10.1007/s11633-007-0141-1

Fault Diagnosis of Nonlinear Systems Using Structured Augmented State Models

Jochen Aβfalg[1*] Frank Allgöwer[2]

[1]Robert Bosch GmbH, Corporate Sector Research, 70442 Stuttgart, Germany

[2]University of Stuttgart, Institute for Systems Theory and Automatic Control, 70550 Stuttgart, Germany

Abstract: This paper presents an internal model approach for modeling and diagnostic functionality design for nonlinear systems operating subject to single- and multiple-faults. We therefore provide the framework of structured augmented state models. Fault characteristics are considered to be generated by dynamical exosystems that are switched via equality constraints to overcome the augmented state observability limiting the number of diagnosable faults. Based on the proposed model, the fault diagnosis problem is specified as an optimal hybrid augmented state estimation problem. Sub-optimal solutions are motivated and exemplified for the fault diagnosis of the well-known three-tank benchmark. As the considered class of fault diagnosis problems is large, the suggested approach is not only of theoretical interest but also of high practical relevance.

Keywords: Fault diagnosis, nonlinear systems, hybrid estimation.

1 Introduction

A common strategy to come up against model uncertainties and disturbances in system control and state estimation is to employ state augmentation. Exogenous input signals are considered to be generated by known dynamical exosystems which are added to the nominal plant dynamics[1, 2]. As fault modeling by exogenous signals is very general and able to describe nearly all types of faults[3], it seems to be obvious to apply exosystems also to fault diagnosis. Insights into the application of exosystems to linear fault diagnosis can be found in [4, 5].

At first glance, this approach proves to be advantageous as established methods for nonlinear state estimation can be utilized for nonlinear fault diagnosis. The internal model formulation guarantees the augmented model to hold for both the fault-free and the faulty operating system. Thus, malfunctions and dangerous plant operation can be monitored by means of physically motivated (fictive) characteristics. Nevertheless, this approach entails a severe drawback: As in real world fault diagnosis problems the considered faults affect the system rarely all at the same time, this *unstructured approach* is conservative and is restricted to problems where only few faults are possible (as a rough rule of thump, the number of diagnosable faults equals the number of measurement signals).

This contribution presents a powerful approach for the fault diagnosis of nonlinear systems operating subject to single- and multiple-faults. The motivation of the suggested approach is quite simple: Keep the advantages of the unstructured approach and overcome the limitations imposed on the number of diagnosable faults. We therefore suggest the framework of *structured augmented state models*: An assumable fault scenario (the most commonly encountered is certainly the single-fault scenario) describes the topology of a nondeterministic state automaton, whose discrete-states are related each to an operating mode which is due to the presence of a fault and to the fault-free case. The mode-specific continuous dynamics are selected from the unstructured model by switching equality constraints in the rule type of the automaton. By this way, the augmented unstructured state-space is constrained, conditioned on the operating mode. The resulting structured augmented state model describes a hybrid automaton. Based on this, the fault diagnosis problem is specified as an optimal hybrid augmented state estimation problem. From the angle of hybrid (augmented) state estimation, the fault diagnosis is then evaluated basically relying on [6].

The paper is organized as follows: In Section 2, a short description of the considered fault diagnosis problem is given and the framework of structured augmented state modeling is introduced. The fault diagnosis problem is specified from an optimal hybrid state estimation point of view and sub-optimal solutions are discussed in Section 3. An illustrative simulation example for the fault diagnosis of the three-tank benchmark is outlined in Section 4.

2 Problem description

We consider nonlinear dynamical systems which can be described by the following model

$$\Sigma\colon \begin{cases} \dot{x}(t) = f\big(x(t), u(t), \delta(t)\big),\ x(0) = x_0 \\ y(t) = h\big(x(t), u(t), \delta(t)\big) \end{cases}. \tag{1}$$

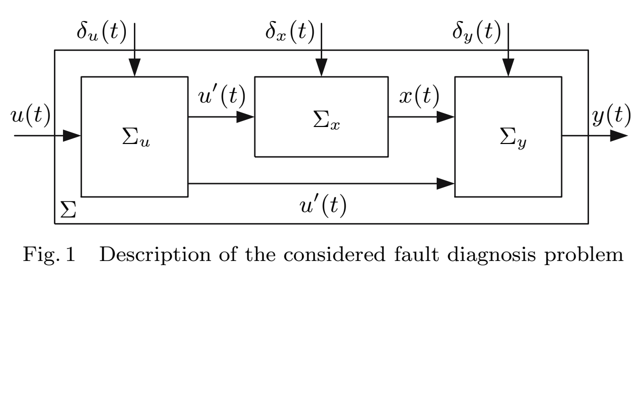

Fig. 1 Description of the considered fault diagnosis problem

Manuscript received January 10, 2006; revised December 11, 2006
*Corresponding author. E-mail: jochen.assfalg@de.bosch.com

The vector fields $f:\mathbb{R}^n\times\mathbb{R}^p\times\mathbb{R}^v\mapsto\mathbb{R}^n$ and $h:\mathbb{R}^n\times\mathbb{R}^p\times\mathbb{R}^v\mapsto\mathbb{R}^q$ are sufficiently smooth functions of the state $x(t)\in\mathbb{R}^n$, the control input $u(t)\in\mathbb{R}^p$, the measured output $y(t)\in\mathbb{R}^q$, and the exogenous input $\delta(t)\in\mathbb{R}^v$ representing deviations due to potential faults.

Depending on the position, the signals contained in $\delta(t)$ are going into (1), it may be distinguished between signals $\delta_u(t)$, $\delta_x(t)$, and $\delta_y(t)$ representing deviations due to potential input-faults, system-faults and output-faults, respectively. According to this, system (1) can be decomposed into the input-sided sub-system

$$\Sigma_u: u'(t)=g\big(u(t),\delta_u(t)\big) \tag{2}$$

being subject to the fault-signal $\delta_u(t)$, the central sub-system

$$\Sigma_x: \dot{x}(t)=f'\big(x(t),u'(t),\delta_x(t)\big) \tag{3}$$

which is subject to $\delta_x(t)$, and the output-sided sub-system

$$\Sigma_y: y(t)=h'\big(x(t),u'(t),\delta_y(t)\big) \tag{4}$$

underlying the influence of $\delta_y(t)$ (see Fig. 1). In (1), $f=f'\circ g$ and $h=h'\circ g$, where "$\circ$" denotes composition (for example $f'\circ g=f'(g)$). To simplify matters, further considerations use the shorter notation of (1) denoting $\delta(t)$ instead of $\delta_{u,x,y}(t)$.

2.1 Fault modeling and state vector augmentation

Many different principles for fault representation have been used in the literature. One of the most commonly encountered is to model faults by exogenous fault-signals. Because this principle has proved to be very general and to be suitable for describing all types of faults[3], this will be employed as basis for the subsequent fault representations.

In (1), deviations due to potential faults have been already introduced by the exogenous input-term $\delta(t)$. In the following, this is considered to be the vectorial concatenation

$$\delta(t)=\big[\delta_1^{\mathrm{T}}(t),\ldots,\delta_M^{\mathrm{T}}(t)\big]^{\mathrm{T}} \tag{5}$$

of M exogenous fault-signals $\delta_\rho(t)\in\mathbb{R}^{v_\rho}$, each characterizing a candidate $\mathcal{S}_\rho$ for a potential single-fault $\mathcal{S}$. With regard to the set defined by the fault-candidates $\mathcal{S}_\rho$, the following assumption will be made.

Assumption 1. The set

$$\mathbb{S}=\big\{\mathcal{S}_\rho,\ \rho\in\{1,\ldots,M\}\big\} \tag{6}$$

containing M candidates $\mathcal{S}_\rho$ for a potential single-fault $\mathcal{S}$ is known and complete.

Remark 1. In general, there is no possibility to prove whether a certain fault is present or not[5]. This necessitates the above assumption considering a single-fault $\mathcal{S}$ as an element of the given fault set $\mathbb{S}$. It is important to note that in practical applications the completeness of the set $\mathbb{S}$ can never be assured since the occurrence of an unforeseen fault can be excluded noway. Hence the risk of false diagnosis always exists, but may be decisively reduced by a judicious choice of the set $\mathbb{S}$. The practical design of $\mathbb{S}$ would typically not only include faults having been previously experienced but also potential fault-candidates that are possibly predictable from analyzing physical weak spots of the system.

2.1.1 Fault modeling by exosystems

Up to now, the fault-specific deviations $\delta_\rho(t)$ have been considered as unrestricted exogenous inputs. However, arbitrary fault-signals are rather implausible in the majority of practical cases[7], hence would model faults rather conservatively for most applications. In the following, each signal $\delta_\rho(t)\in\mathbb{R}^{v_\rho}$ is considered as the output generated by a known and stable dynamical exosystem

$$\Delta_\rho:\begin{cases}\dot{\xi}_\rho(t)=\phi_\rho\big(\xi_\rho(t)\big),\ \xi_\rho(0)=\xi_{\rho,0} & \text{(7a)}\\ \delta_\rho(t)=\psi_\rho\big(\xi_\rho(t),x(t),u(t)\big) & \text{(7b)}\end{cases}$$

where $\xi_\rho(t)\in\mathbb{R}^{n_\rho}$ is the (fault-) state of the exosystem and $\phi_\rho:\mathbb{R}^{n_\rho}\mapsto\mathbb{R}^{n_\rho}$, $\psi_\rho:\mathbb{R}^{n_\rho}\times\mathbb{R}^n\times\mathbb{R}^p\mapsto\mathbb{R}^{v_\rho}$ are sufficiently smooth vector fields.

The exosystem (7) operates as a signal-generator whose nonlinear function ϕ_ρ and initial condition $\xi_{\rho,0}$ uniquely determine the fault-state $\xi_\rho(t)$, thus, with ψ_ρ, $x(t)$, and $u(t)$, uniquely determines the fault-signal $\delta_\rho(t)$. The set $\mathbb{R}^{n_\rho}$ of admissible initial conditions and the exosystem Δ_ρ in conjunction specify a class of possible signals $\delta_\rho(t)$. In this way, the so far unrestricted fault-signals $\delta_\rho(t)$ get constrained to a smaller and more realistic set.

Remark 2. As discussed in [8,9], many practical system-faults are nonlinear functions of the system state $x(t)$ and/or the input $u(t)$. For example, the outflow due to a leak in a pressure vessel is a nonlinear function of the interior pressure and the temperature. Such fault characteristics are captured in (7) by allowing the deviations $\delta_\rho(t)$ to be nonlinear functions of $x(t)$ and $u(t)$. Other faults which can be modeled by (7) are, for instance, additive and multiplicative actuator-/sensor-faults, parametric faults, and system-faults causing the dynamics of (1) to change from f to another nonlinear function.

Since in general, the nonlinear output-function (7b) is physically motivated, this is known in the majority of practical cases. In contrast, the dynamics of (7a) is rarely known in practice. Therefore, one normally has to rely on simple exosystem models expressing the expected class of signals $\xi_\rho(t)$ sufficiently accurate. Such predominantly linear exosystems have received significant attention in the literature concerning disturbance modeling for system control and state estimation (see, for example [2,10]), but have been also applied to fault modeling for diagnostic applications[4,7]. A tabular summary of several linear exosystems for representing exogenous input-signals is given in [1].

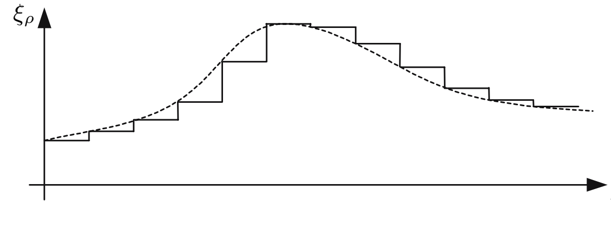

Fig. 2 Piece-wise constant approximation of an arbitrary signal

Remark 3. When time and frequency characteristics of an exogenous input are largely unknown, the assumption of piece-wise constant bias (see Fig. 2) has proved to be very effective for a variety of applications[11−13]. Due to the generally unpredictable nature of faults, such piece-wise constant approximations of the fault-states $\xi_\rho(t)$ are due to be indispensable in most diagnostic applications. In this case, the above introduced exosystem formulation reduces to

$$\Delta_\rho : \begin{cases} \dot{\xi}_\rho(t) = 0,\ \xi_\rho(0) = \xi_{\rho,0} \\ \delta_\rho(t) = \psi_\rho\big(\xi_\rho(t), x(t), u(t)\big) \end{cases} \tag{8}$$

which is a special case of (7). To keep generality, further considerations will, nevertheless, base on the general expression.

2.1.2 State vector augmentation

Given that system (1) is considered subject to the set of fault-candidates $\mathbb{S}$, the possible dynamics of (1) can be described by an augmented state model including the M exosystems. The extension of (1) by the M exosystems (7) results in an *unstructured augmented state model* which is given by

$$\tilde{\Sigma} : \begin{cases} \dot{\tilde{x}}(t) = \tilde{f}\big(\tilde{x}(t), u(t)\big),\ \tilde{x}(0) = \tilde{x}_0 \\ y(t) = \tilde{h}\big(\tilde{x}(t), u(t)\big) \end{cases} \tag{9}$$

where $\tilde{x}(t) = [x^{\mathrm{T}}(t)\ \xi_1^{\mathrm{T}}(t), \dots, \xi_M^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathbb{R}^{\tilde{n}}$ is the augmented state, $\tilde{f} = [(f \circ (\psi_1, \dots, \psi_M))^{\mathrm{T}}\ \phi_1^{\mathrm{T}}, \dots, \phi_M^{\mathrm{T}}]^{\mathrm{T}}$ is the augmented dynamics, and $\tilde{h} = h \circ (\psi_1, \dots, \psi_M)$ denotes the modified output term.

From a methodical point of view, (9) is particularly useful. At first glance, crucial diagnostic problems like nonlinear system diagnosis and diagnosability analysis can be efficiently tackled by using established methods of nonlinear state estimation theory. For instance, the fault diagnosis of (1) is accomplishable by means of observing the augmented state $\tilde{x}(t)$[4]. Accordingly, the diagnosability of (1) can be analyzed by verifying observability of the augmented system (9)[7]. That this approach, anyhow, hasn't become widely accepted in fault diagnosis so far may be traced back to the drastic limitation on the number of considerable faults. By the augmented state observability of (9), this unstructured approach is restricted to problems where only few faults may occur. For an extensive study of this issue, the interested reader is directed to [14]. The goal of the following subsection is to provide a less conservative model allowing to overcome these limitations.

2.2 Structured augmented state modeling for fault diagnosis

If in the fault-free condition of (1) each signal $\delta_\rho(t)$ takes a constant value $\delta_\rho^\star \in \mathbb{R}^{v_\rho}$, i.e., $\xi_\rho(t) = 0$, a (single- or multiple-) fault $\mathcal{F}$ can be closer specified as in the following definition.

Definition 1. A single- or multiple-fault $\mathcal{F}$ is defined as a deviation

$$\delta(t) \neq \delta^\star,\ \delta^\star = [\delta_1^{\star\mathrm{T}}, \dots, \delta_M^{\star\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^v \tag{10}$$

changing the behavior of system (1) such that this no longer satisfies its purpose appropriately.

Remark 4. Depending on the number of signals $\delta_\rho(t) \neq \delta_\rho^\star$ causing a fault by (10), $\mathcal{F}$ may be of both single- and multiple-fault type. More precisely, if $\delta_\rho(t) \neq \delta_\rho^\star$ is contained in $\delta(t)$ just once, the fault $\mathcal{F}$ is a single-fault. On the other hand, $\mathcal{F}$ is a multiple-fault if there are at least two signals $\delta_\rho(t) \neq \delta_\rho^\star$ contained in $\delta(t)$.

As an important implication of Definition 1, the case of system (1) operating subject to several faults $\mathcal{F}$ is ruled out. Therefore, a potential fault $\mathcal{F}$ can be treated again as an element of a set of $N \geq M$ single- or multiple-fault candidates $\mathcal{F}_\nu$. As regards the set defined by $\mathcal{F}_\nu$, $\nu \in \{1, \dots, N\}$, the following assumption, similar to Assumption 1, will be used.

Assumption 2. The set

$$\mathbb{F} = \big\{\mathcal{F}_\nu,\ \nu \in \{1, \dots, N\}\big\},\ N \geq M \tag{11}$$

containing N candidates $\mathcal{F}_\nu$ for a potential single- or multiple-fault $\mathcal{F}$ is known and complete.

Remark 5. The above assumption is similar to Assumption 1, hence it is worth to allude to the explanatory notes of Remark 1.

In the following, to any fault $\mathcal{F}_\nu$ and to the fault-free case, an operating mode of system (1) is related. The resulting $N+1$ operating modes are closer specified by the following definition.

Definition 2. The system Σ is defined to operate in the fault-free operating mode $z(t) = m_0$, if, *for all* $\nu \in \{1, \dots, N\}$,

$$\Sigma\big(\delta_\nu(t) = \delta_\nu^\star\big) \equiv \big\{\tilde{\Sigma}_0 := \tilde{\Sigma}\big(\xi_\nu(t) = 0\big)\big\}. \tag{12}$$

The system Σ is defined to operate in the ν-th single- or multiple-faulty operating mode $z(t) = m_\nu$, if

$$\Sigma\big(\delta_\nu(t) \neq \delta_\nu^\star\big) \equiv \big\{\tilde{\Sigma}_\nu := \tilde{\Sigma}\big(\xi_\nu(t) \neq 0\big)\big\} \tag{13}$$

where only one operating mode is allowed to be present at the same time. Thereby, $\delta_\nu^\star$ denotes the value of $\delta_\nu(t)$ in the fault-free condition and $z(t) \in \mathbb{Z} = \{m_\sigma\}_{\sigma=0}^N$ is a discrete-state variable taking values of the set of all possible operating modes $\mathbb{Z}$.

According to Definition 2, any operating mode m_σ is defined by a corresponding continuous augmented dynamics $\tilde{\Sigma}_\sigma$ contained in (9) as a special case. Each of these dynamics can be selected from (9) by switching a mode-specific equality constraint

$$E\big(z(t) = m_\sigma\big)\tilde{x}(t) = 0 \Rightarrow \tilde{x}_\sigma(t) \in \tilde{\mathbb{X}}_\sigma \subseteq \mathbb{R}^{\tilde{n}} \tag{14}$$

constraining mismatched fault-states $\xi_\rho(t)$ contained in $\tilde{x}(t)$ to zero. Thereby, $E(z(t)) \in \mathbb{E} = \{S_0, \dots, S_N\}$, where $S_\sigma \in$

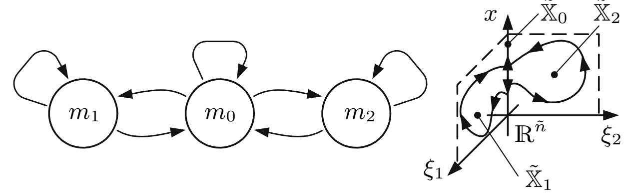

Fig. 3 Nondeterministic automaton and related structured augmented state space of an exemplary single-fault scenario

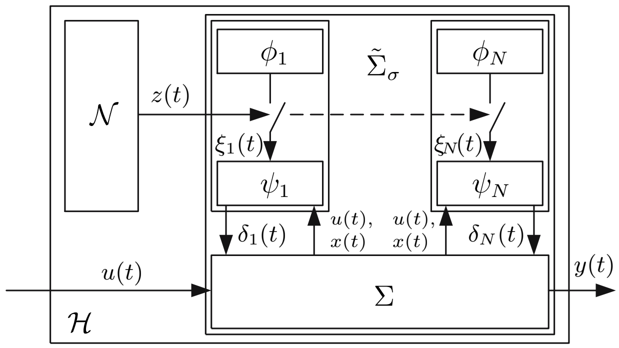

Fig. 4 Structured augmented state model

$\mathbb{R}^{\tilde{n}\times\tilde{n}}$ is a mode-specific diagonal matrix with elements $s_\sigma \in \{0,1\}$. By this means, the augmented state space $\mathbb{R}^{\tilde{n}}$ gets structured into subspaces $\tilde{\mathbb{X}}_\sigma$ restricting possible augmented state trajectory evolutions mode-specifically (see Fig. 3, right figure). Hence, the structured continuous dynamics of the potential faulty system (1) is given by

$$\tilde{\Sigma}_\sigma: \begin{cases} \dot{\tilde{x}}(t) = \tilde{f}\big(\tilde{x}(t), u(t)\big),\ \tilde{x}(0) = \tilde{x}_0 & \text{(15a)} \\ E\big(z(t)= m_\sigma\big)\tilde{x}(t) = 0 \Rightarrow \tilde{x}_\sigma(t) \in \tilde{\mathbb{X}}_\sigma \subseteq \mathbb{R}^{\tilde{n}} & \text{(15b)} \\ y(t) = \tilde{h}\big(\tilde{x}_\sigma(t), u(t)\big). & \text{(15c)} \end{cases}$$

An important role in structuring the $\mathbb{R}^{\tilde{n}}$ is played by the rule type of switching $z(t)$. We therefore suggest to employ a nondeterministic automaton directly expressing an assumable fault scenario (the most commonly encountered is certainly a single-fault scenario (see Fig. 3, left figure). The nondeterministic automaton is described by the tuple

$$\mathcal{N}\big\langle \mathbb{Z}, G: \mathbb{Z} \mapsto \mathbb{Z}, z_0\big\rangle \tag{16}$$

where G is the discrete-state transition matrix with elements $g_{ab} \in \{0,1\}$ and $z_0 = z(0)$ is the initial discrete-state.

Based on (15) and the nondeterministic automaton (16), the structured augmented state model (see Fig. 4) can be described as follows.

Definition 3. A structured augmented state model is described by a hybrid automaton

$$\mathcal{H}\big\langle\, \chi(t), u(t), y(t), \tilde{\Sigma}_\sigma, \mathbb{E}, \mathcal{N}, \chi_0\big\rangle \tag{17}$$

which is a tuple of the hybrid state $\chi(t) = \langle z(t), \tilde{x}(t)\rangle$ with initial value $\chi_0 = \langle z(0), \tilde{x}(0)\rangle$, the set $\mathbb{E}$ of $N+1$ switching matrices S_σ, and the tuple $\langle u(t), y(t), \tilde{\Sigma}_\sigma, \mathcal{N}\rangle$.

Remark 6. The structure of the automaton (16) can be easily adapted to any arbitrary multiple-fault scenario, i.e., where a fault may occur, remain and may cause a consecutive fault.

3 Fault diagnosis

In this section, the fault diagnosis of the structured augmented system is considered. As in most practical applications the vector of measurements is taken with a sampling rate dT at discrete-time, further considerations are based on the following continuous-discrete form of (17), given by

$$\mathcal{H}_{dT}\big\langle\, \chi(t), u(t), y_k, \tilde{\Sigma}_\sigma^{dT}, \mathbb{E}, \mathcal{N}, \chi_0\big\rangle \tag{18}$$

where

$$\tilde{\Sigma}_\sigma^{dT}: \begin{cases} \dot{\tilde{x}}(t) = \tilde{f}\big(\tilde{x}(t), u(t)\big),\ \tilde{x}(0) = \tilde{x}_0 \\ E\big(z(t)= m_\sigma\big)\tilde{x}(t) = 0 \\ y_k = \tilde{h}\big(\tilde{x}_{\sigma,k}, u_k\big) \end{cases} \tag{19}$$

and $y_k := y(k\,dT)$, $\tilde{x}_{\sigma,k} := \tilde{x}_\sigma(k\,dT)$, and $u_k := u(k\,dT)$. For the sake of well-posedness of (18), the following assumption will be made.

Assumption 3. The sampling rate dT is chosen adequately; such that (18) remains diagnosable and the contained nondeterministic automaton $\mathcal{N}$ is able to switch maximum once within the intervals $[k\,dT,\ (k+1)\,dT]$.

Based on (18), the fault diagnosis problem consists in estimating the hybrid state χ_k containing the complete information for fault-isolation and -identification (determining the fault-magnitude) in z_k, $\tilde{x}_k$ respectively. As (18) allows switching among the operating modes m_σ, the hybrid estimator has to track possible trajectories involving mode changes $\mathcal{X}_k = \{\chi_0, \ldots, \chi_k\}$ and determine their estimates $\hat{\mathcal{X}}_k = \{\hat{\chi}_0, \ldots, \hat{\chi}_k\}$. Against this background, the fault diagnosis problem can be stated as follows.

Problem 1. Given the continuous-discrete structured augmented state model (18), the continuous-time input trajectory $U_k = u(t_0, t_k)$, and the discrete-time measurement sequence $Y_k = \{y_1, \ldots, y_k\}$, determine the optimal hybrid state trajectory

$$\hat{\mathcal{X}}_k^{(\zeta)} : \zeta = \underset{j=\{1,\ldots,\lambda_k\}}{\arg\min}\ \Psi_k^{(j)} \tag{20}$$

where $\Psi_k^{(j)} = \mathcal{C}\big(\hat{\mathcal{X}}_k^{(j)}; Y_k, U_k\big)$ defines the path cost of the j-th trajectory hypothesis $\hat{\mathcal{X}}_k^{(j)}$ quantifying the level of agreement between $\hat{\mathcal{X}}_k^{(j)}$ and Y_k, U_k.

Remark 7. Since it is dealt with a set of trajectory hypotheses, a superscript index in parentheses is used, for example, $\hat{\mathcal{X}}_k^{(j)}$, to refer to the j-th trajectory hypothesis $\hat{\mathcal{X}}_k^{(j)} = \{\ldots, \hat{\chi}_{k-1}^{(i)}, \hat{\chi}_k^{(j)}\}$ that estimates a possible trajectory up to time k. This trajectory hypothesis includes the hybrid estimate $\hat{\chi}_k^{(j)}$ at its fringe.

The optimal solution of the above stated fault diagnosis problem requires to build a full-hypothesis tree (see Fig. 5) which encodes the estimates of all possible trajectories of (18). Therefore, any hypothesis $\hat{\mathcal{X}}_{k-1}^{(i)}$ is extended with respect to the topology (given by the transition matrix G, with elements g_{ab}) of the automaton $\mathcal{N}$ by possible successor states $\hat{\chi}_k^{(j)}$ and new trajectory hypotheses $\hat{\mathcal{X}}_k^{(j)} = \{\ldots, \hat{\chi}_{k-1}^{(i)}, \hat{\chi}_k^{(j)}\}$ are obtained. The resulting λ_k hypotheses are ranked according to their path costs $\Psi_k^{(j)}$. The fringe estimates $\hat{\chi}_k^{(j)}$ and their associated costs $\Psi_k^{(j)}$ are calculated according to the following two step process (see Fig. 6):

1) The first step deduces possible transitions $\hat{z}_{k-1}^{(i)} = m_a \rightarrow \hat{z}_k^{(j)} = m_b$, $a, b \in \{0, \ldots, N\}$, and predicts λ_k tra-

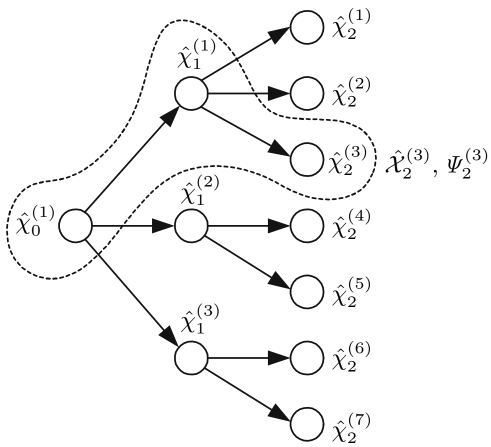

Fig. 5 Exemplary full-hypothesis tree and trajectory path for behavior $\hat{\mathcal{X}}_2^{(3)}$ with cost $\Psi_2^{(3)}$ for the automaton, exemplified in Fig. 3 ($\hat{z}_0^{(1)} = m_0$)

jectory hypotheses

$$\hat{\mathcal{X}}_{k|k-1}^{(j)} = \Big\{ \ldots, \hat{\chi}_{k-1}^{(i)} = \langle \hat{z}_{k-1}^{(i)} = m_a, \hat{\tilde{x}}_{a,k-1}^{(i)} \rangle, \\ \hat{\chi}_{k|k-1}^{(j)} = \langle \hat{z}_k^{(j)} = m_b, \hat{\tilde{x}}_{b,k|k-1}^{(j)} \rangle \Big\}. \tag{21}$$

Thereby, the j-th prior fringe estimate $\hat{\chi}_{k|k-1}^{(j)}$ is a tuple of the mode estimate $\hat{z}_k^{(j)} = m_b$ and the continuous augmented state $\hat{\tilde{x}}_{b,k|k-1}^{(j)}$ of the prediction $\hat{\tilde{x}}_{a,k-1}^{(i)} \to \hat{\tilde{x}}_{b,k|k-1}^{(j)}$. This is obtained by integrating the augmented dynamics $\tilde{f}$ subject to the equality constraint (14) as

$$\begin{aligned} &E(\hat{z}_k^{(j)} = m_b)\, \hat{\tilde{x}}_{a,k-1}^{(i)} = 0 \\ &\hat{\tilde{x}}_{b,k|k-1}^{(j)} = \hat{\tilde{x}}_{a,k-1}^{(i)} + \int_{t_{k-1}}^{t_k} \tilde{f}\big(\tilde{x}(t), u(t)\big)\, \mathrm{d}t \end{aligned}. \tag{22}$$

2) The second step takes the current measurement y_k into account. It performs the correction of the predicted continuous state estimate $\hat{\tilde{x}}_{b,k|k-1}^{(j)} \to \hat{\tilde{x}}_{b,k}^{(j)}$. The cost of the corrected trajectory hypothesis $\hat{\mathcal{X}}_k^{(j)} = \big\{ \ldots, \hat{\chi}_k^{(j)} = \langle \hat{z}_k^{(j)} = m_b, \hat{\tilde{x}}_{b,k}^{(j)} \rangle \big\}$ is calculated recursively by

$$\Psi_k^{(j)} := \Psi_{k|k-1}^{(j)} + \Delta\mathcal{C}_k^{(j)}, \tag{23}$$

where $\Psi_{k|k-1}^{(j)} = \Psi_{k-1}^{(i)} + \mathcal{P}_k^{(j)}$ is the predicted cost including the deterministic weight $\mathcal{P}_k^{(j)}$ serving for penalizing the corresponding mode transition $m_a \to m_b$, and

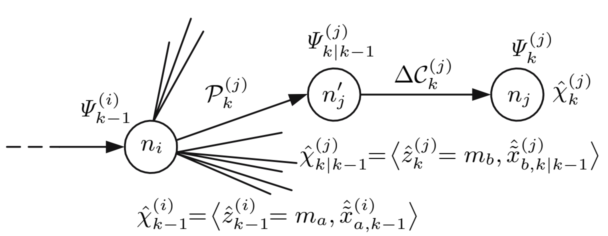

Fig. 6 Two-step recursion in optimal hybrid estimation

$\Delta\mathcal{C}_k^{(j)} = \mathcal{C}\big(y_k; \hat{\chi}_{k|k-1}^{(j)}, u_k\big)$ is a cost term which quantifies the level of agreement between the prior hybrid state estimate $\hat{\chi}_{k|k-1}^{(j)}$ and u_k, y_k.

Remark 8. As the number of involved trajectory hypotheses λ_k is exponential in the number of considered time steps k, an optimal solution of the fault diagnosis problem is computationally intractable. Therefore, the use of sub-optimal estimation approaches is indispensable for practical applications. In literature, this problem has been addressed by a variety of authors, e.g. [6,15,16].

4 Simulation results

Let us consider the three-tank system illustrated in Fig. 7 (for further discussions regarding this well-known fault diagnosis benchmark, the interested reader is directed to [9,17-20]). The three tanks, T_1, T_2, and T_3 are identical and are cylindrical in shape with a cross section $A = 4\cdot 10^{-2}$ m^2. The cross section of the pipes connecting the tanks is $S = 1.6 \cdot 10^{-4}$ m^2, and the liquid levels in the three tanks are denoted by x_1, x_2, and x_3 (in (m)) respectively. The incoming flow rates supplied by the pumps 1 and 2 are denoted by u_1 and u_2 (in (m^3/s)) respectively.

By using balance equations and Torricelli's rule, the nonlinear dynamics of the three-tank is obtained as

$$\Sigma^{dT}: \begin{cases} u_1' = u_1 + \delta_1 \\ u_2' = u_2 + \delta_7 \\ \dot{x}_1 = \dfrac{1}{A}\big(u_1' - q_1(x_1, x_2, \delta_3)\big) \\ \dot{x}_2 = \dfrac{1}{A}\big(q_1(x_1, x_2, \delta_3) - q_2(x_2, x_3, \delta_6) - \delta_5\big) \\ \dot{x}_3 = \dfrac{1}{A}\big(u_2' + q_2(x_2, x_3, \delta_6) - q_3(x_3)\big) \\ y_1 = x_1 + \delta_2 \\ y_2 = x_2 + \delta_4 \\ y_3 = x_3 + \delta_8 \end{cases}$$

where $x = [x_1\ x_2\ x_3]^{\mathrm{T}}$ denotes the state vector, $u = [u_1\ u_2]^{\mathrm{T}}$ denotes the input vector, and $y = [y_1\ y_2\ y_3]^{\mathrm{T}}$ is the vector of measurements, taken with a sampling rate $dT = 1$ seconds at discrete-time (to simplify matters, it is refrained from explicitly denoting the time-dependencies of continuous-/discrete-time-variant variables). The flow-rates q_1 and q_2 between the tanks and the outflow-rate q_3 are given by

$$\begin{aligned} q_1 &= c \operatorname{sign}(x_1 - x_2)(S - \delta_3)\sqrt{2g|x_1 - x_2|} \\ q_2 &= c \operatorname{sign}(x_2 - x_3)(S - \delta_6)\sqrt{2g|x_2 - x_3|}. \\ q_3 &= c\, S\sqrt{2gx_3}. \end{aligned}$$

Thereby, $c = 0.8$ denotes a nondimensional outflow coefficient, and g is the gravity acceleration.

In the following, the case of abrupt faults satisfying a classical single-fault scenario is considered (the case of other faults, like incipient, periodic or multiple faults, is completely analogous and is not addressed here for the sake of brevity). The set $\mathbb{F}$ is defined by the following eight fault candidates:

1) Input-faults $\mathcal{F}_{1,7}$ (actuator faults in pumps 1 and 2). Two simple additive actuator faults in pumps 1 and 2 are

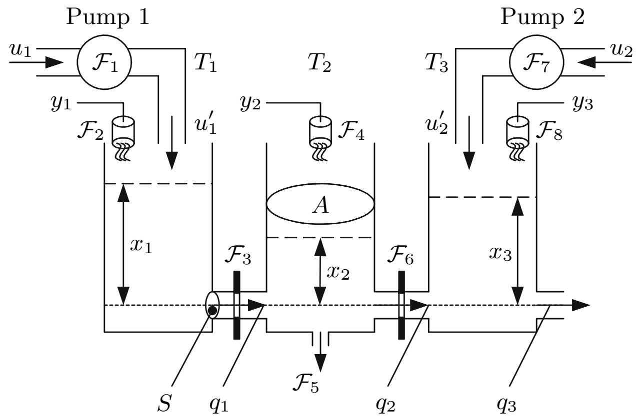

Fig. 7 The three-tank system

considered. Due to the input-faults $\mathcal{F}_i$, $i = 1, 7$, the incoming supply flow rates $u'_i = u_i + \delta_i$, deviate by δ_i from the nominal input values u_i.

2) Output-faults $\mathcal{F}_{2,4,8}$ (defective sensors in tanks 1,2, and 3). Three simple additive sensor faults in tanks 1,2 and 3 are considered. Due to the sensor-faults $\mathcal{F}_i$, $i = 2, 4, 8$, the measured values of the liquid levels $y_i = x_i + \delta_i$ deviate by δ_i from the real levels x_i.

3) System-faults $\mathcal{F}_{3,6}$ (plugging between tanks 1 and 2, and between tanks 2 and 3). Possible plugging of the connection pipes between the three tanks is modeled by the deviations δ_i, $i = 3, 6$, reducing the corresponding cross sections to $(S - \delta_i)$.

4) System-fault $\mathcal{F}_5$ (leakage in tank 2). It is assumed that the leak is circular in shape and of unknown effective area ξ_5. If δ_5 denotes the outflow rate which is due to the leak, this is given by the nonlinear function $\delta_5 = \xi_5\sqrt{2gx_2}$.

According to the foregoing considerations of Section 2, the above introduced faults are now represented by the following dynamical exosystems:

$$\Delta_i : \begin{cases} \dot{\xi}_i = 0 \\ \delta_i = \xi_i \end{cases}, \; i = 1-4, 6-8, \;\; \Delta_5 : \begin{cases} \dot{\xi}_5 = 0 \\ \delta_5 = \xi_5\sqrt{2gx_2} \end{cases}.$$

The unstructured augmented expression $\tilde{\Sigma}^{dT}$, with extended state $\tilde{x} = [x_1 \; x_2 \; x_3 \; \xi_1, \ldots, \xi_8]^{\mathrm{T}}$, can be easily derived by augmenting system Σ^{dT} with the eight exosystems Δ_{1-8}. Adding the switched equality constraint

$$E(z = m_\sigma)\,\tilde{x} = 0 \Rightarrow \tilde{x}_\sigma, \;\; E(z) \in \mathbb{E}$$

where $\mathbb{E} = \{S_0 = \mathrm{diag}(0^{1\times3} \;\; 1^{1\times8}), \ldots, S_8 = \mathrm{diag}(0^{1\times3} \;\; 1^{1\times7} \;\; 0)\}$, yields the structured augmented expression $\tilde{\Sigma}_\sigma^{dT}$. As regards the assumed single-fault scenario, the corresponding nondeterministic automaton $\mathcal{N}$ contains eight fault-modes m_{1-8} that are radiating-adjusted around the fault-free operating mode m_0 (a similar fault-scenario considering two fault-modes is outlined in Fig. 3). The initial hybrid state of the structured augmented system $\mathcal{H}_{dT}$ is chosen as $\chi_0 = \langle z_0, \tilde{x}_0 \rangle$, where $z_0 = m_0$ and $\tilde{x}_0 = [0.36 \; 0.32 \; 0.20 \; 0^{1\times8}]^{\mathrm{T}}$. The flow rates supplied by the pumps 1 and 2 are chosen as $u_1 = 10^{-4}(\mathrm{m}^3/\mathrm{s})$ and $u_2 = 7 \cdot 10^{-5}(\mathrm{m}^3/\mathrm{s})$.

For computing the fault diagnosis of $\mathcal{H}_{dT}$, i.e., for determining the hybrid state estimates $\hat{\chi}_k^{(\zeta)}$, a sub-optimal K-step focused hybrid estimation scheme with $\kappa, \eta = 1$ leading estimates and a fixed window size of $K = 2$ is employed. The estimator is initialized by $\hat{\chi}_0 = \chi_0$. The cost $\Delta\mathcal{C}_k^{(j)}$ of the continuous correction-step is calculated by

$$\Delta\mathcal{C}_k^{(j)} = \left|\left|\epsilon_k^{(j)}\right|\right|^2_{\Lambda_k^{(j)-1}}$$

where $\epsilon_k^{(j)} = y_k - \tilde{h}(\hat{\tilde{x}}_{k|k-1}^{(j)}, u_k, \hat{z}_k^{(j)})$ denotes the measurement residual with associated covariance matrix $\Lambda_k^{(j)}$, both computed by an unscented Kalman observer. An exhaustive discussion of the sub-optimal K-step focused hybrid estimation scheme as well as the employed unscented Kalman observer is clearly beyond the scope of this paper. For detailed discussions, the interested reader is referred to [6,21-23], respectively.

Fig. 8 outlines the excellent results, achieved with this sub-optimal fault diagnosis scheme in the application to the three-tank system. This has been imposed by the eight single-faults $\mathcal{F}_{1-8}$ occurring under the terms of the test-sequence as outlined in Table 1. Beyond the limitation imposed by the unstructured augmented state observability, the suggested approach is able to detect, isolate, and identify the faults $\mathcal{F}_{1-8}$. The operating modes m_{0-8} are estimated without significant time-delays (see Fig. 8 (a)). The estimates of the fault-magnitudes ξ_{1-8} determine an accurate fault-identification (see Fig. 8 (b)). The internal model formulation gives also estimates of the real values $y'_k = [x_{1,k} \; x_{2,k}]^{\mathrm{T}}$ despite of the output-faults $\mathcal{F}_{2,4,8}$ - which may be very useful, e.g., for fault-tolerant control. Moreover, $\hat{\xi}_{1-8}$ can be used for determining physically motivated thresholds for monitoring critical or dangerous plant operation. As can be noticed from Fig. 8 (c), the number of performed continuous observation-steps is variant in time and significantly underruns the constant number of nine observation-steps which would have been required to calculate using a classical observer-bank approach. Hence, in off-line applications, a given sequence of measurements U_k and Y_k can be evaluated much faster than using a classical observer-bank approach.

Table 1 Overview of the performed test-scenario (thereby, t_{on} and t_{off} denote the time of fault-activation and -deactivation, respectively)

Fault/Mode	Fault magnitude	t_{on} (s)	t_{off} (s)
$\mathcal{F}_1/m_1$	$\xi_1 = 10^{-5}\left(\frac{\mathrm{m}^3}{s}\right)$	150	300
$\mathcal{F}_2/m_2$	$\xi_2 = 0.04$ (m)	400	550
$\mathcal{F}_3/m_3$	$\xi_3 = 4 \cdot 10^{-5}\left(\mathrm{m}^2\right)$	650	800
$\mathcal{F}_4/m_4$	$\xi_4 = 0.04$ (m)	900	1 050
$\mathcal{F}_5/m_5$	$\xi_5 = 3 \cdot 10^{-5}\left(\mathrm{m}^2\right)$	1 150	1 300
$\mathcal{F}_6/m_6$	$\xi_6 = 3 \cdot 10^{-5}\left(\mathrm{m}^2\right)$	1 400	1 550
$\mathcal{F}_7/m_7$	$\xi_7 = 10^{-5}\left(\frac{\mathrm{m}^3}{s}\right)$	1 650	1 800
$\mathcal{F}_8/m_8$	$\xi_8 = 0.03$ (m)	1 900	2 050

5 Conclusions

In this paper, the fault diagnosis of nonlinear systems using structured augmented state models has been consid-

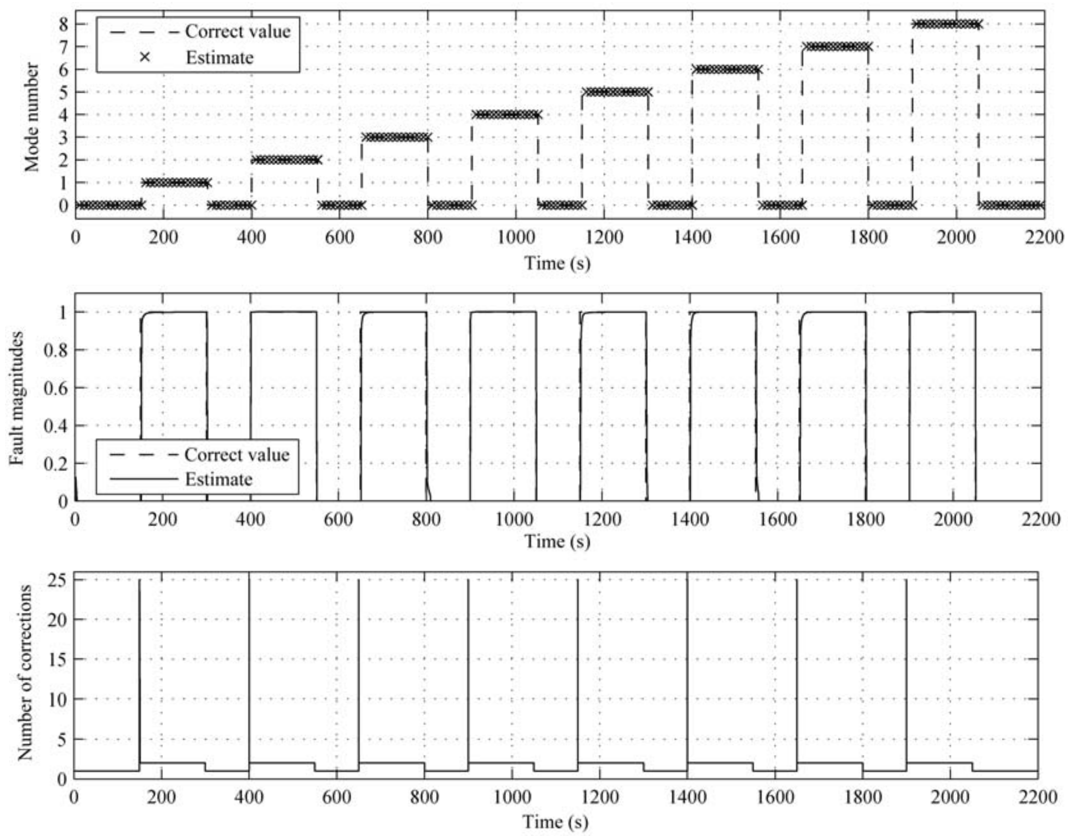

Fig. 8 Fault diagnosis of the three-tank system. (a) Correct and estimated mode sequence (fault isolation); (b) Correct and estimated (normalized) fault magnitudes (fault identification); (c) Number of performed continuous observations

ered. To negotiate the limitation on the number of diagnosable faults (which is imposed by the unstructured state observability), an assumable single- or multiple-fault scenario serves for realistically delimiting the state space of the unstructured model. Based on the structured augmented state model, the fault diagnosis problem has been characterized and extensively discussed from an optimal hybrid state estimation point of view. The use of modern sub-optimal focused hybrid estimation techniques has been motivated and exemplified for the fault diagnosis of the three-tank benchmark. These not only enable the diagnostic functionality to switch back to the fault-free mode in case that a fault is recovered (which is not possible if a classical bank of observers is employed) but also allow to overcome computational burdens of the observer-bank approach and improve the achievable diagnosis performance. Advantageous properties of the unstructured internal model formulation have been maintained and the problem of conservativeness has been vanquished. As the suggested approach is not only simple to apply but also able to detect, isolate, and identify single- and multiple-faults for SISO- and MIMO-type nonlinear systems, the considerable class of fault diagnosis problems is large.

References

[1] A. Gelb, J. F. Kasper, R. A. Nash, C. F. Price, A. A. Sutherland. *Applied Optimal Estimation*, The MIT Press, Cambridge, Massachussets, London, England, 1974.

[2] A. Isidori, L. Marconi, A. Serrani. *Robust Autonomous Guidance: An Internal Model Approach*, Springer Verlag, London, 2004.

[3] M. Nyberg. Model Based Fault Diagnosis Using Structured Hypothesis Tests. In *Proceedings of the IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, Budapest, Hungary, pp. 1026-1031, 2000.

[4] T. S. Liu, J. J. Liou. An Augmented System Model for Fault Detection and Identification. In *Proceedings of the American Control Conference*, Washington, USA, pp. 602-606, 1995.

[5] M. Blanke, M. Kinnaert, J. Lunze, M. Staroswiecki. *Diagnosis and Fault-Tolerant Control*, Springer Verlag, Berlin, Heidelberg, 2003.

[6] M. W. Hofbaur, B. C. Williams. Hybrid Estimation of Complex Systems. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 34, no. 6, pp. 2178-2191, 2004.

[7] J. Lunze. *Automatisierungstechnik: Methoden Für Die Überwachung und Steuerung Kontinuierlicher und Ereignisdiskreter Systeme*, Oldenbourg, München, Wien, 2003.

[8] J. J. Gertler. Survey of Model-based Failure Detection and Isolation in Complex Plants. *IEEE Control Systems Magazine*, vol. 8, no. 6, pp. 3-11, 1998.

[9] X. Zhang, M. M. Polycarpou, T. Parisini. A Robust Detection and Isolation Scheme for Abrupt and Incipient Faults in Nonlinear Systems. *IEEE Transactions on Automatic Control*, vol. 47, no. 4, pp. 576-593, 2002.

[10] H. Cox. On the Estimation of State Variables and Parameters for Noisy Dynamic Systems. *IEEE Transactions on Automatic Control*, vol. 9, no. 1, pp. 5-12, 1964.

[11] G. S. Buja, R. Menis, M. I. Valla. Disturbance Torque Estimation in a Sensorless Dc Drive. *IEEE Transactions on Industrial Electronics*, vol. 42, no. 4, pp. 351-357, 1995.

[12] H. Aschemann, O. Sawodny, S. Lahres, E. P. Hofer. Disturbance Estimation and Compensation for Trajectory Control of an Overhead Crane. In *Proceedings of the IEEE American Control Conference*, Chicago, USA, pp. 1027-1031, 2000.

[13] J. Aβfalg, F. Allgöwer, M. Fritz. Constrained Derivative-free Augmented State Estimation for a Diesel Engine Air Path. In *Proceedings of the IFAC Symposium on System Identification*, New Castle, Australia, pp. 1382-1387, 2006.

[14] J. Aβfalg, F. Allgöwer. Fault Diagnosis with Structured Augmented State Models: Modeling, Analysis, and Design. In *Proceedings of the IEEE Conference on Decision and Control*, San Diego, USA, pp. 1165-1170, 2006.

[15] G. Ackerson, K. Fu. On State Estimation in Switching Environments. *IEEE Transactions on Automatic Control*, vol. 15, no. 1, pp. 10-16, 1970.

[16] H. A. Blom, Y. Bar-Shalom. The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients. *IEEE Transactions on Automatic Control*, vol. 33, no. 8, pp. 780-783, 1988.

[17] R. Isermann. Supervision, Fault Detection, and Fault Diagnosis Methods – An Introduction. *Control Engineering Practice*, vol. 5, no. 5, pp. 639-652, 1997.

[18] B. Köppen-Seliger, E. A. Garcia, P.M. Frank. Fault Detection: Different Strategies for Modelling Applied to the Three Tank Benchmark - A Case Study. In *Proceedings of the IEEE European Control Conference*, Karlsruhe, Germany, 1999.

[19] C. Join, J. C. Ponsart, D. Sauter, D. Theilliol. Nonlinear Filter Design for Fault Diagnosis: Application to the Three-tank System. *IEE Proceedings of Control Theory Applications*, vol. 152, no. 1, pp. 55-64, 2005.

[20] P. D. Hanlon, P. S. Maybeck. Interrelationship of Single-filter and Multiple-model Adaptive Algorithms. *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 3, pp. 934-946, 1998.

[21] M. W. Hofbaur. *Hybrid Estimation of Complex Systems*, Springer Verlag, Berlin, Heidelberg, 2005.

[22] S. J. Julier, J. K. Uhlmann, H. F. Durrant-Whyte. A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators. *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477-482, 2000.

[23] R. Merwe, E. Wan. The Square-root Unscented Kalman Filter for State and Parameter-estimation. *In Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, USA, pp. 3461-3464, 2001.

Jochen Aβfalg was born in Biberach a.d. Riβ, Germany in 1978. He received his Diploma in mechanical engineering from the University of Stuttgart, Germany, in 2004. He is currently an external Ph.D. student of the Institute for Systems Theory and Automatic Control at the University of Stuttgart, Germany, working as a research-engineer at the Department of Mechatronic Systems Engineering at the Robert Bosch GmbH, Germany.

His research interests include the development of new methods for the fault diagnosis and fault-tolerant control of nonlinear systems. His applied research interests include mechatronic and automotive systems.

Frank Allgöwer studied engineering cybernetics and applied mathematics at the University of Stuttgart and the University of California at Los Angeles respectively. He received his Ph.D. degree in chemical engineering from the University of Stuttgart. Prior to his present appointment he held a professorship in the Electrical Engineering Department at ETH Zurich. He also held visiting positions at Caltech, the NASA Ames Research Center, the DuPont Company and the University of California at Santa Barbara. He is the director of the Institute for Systems Theory and Automatic Control and a professor in the Mechanical Engineering Department at the University of Stuttgart in Germany.

His research interests include the development of new methods for the analysis and control of nonlinear systems and the identification of nonlinear systems. His applied research interests range from chemical process control and control of mechatronic systems to AFM control and systems biology.

Prof. Allgöwer is editor for the journal *Automatica*, associate editor of the *Journal of Process Control* and the *European Journal of Control* and is on the editorial board of several other journals. He currently serves on the scientific council of the German Society for Measurement and Control, is on the Board of Governors of the IEEE Control System Society, is chairman of the IFAC Technical Committee on Nonlinear Systems and member of the IFAC Policy Committee. He has been organizer or co-organizer of several international conferences and has published over 150 scientific articles. He received several prizes for his work including the Leibniz prize, which is the most prestigious prize in science and engineering awarded by the German National Science Foundation (DFG).