Fault detection and isolation using viability theory and interval observers

Majid Ghaniee Zarch[a], Vicenç Puig[b], Javad Poshtan[a] and Mahdi Aliyari Shoorehdeli[c]

[a]Electrical Engineering Department, Iran University of Science and Technology (IUST), Tehran, Iran; [b]Advanced Control Systems (SAC), Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Barcelona, Spain; [c]Department of Mechatronic, K.N. Toosi University of Technology, Tehran, Iran

ABSTRACT

This paper proposes the use of interval observers and viability theory in fault detection and isolation (FDI). Viability theory develops mathematical and algorithmic methods for investigating the viability constraints characterisation of dynamic evolutions of complex systems under uncertainty. These methods can be used for checking the consistency between observed and predicted behaviour by using simple sets that approximate the exact set of possible behaviour (in the parameter or state space). In this paper, FDI is based on checking for an inconsistency between the measured and predicted behaviours using viability theory concepts and sets. Finally, an example is provided in order to show the usefulness of the proposed approach.

KEYWORDS

Fault detection; fault isolation; interval observer; viability theory

1. Introduction

Conventional feedback control systems are vulnerable to malfunctions in sensors, actuators or other system components. Therefore, diagnosing which kind of faults are developing is an important task to prevent physical damage and performance degradation. Fault detection and isolation (FDI) could also lead to more reliable and efficient systems (Mohajerpoor, Abdi, & Nahavandi, 2015). In this area of study, a lot of different methods have been proposed in the literature (Gao, Cecati, & Ding, 2015) including observer-based methods (Mondal, 2017; Wang, Yang, & Liu, 2007), parity space (Blesa, Puig, Saludes, & Fernández-Cantí, 2016; Ghaniee Zarch & Aliyari Shoorehdeli, n.d.), parameter estimation (Iurinic, Herrera-Orozco, Ferraz, & Bretas, 2016) and artificial intelligence methods (Li & Yang, 2014).

Set theory has been started to be used in fault detection and fault tolerant context since more than one decade (Puig, 2010). A system can switch among several modes (a healthy one and at least a faulty one). Using set-theoretic methods, it is possible to calculate sets which define healthy and faulty functioning. As long as there exists a (partial) separation between these sets, a FDI scheme can be designed (Stoican & Olaru, 2013). There are two main approaches that use set theory for fault detection: the set-membership (Blesa et al., 2016; Fernández-Cantí, Blesa, Puig, & Tornil-Sin, 2016) and the set-invariance approach (Hanafi, Seron, & De Dona, 2015). One of the most used techniques in the set-membership approach is based on interval observers. Interval observer-based fault detection (FD) consists in generating adaptive intervals for system outputs by considering the bounds of uncertainties, propagating their effect through the mathematical models of the system and testing the consistency between the predicted output intervals and the corresponding output measurements (Mazenc & Bernard, 2011; Meseguer, Puig, & Escobet, 2017). The main advantage of using interval observers in FDI is the inclusion of uncertainties of the system in the computations, which provides a robust method. The problem with this type of analysis is that the feasibility of FDI cannot be guaranteed a priori for all future time instants. This is because the sets are estimated at each iteration and they may conduct to empty sets (Stoican & Olaru, 2013). Another challenging issue is the computational burden of this approach that limits its practical implementation (Xu, Stoican, Puig, Ocampo-Martinez, & Olaru, 2013).

Another set-theoretic FDI approach is to consider invariant sets that can overcome difficulties with interval observer approach. In this approach, for each mode (healthy or faulty), an invariant set for the residual can be obtained (Olaru, De Doná, Seron, & Stoican, 2010). Once the system operates in steady state, it is possible to confine the residual to one of these invariant sets and, as long as all the invariant sets are disjoint, FDI can be performed. Most importantly, in the case that the invariant sets intersect, FD can still be done whenever the residual exits its healthy invariant set (Olaru et al., 2010; Seron, Zhuo, De Doná, & Martínez, 2008). Although both approaches

follow a similar FDI principle, set-invariance methods are more conservative in the size of sets as they consider uncertainties in off-line set computations. Also to the authors' best knowledge, the usage of invariant sets in transient state has not been reported in the literature until now. A detailed comparison between interval observer and set-invariance approaches has been presented in Xu et al. (2013).

Viability theory develops mathematical and algorithmic methods for investigating the viability constraints characterisation of dynamic evolutions of complex systems under uncertainty (Aubin, Bayen, & Saint-Pierre, 2011). Viability is a theory that until now has mostly been used in safety verification of control systems (Maidens, Kaynama, Mitchell, Oishi, & Dumont, 2013). It provides some concepts that are actually more general than those used in set theory for fault detection. Viability kernel is an acceptable tool for safety verification. However, the problem with this theory is how to compute the involved sets. Nowadays, some algorithms have been proposed that can approximate these sets (named kernels in viability theory) effectively. Viability theory has already been used in different areas of study as e.g. economics or biology (Aubin, 2013; Deffuant & Gilbert, 2011). Adapting the use of the viability concepts to address the FDI problem is the major goal of this paper. Moreover, these concepts will be related with those used of set theory that are already used for FDI.

The main contribution of this paper is to propose the combined use of interval observers and viability theory in FDI. Fault detection is based on checking for an inconsistency between the measured and predicted behaviours using viability theory concepts and sets. Fault isolation will be based on checking the faulty system into which faulty invariant kernel is evolving. The main advantage of combining interval observers with viability theory is that FDI can be guaranteed. Actually, by computing kernels in different modes of system operation (one healthy and at least one faulty mode) and showing their separability in offline computations, it can be guaranteed that the faults can be detected and isolated in online implementation. Another contribution of this study is to provide algorithms to find viability sets for non-linear systems that can be expressed in linear parameter varying (LPV) form. These algorithms are the extension of the idea proposed in Maidens et al. (2013) to LPV systems using zonotopic set representation.

Although the fault diagnosis problem in LPV systems has been studied widely, the proposed method has some advantages over existing approaches:

- A method for generating an adaptive threshold for evaluating the residual in different modes (one healthy and at least one faulty mode) dealing explicitly with uncertainties in the model.
- Fault detectability and isolability properties can be guaranteed as the separability of predefined healthy and faulty sets can be obtained offline.
- Set-based methods can provide a framework for fault recovery. In this case, the system can be analysed if the fault condition has been eliminated and the system works in healthy mode. This can be done by set definition in healthy and faulty situations (Seron, De Doná, & Olaru, 2012).

A preliminary form of the results presented in this paper appeared in a conference paper (Zarch, Puig, & Poshtan, 2017). In this paper, some improvements have been done to clarify the approach: First, the detailed algorithms for finding kernels in LPV systems have been presented in Section 3. Also, more results and comparisons are provided in Section 6 to show the effectiveness of the approach.

This paper is organised as follows. In Section 2, some definitions and preliminary concepts of viability theory and zonotopic sets are recalled. The algorithms for finding associated sets (kernels) are proposed in Section 3. The way how viability theory can be used in FDI is a task that is discussed in Section 4. In Section 5, the viability-based FDI approach is integrated with the interval observer approach. An illustrative example is provided in Section 6 in order to illustrate the proposed approach. In Section 7, concluding remarks are drawn.

2. Preliminary concepts

2.1. Concepts definition

In control engineering, a state space representation is a mathematical model of a physical system in the form of set of first-order differential equations represented by

$$\begin{cases} \mathscr{L}(x(t)) = f(x(t), u(t), w(t)) \\ x(t) \in X \\ u(t) \in U \\ w(t) \in W \end{cases} \tag{1}$$

where x are the state variables, u are the inputs and w are the disturbances. The time t ranges throughout a time range $[0, N]$ that can be either continuous or discrete. $\mathscr{L}$ is the differential operator corresponding to the given time domain (differentiation in the case of a continuous-time system and differencing in the case of a discrete-time system). It is assumed that the above system is defined in a proper open set $\mathrm{O} \subseteq \mathbb{R}^n$ and that there exists a globally defined solution for every initial condition $x(0) \in \mathrm{O}$. The

dynamic evolution of the system:

$$S : X \rightarrow \mathbb{C}\,(0, +\infty; X) \tag{2}$$

maps any initial state $x \in X$ to the set $S(x)$ of evolutions $x(.)$ starting from $x(0)$ and governed by Equation (1).

Viability theory goal is to prove if the dynamical system evolution (1) can be maintained inside a viability constraint set $K \subseteq \mathbb{R}^d$. Any trajectory of system (1) that leaves the set K at some point in time is considered to be no longer viable.

Definition 2.1 Viability Kernel (Aubin et al., 2011)**:** The viability kernel of K under the evolutionary system S is the set $Viab_S(k)$ of initial states $x(0) \in K$ from which starts at least one evolution $x(t) \in S(x)$ viable in K for all times $t \geq 0$:

$$Viab_S\,(K) \triangleq \left\{ \begin{array}{l} x\,(0) \in K | \exists x\,(.) \in S\,(x) \\ \text{such that} \quad \forall t \geqslant 0,\ x(t) \in K \end{array} \right\} \tag{3}$$

That is, from any point x_0 in the viability kernel starts at least one evolution that stays inside K forever. It is equivalent to say that all evolutions starting from a state belonging to the complement of the viability kernel of K leave the environment in finite time. Sometimes, from the engineering point of view, the existence of at least one solution in S is not enough, since nothing is said about all other possible solutions. Therefore, another important concept is defined known as the invariance kernel.

Definition 2.2 (Invariance Kernel; Aubin et al., 2011)**:** Let $K \subset X$ be an environment and $C \subset K$ be a target. The subset $Inv_S(K, C)$ of initial states $x(0) \in K$ such that all evolutions $x(t) \in S(x)$ starting at $x(0)$ are viable in K for all $t \geq 0$ or viable in K until they reach C in finite time is called the invariance kernel of K with target C under S.

$$Inv_S\,(K) \triangleq \{\, x\,(0) \in K | \forall x\,(.) \in S\,(x)\,, \forall t \geqslant 0,\ x(t) \in K\} \tag{4}$$

A state x_0 belongs to the invariance kernel of the environment K under an evolutionary system if all the evolutions starting from it are viable in K forever. In spite of viability kernel, invariance kernel can guarantee that every system trajectory will remain in the set forever. This concept is widely accepted as a useful tool for FDI (Seron & De Doná, 2015). Positive invariance in set theory has the same definition as invariance kernel. Viability kernel and weak positive invariance are also equivalent definitions in viability and set theories, respectively (Blanchini & Miani, 2008). Capture basin is another concept that has a wide range of applications, for example, in process control (Spiteri, Pai, & Ascher, 2000) and economics (Saint-Pierre, 2004).

Definition 2.3 Capture Basin(Aubin et al., 2011)**:** The capture basin of C (viable in K) under the evolutionary system S is the set $Capt_S(K, C)$ of initial states $x(0) \in K$ from which starts at least one evolution $x(t) \in S(x)$ viable in K on $[0, N)$ until the finite time N when the evolution reaches the target at $x(N) \in C$.

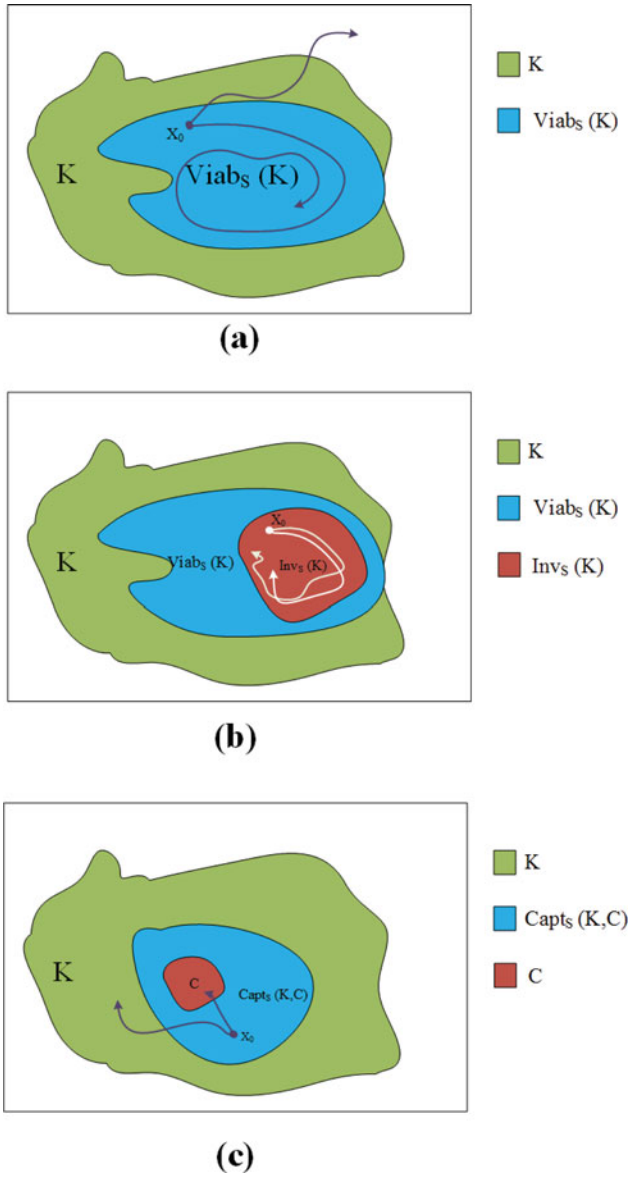


Figure 1. A sample of: (a) viability kernel; (b) invariance kernel; (c) capture basin.

From a state x_0 in the capture basin of the target C viable in the environment K starts at least one evolution viable in K until it reaches C in finite time. It is equivalent to say that, starting from a state belonging to the complement of the capture basin, all evolutions remain outside the target C until they leave the environment K.

Development of the methods for obtaining these three sets is still an important area and not an easy task (Maidens et al., 2013). A schematic diagram of these kernels is shown in Figure 1.

2.2. Zonotopic sets

There exist several families of geometric shapes which can be used to describe sets used in the viability theory with varying degrees of accuracy. An important limiting factor is the numerical reliability of their representation. That is, a particular family may be able to represent a great number of shapes but due to computationally expensive manipulations will be useless in practice. Usually there exists an inverse relation between flexibility of a given type of approximating sets and the numerical cost of the representation.

Zonotopes represent a particular class of polytopes which exhibit symmetry with respect to their centre. In realistic situations, often the constraints that are given in polytopic form have enough symmetry to be described as zonotopic sets. Even when this is not the case, zonotopic approximations may be constructed. For polytopic sets, Alamo, Bravo, and Camacho (2005) give the tightest approximations in fixed directions and Dang (2006) discusses an iterative algorithm. In Linhart (1989), it is proven that any Euclidean ball can be approximated arbitrarily close, in the sense of the Hausdorff distance, by a zonotope.

Definition 2.4 (Minkowski sum; Le, Stoica, Alamo, Camacho, & Dumur, 2013)**:** The Minkowski sum of two sets X and Y is defined by

$$X \oplus Y = \left\{x + y : x \in X, y \in Y\right\}$$

Definition 2.5 (Zonotope; Le et al., 2013)**:** An m-order zonotope Z is defined as

$$Z = c \oplus H\beta^m$$

where c and H are called the centre and segment matrix (also generator matrix), respectively.

Definition 2.6 (Interval hull; Le et al., 2013)**:** The interval hull $\square X$ of a closed set X is the smallest interval box that contains X.

Given a zonotope $X = \pi \oplus H\beta^{m_z}$, its interval hull can be easily computed by

$$\square X = \{x \quad \forall i = 1, \ldots, n : |x_i - \pi_i| \leqslant \|H_i\|_1\}$$

where x_i and π_i are the ith components of x and π, respectively, and H_i is the ith row of H. $\|H_i\|_p$ is the p-norm of vector $H_i = [H_{i,1}, \ldots, H_{i,n}]$ that is defined by

$$\|H_i\|_p := \left(\sum_{j=1}^{n} \left|H_{i,j}\right|^p\right)^{1/p}$$

Zonotopes have a lot of appealing properties (Le et al., 2013). Some properties that have been used in the Section 3 are introduced in Appendix.

3. Computing viability kernels using zonotopes

3.1. Problem set-up

In this section, the way how the viability kernels recalled in previous section can be computed using zonotopes is presented.

The proposed approach considers that the non-linear system (1) can be transformed to the (quasi-)LPV form in discrete-time

$$x(t+1) = A(\rho(t))\,x(t) + B(\rho(t))\,u(t) + E(\rho(t))\,w(t) \tag{5}$$

where $x(t) \in X$ is state, $u(t) \in U$ is control input and $w(t) \in W$ are unknown inputs (disturbances). The bounding sets X, U and W are defined as

$$\begin{aligned} X &= \left\{x \in \mathbb{R}^n : \left|x - x^c\right| \leqslant \bar{x}, x^c \in \mathbb{R}^n, \bar{x} \in \mathbb{R}^n\right\} \\ U &= \left\{u \in \mathbb{R}^m : \left|u - u^c\right| \leqslant \bar{u}, u^c \in \mathbb{R}^m, \bar{u} \in \mathbb{R}^m\right\} \\ W &= \left\{w \in \mathbb{R}^q : \left|w - w^c\right| \leqslant \bar{w}, w^c \in \mathbb{R}^q, \bar{w} \in \mathbb{R}^q\right\} \end{aligned}$$

where x^c, u^c, w^c, $\bar{x}$, $\bar{u}$ and $\bar{w}$ are constant vectors. The sets X, U and W can be rewritten as zonotopes

$$\begin{aligned} X &= x^c \oplus H^{\bar{x}}\beta^n \\ U &= u^c \oplus H^{\bar{u}}\beta^m \\ W &= w^c \oplus H^{\bar{w}}\beta^q \end{aligned}$$

where $H^{\bar{x}} \in \mathbb{R}^{n\times n}$, $H^{\bar{u}} \in \mathbb{R}^{m\times m}$ and $H^{\bar{w}} \in \mathbb{R}^{q\times q}$ are diagonal matrices with their diagonal entries composed of $\bar{x}$, $\bar{u}$ and $\bar{w}$, respectively. It must be noted that the parameters of the sets X and U have been determined using physical constraints of the system, but the parameters of the uncertainties set W have been determined using obtained data from the system (Alamo et al., 2005; Brito, 2009). The parameter ρ is a time-varying parameter whose measurement is available online and that is used to hide non-linear terms in Equation (5). Due to several possibilities of assignment, the result of the transformation is non-unique. The number of the associated scheduling variables increases rapidly with the system order. As it involves no approximation of the system dynamics, efficient modelling solutions can be achieved in many applications (Gáspár, Szabo, & Bokor, 2007; Tóth, 2010).

3.2. Invariance and viability kernel

Reachability analysis identifies the set of states backward (forward) reachable by a constrained dynamical system from a given target (initial) set of states. The notions of maximal and minimal reachability analysis were introduced in Mitchell (2007). Their corresponding constructs differ in how the time variable and the bounded input are quantified. In the formation of the maximal reachability construct, the inputs try to steer as many states as possible to the target set. On the other hand, in the formation of the minimal reachability construct, the trajectories reach the target set regardless of the input applied. Based on these differences, the maximal and minimal reachable sets and tubes (the set of states traversed by the trajectories over the time horizon (Mitchell, 2007)) are formed.

Definition 3.1 (Forward maximal reachable set)**:** The forward maximal reachable set at time instant t is the set of states for which there exists an input such that the trajectories emanating from initial states in T reach that set exactly at time instant t:

$$\text{Reach}_t^F(t) \triangleq \left\{x(t) \in \mathbb{R}^n | \exists u(.) \in U_{[0,t]}, x(0) \in T\right\} \tag{6}$$

Definition 3.2 (Backward maximal reachable set)**:** The backward maximal reachable set at time instant t is the set of initial states for which there exists an input such that the trajectories emanating from those states reach T exactly at time instant t:

$$\text{Reach}_t^B(t) \triangleq \left\{x(0) \in \mathbb{R}^n | \exists u(.) \in U_{[0,t]}, x(t) \in T\right\} \tag{7}$$

The forward reachable set over a single time step is computed as

$$\text{Reach}_1^F(X) = A(\rho(t))X \oplus B(\rho(t))U \oplus E(\rho(t))W$$

Following computation algorithm in Montes de Oca, Puig, and Blesa (2012), we can find this reachable set using zonotopes by

$$X_{t+1} = \text{Reach}_1^F(X_t) = x_{t+1}^c \oplus H_{t+1}^{\bar{x}}\beta^r \tag{8}$$

where

$$\begin{aligned}
x_{t+1}^c &= \text{mid}(A(\rho(t)))x_t^c + \text{mid}(B(\rho(t)))u^c \\
&\quad + \text{mid}(E(\rho(t)))w^c \\
H_{t+1}^{\bar{x}} &= \left[J_1\ J_2\ J_3\ J_4\ J_5\ J_6\right] \\
J_1 &= \text{seg}\left(\diamond A(\rho(t))H_t^{\bar{x}}\right) \\
J_2 &= \frac{\text{diam}(A(\rho(t)))}{2}x_t^c
\end{aligned}$$

Algorithm 1. Invariance kernel computation

```
K_0 ← X
t ← 1
while t ≤ N do
  if K_t = ∅ then
  K_N ← ∅
  break
  end if
  if K_t = K_{t−1} then
  K_N ← K_t
  break
  end if
  K_t ← Reach_1^F(K_{t−1})
  t ← t + 1
end while
return (K_N)                    K_N = Inv(X)
```

$$\begin{aligned}
J_3 &= \text{seg}\left(\diamond B(\rho(t))H^{\bar{u}}\right) \\
J_4 &= \frac{\text{diam}(B(\rho(t)))}{2}u^c \\
J_5 &= \text{seg}\left(\diamond E(\rho(t))H^{\bar{w}}\right) \\
J_6 &= \frac{\text{diam}(E(\rho(t)))}{2}w^c
\end{aligned} \tag{9}$$

where 'mid' denotes the centre and 'diam' the diameter of the interval, $\diamond$ is zonotope inclusion introduced in the Appendix (Property A.1.) and $\text{seg}(Q) = H$ considering that $Q = \pi + H\beta^r$ is a zonotope (Montes de Oca et al., 2012). To compute the invariance kernel, this forward reachable set is calculated step by step. This reachable tube will finally converge toward the invariance kernel

$$\text{Inv}(X) = \bigoplus_{t=0}^{\infty} \text{Reach}_1^F(X_t)$$

It is important to note that the set of estimated states using this method has an increasing number of segments generating the zonotope $\text{Reach}_1^F(X)$. In order to control the domain complexity, a reduction step is thus implemented. Here, we use the method proposed in Combastel (2003) to reduce the zonotope complexity. Algorithm 1 is proposed in order to calculate invariance kernel based on above discussion.

In Maidens et al. (2013), the following iterative approach is proposed to approximate viability kernel using maximal reach sets and such that the sequence defined in Equation (10) will converge to viability kernel as t goes to infinity

$$\begin{aligned}
K_h^0 &= K \\
K_h^{t+1} &= K_h^0 \cap \text{Reach}_1^B\left(K_h^n\right)
\end{aligned} \tag{10}$$

Algorithm 2. Viability kernel computation

$K_0 \leftarrow X$
$t \leftarrow 1$
while $t \leq N$ do
  if $K_t = \varnothing$ then
  $K_N \leftarrow \varnothing$
  break
  end if
  if $K_t = K_{t-1}$ then
  $K_N \leftarrow K_t$
  break
  end if
  $L \leftarrow \text{Reach}_1^B(K_{t-1})$ (see Equation (12))
  $K_{t+1} \leftarrow K_0 \cap L$ (see Property A.2.)
  $t \leftarrow t+1$
end while
return (K_N) $K_N = \text{Viab}(X)$

Algorithm 3. Capture basin computation

$K_0 \leftarrow C$
$t \leftarrow 1$
while $t \leq T$ do
  if $K_t = \varnothing$ then
  $K_T \leftarrow \varnothing$
  break
  end if
  $K_t \leftarrow \text{Reach}_1^B(K_{t-1})$
  $t \leftarrow t+1$
end while
$K_N = K_T \cap X$
return (K_N) $K_N = \text{Capt}(X, C)$

The backward reachable set for system (5) over a single time step is computed as

$$\text{Reach}_1^B(X) = A(\rho(t))^{-1}\{X \oplus (-B(\rho(t)))\,U \oplus (-E(\rho(t)))\,W\} \tag{11}$$

Here $A^{-1}(.)$ denotes the pre-image of a set under the map $A : \mathbb{R}^n \to \mathbb{R}^n$. Note that we consider that A is invertible. This is a fair assumption because we are mainly concerned with discrete-time systems that arise from the discretisation of continuous time systems. Such systems have a dynamics matrix of the form $A = \exp(A_c)$ which is always invertible (Maidens et al., 2013). Following the computation algorithm in Montes de Oca et al. (2012), we can find this reachable set using zonotopes as

$$\text{Reach}_1^B(X_t) = X_{t-1} = x_{t-1}^c \oplus H_{t-1}^{\bar{x}}\beta^n \tag{12}$$

where

$$\begin{aligned}
x_{t-1}^c &= \text{mid}\left(A(\rho(t))^{-1}\right)x_t^c + \text{mid}\left(-A(\rho(t))^{-1}B(\rho(t))\right)u^c \\
&\quad + \text{mid}\left(-A(\rho(t))^{-1}E(\rho(t))\right)w^c \\
H_{t-1}^{\bar{x}} &= \left[J_1\ J_2\ J_3\ J_4\ J_5\ J_6\right] \\
J_1 &= \text{seg}\left(\diamond A(\rho(t))^{-1}H_t^{\bar{x}}\right) \\
J_2 &= \frac{\text{diam}\left(A(\rho(t))^{-1}\right)}{2}x_t^c \\
J_3 &= \text{seg}\left(\diamond - A(\rho(t))^{-1}B(\rho(t))\,H^{\bar{u}}\right) \\
J_4 &= \frac{\text{diam}\left(-A(\rho(t))^{-1}B(\rho(t))\right)}{2}u^c \\
J_5 &= \text{seg}\left(\diamond - A(\rho(t))^{-1}E(\rho(t))\,H^{\bar{w}}\right) \\
J_6 &= \frac{\text{diam}\left(-A(\rho(t))^{-1}E(\rho(t))\right)}{2}w^c
\end{aligned} \tag{13}$$

where 'mid' denotes the centre and 'diam' the diameter of the interval, $\diamond$ is zonotope inclusion (Property A.1.) and $\text{seg}(Q) = H$ considering that $Q = \pi + H\beta^r$ is a zonotope (Montes de Oca et al., 2012). For the computation of viability kernel, using this backward reachable set in Equation (10), the reachable tube is calculated step by step. This tube will finally converge toward the viability kernel. Here, a similar algorithm (Algorithm 2) for computing viability kernel for system (5) based on Equations (10) and (12) is presented.

3.3. Capture basin

Finally, for computation of the capture basin, we can find backward reachable tube using Equation (12) for desired time steps. Final reach set is the capture basin. Actually, we must find backward reachable tube for each time instant. In this manner, Algorithm 3 is proposed.

4. FDI using viability theory

4.1. Principles of FDI using set theory

A fault in a dynamical system is a deviation of the system structure or the system parameters from the nominal situation. The principle of model-based fault detection is to test whether the measured system inputs and outputs are consistent with the system behaviour described by a faultless model. If the measurements are inconsistent with the model of the healthy system, the existence of a fault is proved.

For a dynamical system (1), consider that the output $y(t)$ is the reaction of the plant to the input $u(t)$. The pair (u, y) is called input/output (I/O) pair. (U, Y) is the set of all possible I/O pairs. Faults lead to deviations of the dynamical input/output (I/O) properties of the plant from the nominal ones, and hence, change the performance of the closed-loop system which further results in a degradation or even the loss of the system function.

A fault changes the system behaviour as illustrated in Figure 2. If the system works in set A, it is working in healthy mode. The system behaviour can be moved by a fault towards the set B. If a common input u is applied to the healthy and faulty systems, then both systems answer

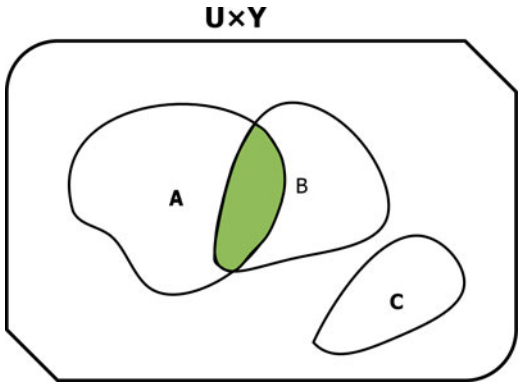


Figure 2. System behaviour with and without fault.

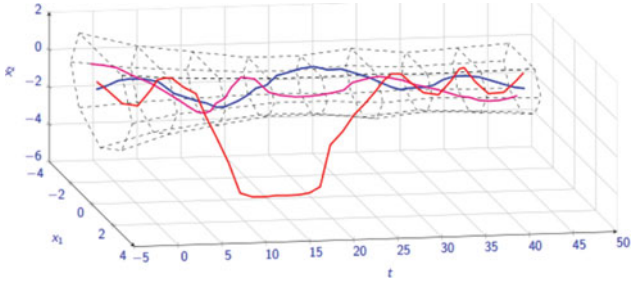


Figure 3. Basic idea of fault detection using sets (Stoican, 2011).

with different outputs $A = (u, y_A)$ and $B = (u, y_B)$. This change in the system behaviour makes the detection of the fault possible, unless the faulty I/O pair lies in the intersection of A and B.

For isolation of more than one fault, in each faulty situations, a proper set must be determined. For example, in Figure 2, there are two faulty cases, f_1 and f_2. If the I/O pair lies in the set B, our suggestion is that the system is working in faulty mode f_1. But, if the I/O pair lies in set C, it is more probable that the system works in the faulty mode f_2. If the I/O pair is in the intersection and outside of these predefined sets, nothing can be said about the system status. The basic idea of using sets in fault diagnosis is to find a predefined set that can assure safety if the system works on that set (see Figure 3) and the different faulty situations (sets) could be separated.

In the framework of the analytical redundancy concept, the process model that is driven by the same process input will run in parallel to the process. It is reasonable to expect that, in the fault-free case, both process and its model shows similar behaviour. Comparing outputs leads to a signal that can be used for fault detection. The difference between the measured process variables and their estimates is called residual. This residual signal carries information about faults that have occurred in the system.

Residual signal information is generally affected by model uncertainties and unknown disturbances. Moreover, fault isolation and identification techniques need additional residual analysis. Hence, the main problem with the application of model-based fault diagnosis techniques can be expressed as extracting the needed and useful information about the faults of interest from residual signal. This step is called residual evaluation and can be performed using set-based approaches (Puig, 2010).

4.2. Interval observer approach

Consider that the system to be monitored (1) can be described by an LPV uncertain dynamic model in a state-space form as follows:

$$x(t+1) = A(\rho)x(t) + B(\rho)F_u(t)u(t) + w(t)$$
$$y(t) = C(\rho)F_y(t)x(t) + \eta(t) \qquad (14)$$

where $x(t)$ are the states, $y(t)$ are the plant measured outputs, $u(t)$ are the control inputs, $w(t)$ are the disturbances and $\eta(t)$ are the measurement noises. Disturbances and noises are both assumed to be unknown but bounded, i.e.

$$w_i \in \left[\underline{\delta}_i, \bar{\delta}_i\right], \eta_i \in \left[\underline{\sigma}_i, \bar{\sigma}_i\right]$$

The bounds of the disturbances and noises are obtained from practical data. If these bounds are very large, it means that the system can go far away from steady state. Therefore, it may lead to misdetection of small faults. This is the drawback with all set-based methods (Puig et al., 2006). This problem has been addressed in some other references by computing minimum detectable faults (Pourasghar, Puig, & Ocampo-Martinez, 2016). $F_u(t)$ and $F_y(t)$ are actuator and sensor faults, respectively. Their values can range in the interval [0,1) in failure mode to 1 in healthy mode. The state, input, output matrices are $A(\rho)$, $B(\rho)$ and $C(\rho)$, respectively. The parameter ρ is a time-varying parameter that can be measured/estimated and whose value belongs to Θ that is a bounded set (of interval box type) such that for each component $\left[\underline{\rho}_i, \overline{\rho}_i\right]$. This is why the resulting model is known as an interval model (Montes de Oca et al., 2012). The set Θ contains all possible values of ρ when the system operates normally. Note that when the parameters $\boldsymbol{\theta}$ are scheduled with the operating point using some known scheduling function and variable then system (14) is known as a *linear parameter varying* (LPV) system (Rugh & Shamma, 2000). Intervals for uncertain parameters can also be inferred from real data as discussed in Puig (2010).

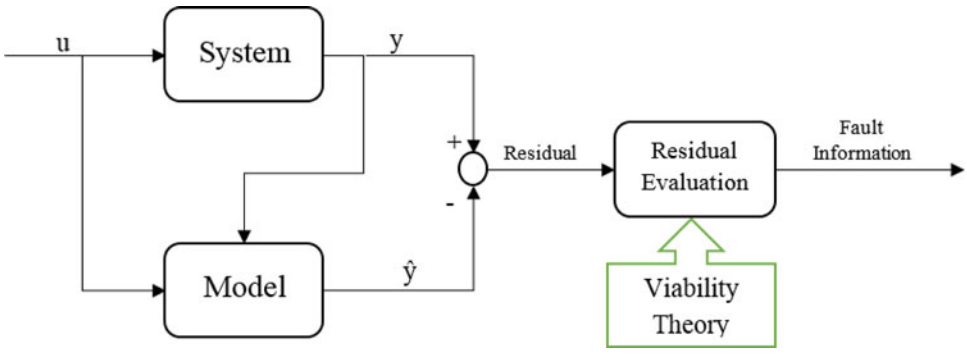


Figure 4. A general fault detection scheme.

The residual vector, known also as *analytical redundant relation*, defined as the difference between measured $y(t)$ and predicted system outputs $\hat{y}(t)$

$$r(t) = y(t) - \hat{y}(t) \tag{15}$$

is usually used to check the consistency. Ideally, the residuals should only be affected by the faults. However, the presence of disturbances, noise and modelling errors causes the residuals to become non-zero and thus interferes with the detection of faults. Therefore, the fault detection procedure must be robust against these undesired effects (Chen & Patton, 2012). In case of modelling a dynamic system using an interval model, the predicted output is described by a set that can be bounded at any iteration by an interval using an observer scheme (Montes de Oca et al., 2012)

$$\hat{y}_i(t) \in [\underline{\hat{y}}_i(t), \overline{\hat{y}}_i(t)]$$

in a non-faulty case. Such interval is computed independently for each output (neglecting couplings between outputs) as follows:

$$\underline{\hat{y}}_i(t) = \min_{\rho \in \Theta}(\hat{y}_i(t, \rho)) \text{ and } \overline{\hat{y}}_i(t) = \max_{\rho \in \Theta}(\hat{y}_i(t, \rho))$$

Such interval can be computed using zonotopes (Alamo et al., 2005; Montes de Oca et al., 2012). Then, the fault detection test is based on propagating the parameter uncertainty to the residual, and checking if

$$y(t) \in \left[\underline{\hat{y}}(t) - \sigma, \overline{\hat{y}}(t) + \sigma\right]$$

where σ is the noise bound. Equivalently, previous test can be formulated in terms of the residual checking if

$$0 \in \left[\underline{r}(t), \overline{r}(t)\right] = y(t) - \left[\underline{\hat{y}}(t) - \sigma, \overline{\hat{y}}(t) + \sigma\right]$$

holds or not. In case it does not hold, a fault can be indicated. This test is named as *direct test*. According to Isermann (2006), parity equations and observer approaches are more suitable for additive faults while parameter estimation approach is better suited for multiplicative (parametric) faults.

4.3. Fault detection using viability theory

In this section, the main results regarding fault detection using viability theory are derived. The FDI problem presented in previous section using sets can be addressed using viability theory concepts introduced in Section 2. These concepts can be used for evaluating the residual as it is shown in Figure 4 and illustrated in the following subsections.

4.3.1. Steady state

As with all the FDI methods based on the set invariance approach (Stoican, 2011), we can use invariance kernel for fault detection. Therefore, fault detection criteria in steady state can be summarised as follows.

Theorem 4.1 (Fault detection criteria in steady state)**:** *Consider system* (14) *and a residual signal* (15)*. A fault* (F_u *or* F_y) *in the system* (14) *in steady state can be detected at time instant* t_f *if*

$$r\left(t_f\right) \notin Inv_{S_r}(X)$$

Proof: A system without fault (5) works in the invariance kernel as discussed in Seron et al. (2008). Hence, residual signal must lie in the invariance kernel of the system residual under healthy functioning, i.e.

$$r(t) \in Inv_{S_r}(X) \tag{16}$$

in which the evolutionary system $S_r : X \to \mathbb{C}(0, +\infty; X)$ maps any initial state $x \in X$ to the set $S_r(x)$ of evolutions $x(.)$ starting from $x(0)$ and governed

by Equation (15). If the residual exits its healthy invariant set, it indicates that a fault has occurred. □

Corollary 4.1: *Considering $F(t)$ could be $F_u(t)$ or $F_y(t)$. Minimum detectable fault in steady state $F^{ss}_{\min}(t)$ is the smallest value of $F(t)$ that makes the residual $r(t)$ going outside the invariance kernel $Inv_{S_r}(X)$.*

$$F^{ss}_{\min}(t) = \min_{r(t)\notin Inv_{S_r}(X)} F(t) \tag{17}$$

4.3.2. Transient

The system behavior in transient mode is different from steady state and invariance kernel is no longer applicable for fault detection.

Theorem 4.2 (Fault detection criteria in transient without time constraint)**:** *Consider system (14) and a residual signal (15). A fault (F_u or F_y) in the system (14) in transient state can be detected at time instant t_f if*

$$r(t_f) \notin Viab_{S_r}(X)$$

Proof: In transient state, because changes in system (14) states are somewhat unpredictable and also, we do not know initial state of system, we cannot use invariance kernel in fault detection appropriately. In this situation, the question to be answered is: Is there any possible action that brings our system near steady state? Translating this question to the concepts in viability theory, we can use viability kernel for fault detection

$$r(t) \in \text{Viab}_{S_r}(X) \tag{18}$$

That is, while the system is in the viability kernel means that the system can find a way to be safe. Actually in this manner, we cannot say anything about convergence of the system toward steady-state. Note that in the constructing viability kernel, system constraints are considered. Therefore, being in viability kernel also means the system do not violate constraints. Satisfying constraints and having the opportunity to come back to steady state can ensure us that our system is in healthy functioning. □

Corollary 4.2: *Minimum detectable fault in transient without time constraint $F^{t}_{\min}(t)$ is the smallest value of $F(t)$ that makes the residual $r(t)$ going outside the viability kernel $Viab_{S_r}(X)$.*

$$F^{t}_{\min}(t) = \min_{r(t)\notin Viab_{S_r}(X)} F(t) \tag{19}$$

In most practical cases, we have time constraints in the transient mode, which means system must come back near steady state in finite time. In this situation, we can use capture basin for fault detection.

Theorem 4.3 (Fault detection criteria in transient with time constraint)**:** *Consider system (14) and a residual signal (15). A fault (F_u or F_y) in the system (14) in transient state with time constraint can be detected at time instant t_f if*

$$r(t_f) \notin Capt_{S_r}(X, C)$$

Proof: If system (14) has a time constraint, it means that the system has a limited time N to reach the target. Hence, being in the capture basin means that the system can find a way to come back to the target in limited time instants.

$$r(t) \in \text{Capt}_{S_r}(X, C) \tag{20}$$

Therefore, working outside this set can raise a fault alarm.

Note that a target can be chosen arbitrarily according to the application, which means we can choose as e.g. the invariance kernel . □

Corollary 4.3: *Minimum detectable fault in transient with time constraint $F^{tt}_{\min}(t)$ is the smallest value of $F(t)$ that makes the residual $r(t)$ going outside the capture basin $Capt_{S_r}(X, C)$.*

$$F^{tt}_{\min}(t) = \min_{r(t)\notin Capt_{S_r}(X,C)} F(t) \tag{21}$$

4.4. Fault isolation using viability theory

Consider that our system can have $i(=0, 1, 2, ..., n_s)$ different states: The first one ($i = 0$) related to the healthy mode and the others are different faulty modes to be detected. Residual equation is written for every fault scenario as follows:

$$r_i(t) = y_i(t) - \hat{y}_i(t) \tag{22}$$

In this situation, consider that we can construct a kernel for every types of fault. Faults can be isolated if kernels are separable. Hence, the condition for faults to be isolable is that those kernels are separable.

Lemma 4.1 (Guaranteed fault isolability criteria)**:** *Consider system (14) and residual signals (22). Faults (F_u or F_y) in the system (14) are guaranteed to be isolated in different modes of system operation (steady and transient states) if*

$$Inv_{S_{r_0}}(K) \cap Inv_{S_{r_1}}(K) \cap ... = \emptyset$$
(in steady state)

$$Viab_{S_{r_0}}(K) \cap Viab_{S_{r_1}}(K) \cap ... = \emptyset$$
(in transient state without time constraint)

$$Capt_{S_{r_0}}(K, C) \cap Capt_{S_{r_1}}(K, C) \cap ... = \emptyset$$
(in transient state with time constraint)

Considering this condition, for example if

$$r(t) \in \mathrm{In}\upsilon_{S_{r_1}}(K)$$

it means that the system works in mode 1 in steady state, and that is actually a faulty situation. If $\mathrm{In}\upsilon_{S_{r0}} \cap \mathrm{In}\upsilon_{S_{r1}} \neq \emptyset$, then for every $r_i(t) \in \mathrm{In}\upsilon_{S_{r0}} \cap \mathrm{In}\upsilon_{S_{r1}}$, nothing can be said about the healthy or faulty functioning of the system.

5. Integration with interval observers

5.1. Interval observers

The application of the viability theory to FDI requires the generation of the residual signal. Residuals can be generated in many ways as discussed in Blanke, Kinnaert, Lunze, and Staroswiecki (2006). A particular well-established way of residual generation is based on the use of interval observers (Montes de Oca et al., 2012). Viability theory is well suited for dealing with FDI in non-linear systems since most of the concepts have been developed in this context (see Section 2). Designing observers for non-linear systems is a difficult problem. A possible approach to deal with the observer design for non-linear systems is by approximating them as LPV system (Seron & De Doná, 2015) and applying LMI-based (Linear Matrix Inequality) designs. For design purposes, system matrices of the LPV model (14) can be expressed in polytopic form

$$[A(\rho), B(\rho), C(\rho)] = \sum_{j=1}^{N} \zeta_j(\rho) \left[A_j, B_j, C_j\right]$$

for certain constant matrices A_j, B_j, C_j and continuous functions ζ_j such that $\zeta_j(\rho) \geq 0$ and $\sum_{j=1}^{N} \zeta_j(\rho) = 1$ for all ρ. Assume that the pairs (A_j, B_j) are stabilisable and the pairs (A_j, C_j) are detectable for $j = 1, ..., N$. An LPV observer for this system is defined as

$$\begin{aligned} \hat{x}(t+1) &= A(\rho)\hat{x}(t) + B(\rho)u(t) \\ &\quad + L(\rho)\left(y(t) - \hat{y}(t)\right) + \breve{\omega}(t) \\ \hat{y}(t) &= C(\rho)\hat{x}(t) + \breve{\eta}(t) \end{aligned} \tag{23}$$

where $u(t) \in U$ is the measured system input vector, $\hat{x}(t) \in \hat{X}$ is the estimated system state vector, $\hat{y}(t)$ is the estimated system output vector, the uncertain variables $\breve{\omega}(t)$ and $\breve{\eta}(t)$ are used to describe the effect of variable $w(t)$ and $\eta(t)$ on the plant (14), respectively. The uncertain variables $\breve{\omega}(t)$ and $\breve{\eta}(t)$ are different from $w(t)$ and $\eta(t)$, but are defined to have the same bounds, respectively (i.e. $\breve{\omega}(t) \in W$ and $\breve{\eta}(t) \in V$).

According to Montes de Oca et al. (2012), taking into account uncertainty bounds when obtaining the observer estimation, intervals that bound the estimated state and output can be generated. This type of observer is known as an interval observer and it is a well-accepted approach in robust FDI.

The next issue to be addressed here is how to find observer gain in order to stabilise it. The following theorem provides the design procedure to stabilise observer (23).

Theorem 5.1: *Consider an observer of the form (23) for an LPV system (5). To guarantee the stability of the observer, the observer gains L_j can be determined through the following LMIs:*

$$\begin{bmatrix} -rX_j & aX_j + X_j^T A_j(\rho) - W_j^T C_j(\rho) \\ \left(a + A_j(\rho)^T\right) X_j - C_j(\rho)^T W_j & -rX_j \end{bmatrix} < 0 \tag{24}$$

where $L_j = \left(W_j X_j^{-1}\right)^T$. Now, the observer gains $L_j(\rho)$ will be interpolated to obtain the interval LPV observer gain as

$$L(\rho) = \sum_{j=1}^{N} \zeta_j(\rho) L_j(\rho)$$

Proof: See Montes de Oca et al. (2012). □

5.2. Interval observers and set invariance

Recently, a lot of efforts have been done in order to use interval observers and set invariance in FDI; see for example Mazenc and Bernard (2011), Seron and De Doná (2015). The principle for FDI in both approaches is similar. In interval observer based approach, zero must be inside the residual set to assure healthy functioning. In this approach, zero is fixed but the residual sets are computed online. They have the advantage of considering noise and uncertainty in transient phase of fault detection, which makes this method more robust, but more computationally demanding. Set invariance approach is more conservative than the one based on interval observers, but it provides FDI guarantees. In this approach, invariant set is fixed and determined offline but the residual is calculated online. In Xu et al. (2013), a detailed comparison between these two methods is presented. Here, we propose using viability theory in order to make it possible to detect faults in transient states combining interval observers for generating residuals and viability theory concepts for evaluating them in FDI as described in Section 5. The main motivation for this integration is because viability offers a general framework for dealing with non-linear systems. Another fact is that in

viability theory, sets can be defined without considering any specific shape (as e.g. ellipsoids or zonotopes), which makes it less conservative, and more general implementation can be carried out.

6. Illustrative example

In this section, the viability theory based FDI approach developed in previous sections is applied to a two-tank system described by a continuous non-linear model (Seron & De Doná, 2015)

$$\begin{aligned}\dot{h}_1(t) &= -\left({}^{s}/_{S}\right)\sqrt{2g}\sqrt{h_1(t)} + \left({}^{\kappa}/_{S}\right)u(t) + w_1(t)\\ \dot{h}_2(t) &= \left({}^{s}/_{S}\right)\sqrt{2g}\left[\sqrt{h_1(t)} - \sqrt{h_2(t)}\right] + w_2(t)\end{aligned} \quad (25)$$

where $u(t)$ is the voltage applied to the pump, $h_1(t)$, $h_2(t)$ are system states, $w_1(t)$, $w_2(t)$ are bounded state perturbation and the parameters are as follows: $S = 15.5179$ cm^2 is the cross-sectional area of the tanks; $s = 0.1781$ cm^2 is the cross section of the tanks outflow orifice; $\kappa = 3.3$ cm^3 Vs is the gain of the pump; $g = 981$ cm s^2 is the gravitational constant. After Euler discretisation with sampling period $\tau = 1$ s, the whole system with considered faults is formulated in its quasi-LPV form through parameter non-linear embedding approach

$$\begin{aligned}x(t+1) &= A(\rho(t))\,x(t) + BF_u u(t) + Ew(t)\\ y(t) &= CF_y x(t) + \eta(t)\end{aligned} \quad (26)$$

where

$$\begin{aligned}x(t) &= \begin{bmatrix} h_1(t) & h_2(t)\end{bmatrix}^T\\ A(\rho(t)) &= I + \tau\begin{bmatrix} -\rho_1(t) & 0\\ \rho_1(t) & -\rho_2(t)\end{bmatrix}\\ B &= \tau\begin{bmatrix}\kappa/S & 0\end{bmatrix}^T\\ C &= \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\\ E &= \tau\begin{bmatrix}1 & 1\end{bmatrix}^T\\ \eta(t) &= \begin{bmatrix}\eta_1(k) & \eta_2(k)\end{bmatrix}^T\end{aligned}$$

$w(t)$ is disturbance, $\eta_1(t)$ and $\eta_2(t)$ are measurement noises that are considered to be bounded by means of zonotopes

$$\begin{aligned}w &= 0 \oplus 0.01\beta\\ \eta_i &= 0 \oplus 0.01\beta, \quad i = 1, 2\end{aligned}$$

Actuator and sensor faults are modelled by F_u and F_y. They can have values in the range [0, 1], where one is related to the healthy operation of the system and zero is full fault. Any value between these two is considered as partial fault. The varying parameters can be shown as

$$\rho_i(t) = \frac{s}{S}\sqrt{\frac{2g}{h_i(t)}}, \quad i = 1, 2$$

The system states and parameters are bounded as

$$\begin{aligned}h_i^{\min} = 0.5 \rightarrow \rho_i^{\min} = 0.0928 = \underline{\rho}_i\\ h_i^{\max} = 30 \rightarrow \rho_i^{\max} = 0.7189 = \bar{\rho}_i\end{aligned}$$

Hence, an LPV model with convex polytopic description can be obtained from Equation (26) by taking

$$A_1 = A\left(\underline{\rho}_1, \underline{\rho}_2\right), A_2 = A\left(\underline{\rho}_1, \bar{\rho}_2\right), A_3 = A\left(\bar{\rho}_1, \underline{\rho}_2\right),$$
$$A_4 = A\left(\bar{\rho}_1, \bar{\rho}_2\right)$$

and

$$A(\rho(t)) = \sum_{i=1}^{4}\alpha_i A_i$$

These weights can be computed using following equations, which is valid for the case of polytopes with four vertices:

$$\begin{aligned}\rho &= \frac{h_1^{\max} - h_1(t)}{h_1^{\max} - h_1^{\min}}\\ \lambda &= \frac{h_2(t) - h_2^{\min}}{h_2^{\max} - h_2^{\min}}\\ \alpha_1 &= \rho\lambda, \quad \alpha_2 = (1-\rho)\lambda\\ \alpha_3 &= \rho(1-\lambda), \quad \alpha_4 = (1-\rho)(1-\lambda)\end{aligned}$$

Here, for the goal of FDI, two residual signals can be defined. First, output of the system is divided in two parts

$$\begin{aligned}y_1(t) &= C_1F_{y1}x(t) + \eta_1(t)\\ y_2(t) &= C_2F_{y2}x(t) + \eta_2(t)\end{aligned}$$

where

$$\begin{aligned}C_1 &= \begin{bmatrix}1 & 0\end{bmatrix}\\ C_2 &= \begin{bmatrix}0 & 1\end{bmatrix}\end{aligned}$$

and F_{y1} and are first and second output faults, respectively. Based on these two outputs, two residuals are defined as

$$r_1(t) = y_1(t) - \hat{y}_1(t) \quad (27)$$

$$r_2(t) = y_2(t) - \hat{y}_2(t) \quad (28)$$

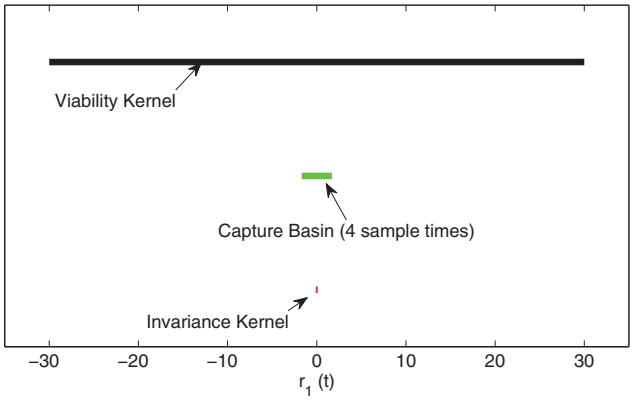


Figure 5. Kernels for residual signal $r_1(t)$ in healthy mode.

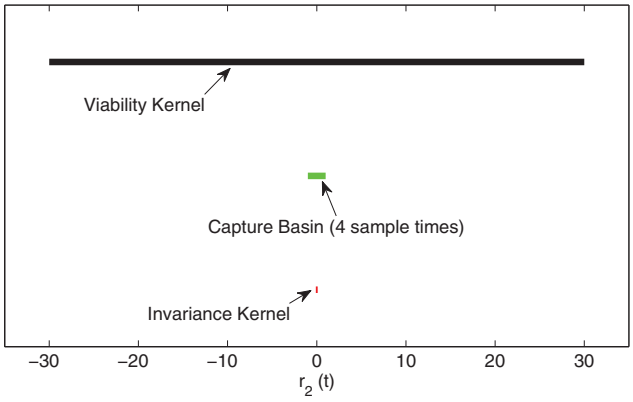


Figure 6. Kernels for residual signal $r_2(t)$ in healthy mode.

where $\hat{y}_1(t)$ and $\hat{y}_2(t)$ are estimated values of $y_1(t)$ and $y_2(t)$, respectively. For deriving $\hat{y}_1(t)$, considering Equation (23) as observer and using Equation (24), observer gains are derived as

$$L_1 = \begin{bmatrix} 0.3532 \\ 0.7561 \end{bmatrix}, L_2 = \begin{bmatrix} 0.3532 \\ 0.7936 \end{bmatrix}, L_3 = \begin{bmatrix} 0.9345 \\ 0.0924 \end{bmatrix},$$
$$L_4 = \begin{bmatrix} 0.9345 \\ 0.2042 \end{bmatrix}$$

In the same manner, observer gains for the second observer $\hat{y}_2(t)$ are

$$L_1 = \begin{bmatrix} 0.4146 \\ 1.1889 \end{bmatrix}, L_2 = \begin{bmatrix} 0.4146 \\ 1.8150 \end{bmatrix}, L_3 = \begin{bmatrix} 1.1435 \\ 0.4889 \end{bmatrix},$$
$$L_4 = \begin{bmatrix} 1.1435 \\ 1.1150 \end{bmatrix}$$

Invariance kernel, viability kernel and capture basin for residual signals (27) and (28) in healthy mode are shown in Figures 5 and 6, respectively. These figures are only for comparing the size of the kernels with one-graded axes. Both residuals in steady state must lie in $[-0.1, 0.1]$, otherwise fault detection scheme raises a fault alarm. For transient state, if no time constraint is considered, then viability kernel must be used for fault detection. We consider that the residuals start in a box given by $[-30, 30]$ per each component. In this application, it is reasonable, because there is always a way to bring the system back to steady state as long as the system constraints are satisfied. For the first residual $r_1(t)$, if we consider a time limitation of four sample times and start from invariance kernel as initial set, the interval $[-1.7, 1.7]$ is capture basin and can be used for fault detection. It means that if absolute value of residual becomes more than 1.7, there is no possibility to come back to steady state in less than four sample times. For the second residual $r_2(t)$ with four sample times constraint, the interval $[-1, 1]$ can be used for fault detection.

For FDI, two different cases are considered: full fault and partial fault.

6.1. Full faults

In this case, three fault scenarios are considered

$$\begin{aligned} &\textit{Fault scenario1}: && \mathrm{F}_{y1} = 0, && 100 \leqslant t^{(\mathrm{sec})} \leqslant 250 \\ &\textit{Fault scenario2}: && \mathrm{F}_{y2} = 0, && 400 \leqslant t^{(\mathrm{sec})} \leqslant 550 \\ &\textit{Fault scenario3}: && \mathrm{F}_{u} = 0, && 700 \leqslant t^{(\mathrm{sec})} \leqslant 850 \end{aligned}$$

In Figure 7, simulation results are presented. The residual signals $r_1(t)$ and $r_2(t)$ are depicted in Figures 8 and 9, respectively. It is clear that residual in healthy mode is inside the invariance kernel. Fault scenarios 1 and 3 make $r_1(t)$ going outside invariance kernel, which means a fault occurs. Only the second fault scenario that is related to the second output does not change this residual. As it can be seen from Figure 8, the third fault also makes the residual going outside the capture basin, which means there is no possibility to come back to invariance kernel in less that four sample times. In all situations, when fault disappears, it takes more than 10 sample times to come back to invariance kernel. But, for the first fault scenario, we cannot say that there is a fault in transient mode, as long as it lies in capture basin. During the whole simulation time, residual lies in viability kernel, which means system is safe and has the opportunity to come back to invariance kernel. The same argument can be done about $r_2(t)$. Only first fault scenario does not change it. Here also, third fault not only makes the residual going outside the invariance kernel, but also makes it to go out of capture basin. Therefore, a fault alarm can raise in transient mode. It must be noted that like first residual, $r_2(t)$ is also in the viability kernel.

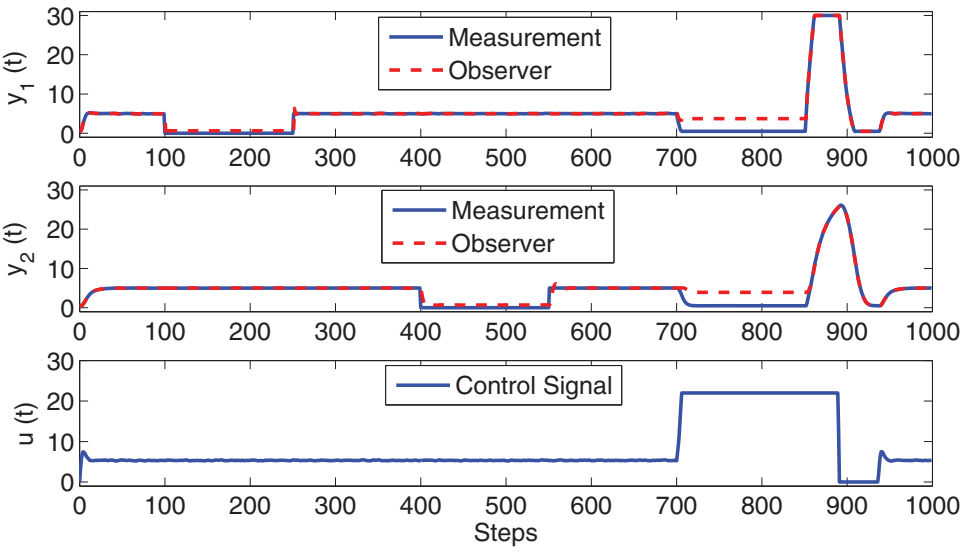


Figure 7. Simulation results.

It is also worth to note that using both residuals, three defined faults can be isolated. If $r_1(t)$ plotted against $r_2(t)$, as in Figure 10, invariance kernel for the applied faults can be used for fault isolation. Because invariance kernel in healthy and faulty modes are separable according to Section 4.4, the faults can be isolated. In Figure 10, it can be seen that when fault occurs, residual signal goes from healthy invariance kernel to one of the faulty invariance kernels. Based on this analysis, it can be said that first fault is detected in 102 seconds (2 seconds after its occurrence), second fault is detected in 410 seconds (10 seconds after its occurrence) and third fault is detected in 718 seconds (18 seconds after its occurrence).

6.2. *Partial faults*

In most practical situations, the size of the fault is not known a priori. Therefore, it is useful to analyse cases when faults can change in a range. Here, these three predefined faults are considered to be in a range; i.e.

$$F_{y1} \in [0,\ 0.5]$$
$$F_{y2} \in [0,\ 0.5]$$
$$F_{u} \in [0,\ 0.5]$$

An advantage of using viability approach is that in this case also FDI is possible. Invariance kernels in these three

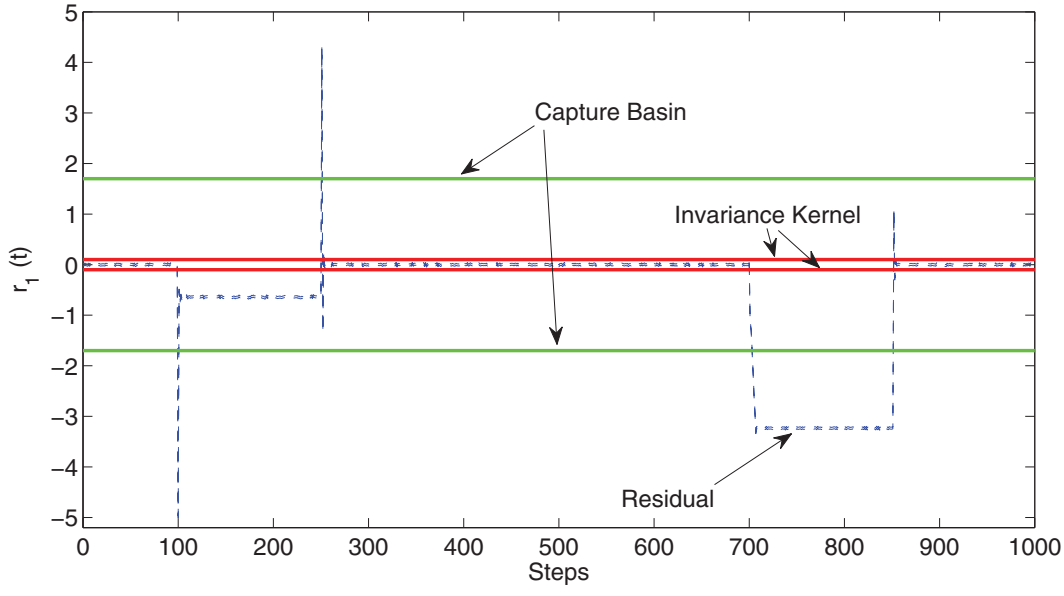


Figure 8. Invariance kernel and capture basin for first residual signal.

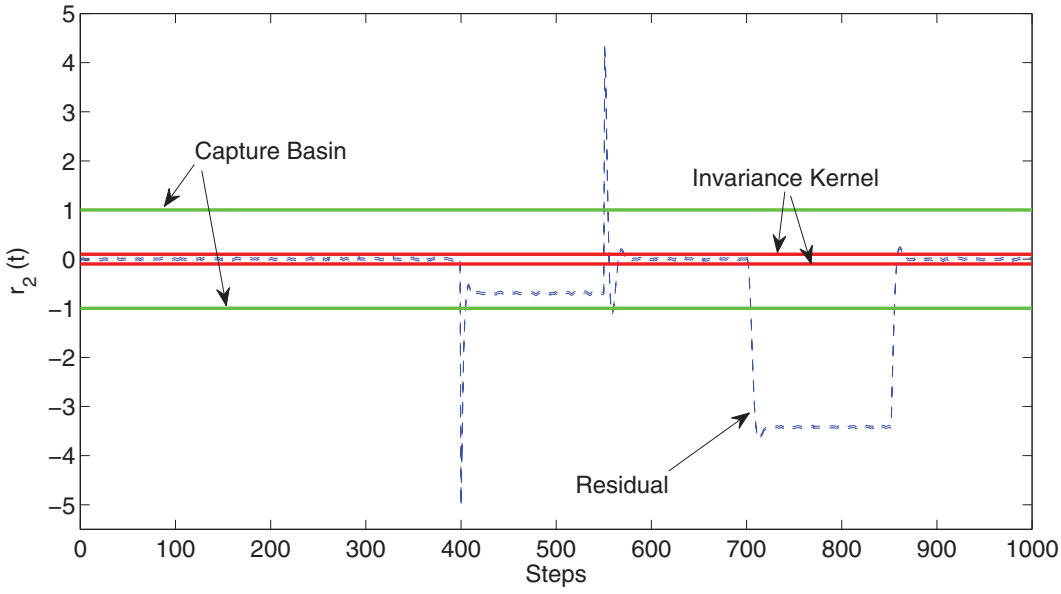


Figure 9. Invariance kernel and capture basin for second residual signal.

fault scenarios and healthy mode of operation of the system are depicted in Figure 11. As it can be seen from Figure 11, kernels are separable. Therefore, according to (4.4), considered faults can be isolated. The evolution of the system is also depicted in Figure 11 according to the following fault scenarios:

$$\textit{Fault scenario1}: \quad \mathrm{F}_{y1} = 0.25, \quad 100 \leqslant t^{(\mathrm{sec})} \leqslant 250$$
$$\textit{Fault scenario2}: \quad \mathrm{F}_{y2} = 0.25, \quad 400 \leqslant t^{(\mathrm{sec})} \leqslant 550$$
$$\textit{Fault scenario3}: \quad \mathrm{F}_{u} = 0.25, \quad 700 \leqslant t^{(\mathrm{sec})} \leqslant 850$$

In this case, first fault scenario is detected after 2 seconds, second fault scenario after 8 seconds is detected and third fault scenario is isolated in 6 seconds after its occurrence. It is clear that faults in this case is detected faster than the case of full fault, because the size of kernels are bigger than when a specific value for fault is considered.

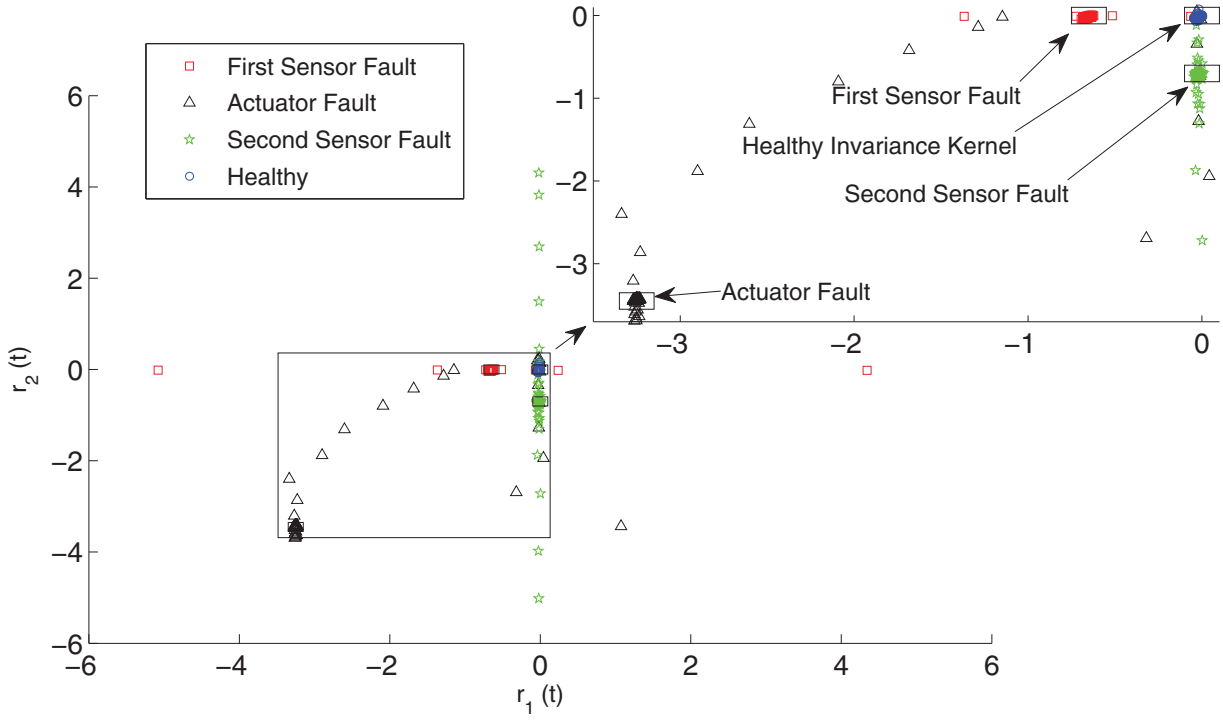


Figure 10. Invariance kernels for fault isolation.

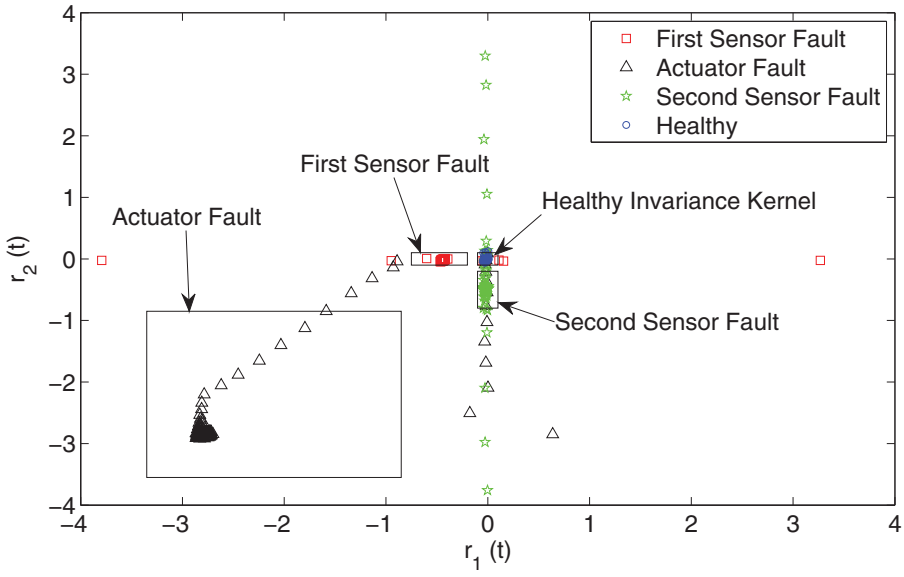


Figure 11. Invariance kernels for different interval faults.

7. Conclusions

In this paper, the application of viability theory to FDI has been developed regarding its application to non-linear systems that can be expressed in LPV form. Concepts of invariance kernel, viability kernel and capture basin are introduced and adapted to be used in FDI. The main drawback of viability theory is the difficulty of computing these kernels. In this paper, this drawback has been overcome with the use of zonotopes. Then, FDI in steady state and transient mode can be done using those proposed viability theory concepts. Moreover, the integration of the proposed FDI approach with interval observers has been presented allowing to work in transient mode and considering noise and uncertainty in residual evaluation phase and achieving robustness in a passive way. Finally, a well-known application example has been used in order to show effectiveness of the proposed approach.

As future work, the proposed approach will be extended to deal with fault tolerant control.

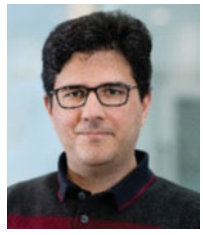
Acknowledgement

This work was partially supported by the Spanish State Research Agency (AEI) and the European Regional Development Fund (ERFD) through the projects ECOCIS (ref. DPI2013-48243-C2-1-R), DEOCS (ref. DPI2016-76493-C3-3-R) and HARCRICS (ref. DPI2014-58104-R).

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors



Majid Ghaniee Zarch received his B.Sc. and M.Sc. degrees in control engineering from Ferdowsi University and K.N. Toosi University of Technology, respectively. He is currently working towards a Ph.D. at Iran University of Science and Technology, Iran. His research interests include fault diagnosis, fault tolerant control and intelligent systems.

Vicenç Puig received his B.Sc./M.Sc. Degree in Telecommunications Engineering in 1993 and Ph.D. degree in Automatic Control, Vision and Robotics in 1999, both from the Technical University of Catalonia (UPC). He is a full professor at the Automatic Control Department of the UPC and a researcher at the Institute of Robotics and Industrial Informatics (IRI), CSIC-UPC. He is the chair of the Automatic Control Department and the head of the research group on Advanced Control Systems (SAC) at the UPC. He has developed important scientific contributions in the areas of fault diagnosis and fault tolerant control, using interval and linear-parameter-varying models using set-based approaches. He has participated in more than 20 European and national research projects in the last decade. He has also led many private contracts with several companies and has published more than 140 journal articles as well as over 400 contributions in international conference/workshop proceedings. He has supervised over 20 Ph.D. dissertations and over 50 Master's theses/final projects. He is currently the vice-chair of the IFAC Safeprocess TC Committee 6.4 (2014–2018). He was the general chair of the 3rd IEEE Conference on Control and Fault-Tolerant Systems (SysTol'2016) and is the IPC chair of IFAC Safeprocess 2018.

Javad Poshtan received his B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Tehran University, Tehran, Iran, in 1987, Sharif University of Technology, Tehran, Iran, in 1991, and the University of New Brunswick, Canada, in 1997, respectively. Since 1997, he has been with the Department of Electrical Engineering at Iran University of Science and Technology. He is involved in academic and research activities in areas such as control systems theory, system identification, and estimation theory.

Mahdi Aliyari Shoorehdeli received his B.Sc. degree in electronics engineering, his M. Sc. degree and Ph.D. degree in control engineering from K. N. Toosi University of Technology, in 2001, 2003 and 2008, respectively. He is currently an Assistant Professor with the Department of Mechatronics Engineering, K. N. Toosi University of Technology, Tehran. He is the author of more than 150 papers in international journals and conference proceedings. His research interests include fault detection and diagnosis, intelligent control of mechatronics systems and multi objective optimization.

References

Alamo, T., Bravo, J. M., & Camacho, E. F. (2005). Guaranteed state estimation by zonotopes. *Automatica, 41*(6), 1035–1043.

Aubin, J.-P. (2013). *Dynamic economic theory: A viability approach*. Berlin: Springer.

Aubin, J.-P., Bayen, A. M., & Saint-Pierre, P. (2011). *Viability theory: New directions*. Berlin: Springer.

Bergen, G. V. D. (1999). A fast and robust GJK implementation for collision detection of convex objects. *Journal of Graphics Tools, 4*(2), 7–25.

Blanchini, F., & Miani, S. (2008). *Set-theoretic methods in control*. Basel: Birkhäuser.

Blanke, M., Kinnaert, M., Lunze, J., & Staroswiecki, M. (2006). *Diagnosis and fault-tolerant control*. Berlin: Springer.

Blesa, J., Puig, V., Saludes, J., & Fernández-Cantí, R. M. (2016). Set-membership parity space approach for fault detection in linear uncertain dynamic systems. *International Journal of Adaptive Control and Signal Processing, 30*(2), 186–205.

Brito, P. L. G. (2009). *Robust fault detection and tolerance evaluation using zonotopes* (Doctoral dissertation). Polytechnic University of Catalonia, Barcelona.

Chen, J., and Patton, R. J. (2012). *Robust model-based fault diagnosis for dynamic systems*. New York, NY: Springer.

Combastel, C. (2003, September). A state bounding observer based on zonotopes. In *Proceedings of the European control conference* (pp. 2589–2594). Cambridge, UK: IEEE.

Dang, T. (2006, March). Approximate reachability computation for polynomial systems. In *Proceedings of the International workshop on hybrid systems: Computation and control* (pp. 138–152). Santa Barbara, Ca, USA: Springer.

Deffuant, G., & Gilbert, N. (2011). *Viability and resilience of complex systems: Concepts, methods and case studies from ecology and society*. Berlin: Springer.

Fernández-Cantí, R. M., Blesa, J., Puig, V., & Tornil-Sin, S. (2016). Set-membership identification and fault detection using a Bayesian framework. *International Journal of Systems Science, 47*(7), 1710–1724.

Gao, Z., Cecati, C., & Ding, S. X. (2015). A survey of fault diagnosis and fault-tolerant techniques – part I: Fault diagnosis with model-based and signal-based approaches. *IEEE Transactions on Industrial Electronics, 62*(6), 3757–3767.

Gáspár, P., Szabo, Z., & Bokor, J. (2007, July). A grey-box identification of an LPV vehicle model for observer-based side slip angle estimation. In *Proceedings of the American control conference* (pp. 2961–2966). New York, NY, USA: IEEE.

Ghaniee Zarch, M., & Aliyari Shoorehdeli, M. (n.d.). Generalization of parity space to fault detection based on Takagi-Sugeno fuzzy models for non-linear dynamic systems. *Expert Systems*, 35(1). Retrieved from http://dx.doi.org/10.1111/exsy.12228

Hanafi, A. N., Seron, M. M., & De Dona, J. A. (2015). Set invariance approach for fault detection and isolation in lure systems by LPV-embedding. *IFAC-PapersOnLine, 48*(21), 1036–1041.

Isermann, R. (2006). *Fault diagnosis systems: An introduction from fault detection to fault tolerance*. New York, NY: Springer.

Iurinic, L. U., Herrera-Orozco, A. R., Ferraz, R. G., & Bretas, A. S. (2016). Distribution systems high-impedance fault location: A parameter estimation approach. *IEEE Transactions on Power Delivery, 31*(4), 1806–1814.

Lalami, A., & Combastel, C. (2006). Generation of set membership tests for fault diagnosis and evaluation of their worst case sensitivity. *IFAC Proceedings Volumes, 39*(13), 569–574.

Le, V. T. H., Stoica, C., Alamo, T., Camacho, E. F., & Dumur, D. (2013). *Zonotopes: From guaranteed state-estimation to control*. Hoboken, NJ, USA: Wiley.

Li, X.-J., & Yang, G.-H. (2014). Fault detection in finite frequency domain for Takagi-Sugeno fuzzy systems with sensor faults. *IEEE Transactions on Cybernetics, 44*(8), 1446–1458.

Linhart, J. (1989). Approximation of a ball by zonotopes using uniform distribution on the sphere. *Archiv der Mathematik, 53*(1), 82–86.

Maidens, J. N., Kaynama, S., Mitchell, I. M., Oishi, M. M., & Dumont, G. A. (2013). Lagrangian methods for approximating the viability kernel in high-dimensional systems. *Automatica, 49*(7), 2017–2029.

Mazenc, F., & Bernard, O. (2011). Interval observers for linear time-invariant systems with disturbances. *Automatica, 47*(1), 140–147.

Meseguer, J., Puig, V., & Escobet, T. (2017). Approximating fault detection linear interval observers using λ-order interval predictors. *International Journal of Adaptive Control and Signal Processing, 31*(7), 1040–1060.

Mitchell, I. M. (2007, April). Comparing forward and backward reachability as tools for safety analysis. In *Proceedings of the International workshop on hybrid systems: Computation and control* (pp. 428–443). Pisa, Italy: Springer.

Mohajerpoor, R., Abdi, H., & Nahavandi, S. (2015). Reduced-order functional observers with application to partial state estimation of linear systems with input-delays. *Journal of Control and Decision, 2*(4), 233–256.

Mondal, S. (2017). Robust adaptive observer for non-linear time-delay systems with disturbances and uncertainties. *Journal of Control and Decision, 4*(2), 100–113.

Montes de Oca, S., Puig, V., & Blesa, J. (2012). Robust fault detection based on adaptive threshold generation using interval LPV observers. *International Journal of Adaptive Control and Signal Processing, 26*(3), 258–283.

Olaru, S., De Doná, J. A., Seron, M., & Stoican, F. (2010). Positive invariant sets for fault tolerant multisensor control schemes. *International Journal of Control, 83*(12), 2622–2640.

Pourasghar, M., Puig, V., & Ocampo-Martinez, C. (2016, September). Characterization of the minimum detectable fault of interval observers by using set-invariance theory. In *Proceedings of the 3rd conference on control and fault-tolerant systems* (pp. 79–86). Barcelona, Spain: IEEE.

Puig, V. (2010). Fault diagnosis and fault tolerant Control using set-membership approaches: Application to real case studies. *International Journal of Applied Mathematics and Computer Science, 20*, 619–635.

Puig, V., Stancu, A., Escobet, T., Nejjari, F., Quevedo, J., & Patton, R. J. (2006). Passive robust fault detection using interval observers: Application to the damadics benchmark problem. *Control Engineering Practice, 14*(6), 621–633.

Rugh, W., & Shamma, J. (2000). A survey of research on gain-scheduling. *Automatica, 36*(10), 1401–1425.

Saint-Pierre, P. (2004). Viable capture basin for studying differential and hybrid games: Application to finance. *International Game Theory Review, 6*(01), 109–136.

Seron, M. M., & De Doná, J. A. (2015). Robust fault estimation and compensation for LPV systems under actuator and sensor faults. *Automatica, 52*, 294–301.

Seron, M. M., De Doná, J. A., & Olaru, S. (2012). Fault tolerant control allowing sensor healthy-to-faulty and faulty-to-healthy transitions. *IEEE Transactions on Automatic Control, 57*(7), 1657–1669.

Seron, M. M., Zhuo, X. W., De Doná, J. A., & Martínez, J. J. (2008). Multisensor switching control strategy with fault tolerance guarantees. *Automatica, 44*(1), 88–97.

Spiteri, R. J., Pai, D. K., & Ascher, U. M. (2000). Programming and control of robots by means of differential algebraic inequalities. *IEEE Transactions on Robotics and Automation, 16*(2), 135–145.

Stoican, F. (2011). *Fault tolerant control based on set-theoretic methods* (Doctoral dissertation). Supélec, Paris.

Stoican, F., & Olaru, S. (2013). *Set-theoretic fault-tolerant control in multisensor systems*. Hoboken, NJ, USA: Wiley.

Tóth, R. (2010). *Modeling and identification of linear parameter-varying systems*. Berlin: Springer.

Wang, J. L., Yang, G.-H., & Liu, J. (2007). An LMI approach to H-index and mixed H-/H^{∞} fault detection observer design. *Automatica, 43*(9), 1656–1665.

Xu, F., Stoican, F., Puig, V., Ocampo-Martinez, C., & Olaru, S. (2013, October). On the relationship between interval observers and invariant sets in fault detection. In *Proceedings of the conference on control and fault-tolerant systems* (pp. 49–54, Nice, France: IEEE).

Zarch, M. G., Puig, V., & Poshtan, J. (2017). Fault detection and isolation using viability theory and interval observers. *Journal of Physics: Conference Series, 783*(1), 1–12.

Appendix

Some relevant properties of zonotopes used in the algorithms developed in the paper have been reviewed in this appendix.

Property A.1. (Zonotope inclusion; Alamo et al., 2005)**:** Consider a family of zonotopes represented by $X = \pi \oplus M\beta^{m_z}$, where $\pi \in \mathbb{R}^{n_z}$ is a real vector and $M \in \mathbb{R}^{n_z \times m_z}$ is an interval matrix. A zonotope inclusion $\diamond(X)$ is defined by

$$\diamond(X) = \pi \oplus \left[mid\,(M)\ G \right] \begin{bmatrix} \beta^{m_z} \\ \beta^{n_z} \end{bmatrix} = \pi \oplus J\beta^{n_z+m_z}$$

where $G \in \mathbb{R}^{n_z \times m_z}$ is a diagonal matrix that satisfies

$$G_{ii} = \sum_{j=1}^{m_z} \left(diam\left(M_{ij}\right)/_2 \right), \quad i = 1, 2, \ldots, n$$

where 'mid' denotes the centre and 'diam' the diameter of the interval according to Le et al. (2013). Under this definition $X \subseteq \diamond(X)$.

Property A.2. (Intersection; Brito, 2009)**:** Given two zonotopes $Z_1 = p_1 \oplus H_1\beta^{r_1}$ and $Z_2 = p_2 \oplus H_2\beta^{r_2}$ and matrix E, let us define

$$\begin{aligned} \hat{p}\,(E) &= Ep_1 + (I - E)\,p_2 \\ \hat{H}\,(E) &= \left[EH_1\ \ (I - E)\,H_2 \right] \end{aligned}$$

then,

$$\begin{aligned} Z_1 \cap Z_2 &\subseteq \hat{Z}\,(E) \\ \hat{Z}\,(E) &= \hat{p}\,(E) \oplus \hat{H}\,(E)\,\beta^{r_1+r_2} \end{aligned} \tag{A1}$$

Testing whether the intersection of two convex sets is empty or not can be done by collision detection algorithms. Some collision detection algorithms can thus be used to test whether a point belongs to a given set. The GJK (Gilbert–Johnson–Keerthi) algorithm is a robust and fast collision detection algorithm that is introduced in Bergen (1999). Alternatively, testing the emptiness of the intersection between two sets is equivalent to test the membership of the origin in the Minkowski difference of the two sets (Lalami & Combastel, 2006).

As previously mentioned, detecting the collision between Z_1 and Z_2 can be reformulated as testing the inclusion of origin in the Minkowski difference between Z_1 and Z_2:

$$0 \in Z_d = Z_1 \oplus (-Z_2) = c_d \oplus H_d \beta^{m_d}$$

The proposed solution is to find a separation vector ω whose direction aims at proving that the 0 is not included in the zonotope. Indeed,

$$\begin{aligned} 0 \notin Z_d \quad &\Leftrightarrow \quad \exists \omega, 0 \notin \omega^T Z_d \\ &\Leftrightarrow \quad \exists \omega, \left|\omega^T c_d\right| > \left\|\omega^T H_d\right\|_1 \end{aligned}$$

Therefore, the collision detection can be reformulated as the maximisation of the criterion J:

$$\begin{aligned} \omega^* &= \max_{\omega} J(\omega) \\ J(\omega) &= \frac{\left|\omega^T c_d\right|}{\left\|\omega^T H_d\right\|_1} \end{aligned} \tag{A2}$$

If $|\omega^{*T} c_d| > \|\omega^{*T} H_d\|_1$ then $0 \notin Z_d$. The problem is addressed solving Equation (A2) that could be addressed by an iterative algorithm. In Lalami and Combastel (2006), a sub-optimal solution based on the optimisation of a criterion involving the Euclidean norm instead of the 1-norm is proposed:

$$J_{\text{subopt}}(\omega) = \frac{\left\|\omega^T c_d\right\|_2^2}{\left\|\omega^T H_d\right\|_2^2}$$