WARSAW UNIVERSITY
OF TECHNOLOGY

Faculty of Mechatronics

Ph.D. THESIS

Kornel Rostek, M.Sc.

Generalized Metric of Fault Distinguishability for Diagnostics of Industrial Processes

Supervisor
Professor Michał Bartyś, Ph.D., D.Sc.

Warsaw, 2017

Streszczenie

Celem niniejszej rozprawy doktorskiej było opracowanie nowej, uniwersalnej i użytecznej miary rozróżnialności uszkodzeń. Opracowana miara znajduje zastosowanie między innymi do rozwiązania problemu optymalnego doboru sygnałów pomiarowych na potrzeby diagnostyki procesów. Miara może być stosowana do analizy różnych rodzajów systemów diagnostycznych, w tym wykorzystujących binarne, wielowartościowe i ciągłe sygnały diagnostyczne. Unikalną cechą opracowanej miary jest jej wrażliwość na słabą i silną rozróżnialność uszkodzeń.

W pracy została zaproponowana metoda budowy systemu diagnostycznego o wielowartościowych sygnałach diagnostycznych o właściwościach odpowiadających systemowi wykorzystującemu sekwencje symptomów. Dzięki temu zaproponowana miara może być także zastosowana do analizy właściwości struktur diagnostycznych wykorzystujących informacje o kolejności (sekwencji) symptomów.

W pracy przedstawiono także metodykę konstrukcji liniowego problemu optymalizacji doboru sygnałów pomiarowych z zastosowaniem wprowadzonej miary rozróżnialności. Metodyka ta pozwala na zdefiniowanie liniowej funkcji celu i liniowych ograniczeń, co ma istotne znaczenie w zastosowaniach praktycznych. W pracy zostały przedstawione przykłady praktycznego zastosowania zaproponowanej miary rozróżnialności do:

- wyznaczenia maksymalnej wartości miary rozróżnialności uszkodzeń dla inteligentnego elektro-pneumatycznego elementu wykonawczego z wbudowanymi funkcjami diagnostycznymi,
- doboru sygnałów pomiarowych i czujników pomiarowych do ogniwa paliwowego w warunkach ograniczeń budżetowych,
- sformułowanie i rozwiązanie problemu optymalnego doboru sygnałów pomiarowych dla laboratoryjnego zestawu trzech zbiorników. Uzyskane wyniki zostały porównane z rezultatami osiąganymi innymi metodami.

W podsumowaniu zostały wskazane i przeanalizowane dalsze kierunki badań.

Słowa kluczowe: rozróżnialność uszkodzeń, miary rozróżnialności uszkodzeń, optymalny dobór sygnałów pomiarowych.

Abstract

The main aim of this thesis was to introduce a novel, generalized and practical metric of single fault isolability. The proposed metric can be used to formulate an optimal sensor placement problem for diagnostic purposes. The metric applies to various approaches to fault isolability. It can be used with diagnostic structures using different diagnostic signals, e.g., binary, multi-valued, continuous. The metric takes into account effects of weak and unidirectional strong isolability.

A method of constructing multi-valued diagnostic signals that provide equivalent isolability properties as symptoms sequences were proposed. It can be used to analyze diagnostic structures based on the order of symptoms.

A method was proposed for constructing linear optimal sensor placement problem using the proposed isolability metric. Using this method, the obtained objective function and constraints using the metric are linear which is important in industrial applications. The resulting optimization problem is then a standard Integer Linear Programming problem. The examples of practical applications of the proposed metric were presented:

- finding the maximal value of the metric of fault isolability for an intelligent electro-pneumatic actuator with embedded diagnostics and determine the optimal binary diagnostic structure,
- optimal sensor placement for a Fuel Cell Stack System with budgetary constraints,
- formulating and solving the optimal sensor placement problem for laboratory station for the Three Tank System.

In the summary, further work was discussed and analyzed.

Keywords: fault isolation, metrics of fault distinguishability, optimal sensor placement.

Contents

Nomenclature

$\boldsymbol{\epsilon}$	Vector of noise in the system
$\boldsymbol{\phi}$	Alternative signature of a fault
$\boldsymbol{\upsilon}$	Direction in parity space
$\boldsymbol{FSig}$	Functional fault signature
$\boldsymbol{G}$	Matrix of input–output transfer functions
$\boldsymbol{H}$	Matrix of faults–output transfer functions
$\boldsymbol{p}$	Vector of process parameters
$\boldsymbol{u}$	Vector of process inputs
$\boldsymbol{y}$	Vector of process outputs
Φ	Set of all alternative signatures
ψ	Proposed measure of isolability
ARR	Analytical Redundancy Relation
$COMPS$	Set of components
D	Set of all elementary diagnoses
d	Diagnosis
D_c	Number of disjoint subsets of unisolable faults (D-classes)
es	Elementary sequence of two symptoms
es^d	Symptom sequence with assigned delay
F	Set of faults
f	Fault
K	Number of faults
OBS	Set of observations

S	Set of diagnostic signals
s	Diagnostic signal
SD	System description
SM	System model
V	Set of values of a diagnostic signal
v	Element of Binary Diagnostic Matrix
x	Decision variable

Chapter 1

Introduction

1.1 Thesis and motivation

The main aim of this thesis was to introduce a novel, generalized metric of single fault isolability. This metric has both theoretical and usability purposes. From a theoretical point of view, it can be used as a synthetic metric of fault isolability applicable to majority of methods of fault isolation developed in recent years. From a usability point of view, it can also be used to formulate an optimal sensor placement as a linear optimization problem.

Since many years there has been observed a strong, rising demand in the Fault Detection and Isolation community for a development of a universal fault distinguishability metric that should make it possible to formulate an optimal sensor placement problem. This metric should be useful for applications in the wide range of diagnostic structural approaches, including those that make use of multivalued discrete and continuous diagnostic signals. Additionally, this metric should take into account effects of weak and unidirectional strong isolability. This thesis responds to this demand.

Various metrics of fault isolability are known. However, there is no general metric. Current definitions are insufficient because their applicability is limited. They are restricted to some types of diagnostic signals, or they are difficult to calculate for complex diagnostic structures.

In order to propose a new of isolability metric, it was necessary to evaluate all known types of fault isolability metrics in terms of their:

- applicability to diagnostic structures based on multi-valued and continuous diagnostic signals and on sequences of symptoms,
- ability to expose strength of isolability,
- applicability for constructing and solving of an optimal sensor placement problem.

1.2 Model-based fault detection and isolation

Different types of diagnostic methods are considered in the literature. In general, methods of fault detection and isolation can be classified into three categories (Papadopoulos and McDermid 2001; Roth 2010):

- data-driven, e.g. (Rostek *et al.* 2015; Yin *et al.* 2015; Yin *et al.* 2016),
- model-based, e.g. (Blanke *et al.* 2006; Górny and Ligęza 2002; Patton *et al.* 2000a),
- expert knowledge based, e.g. (Kościelny 1999; Syfert and Kościelny 2009).

Fault detection approaches based on analysis of statistical properties of signals are also studied. Examples can be found in (Fillatre and Nikiforov 2007; Flouladirad and Nikiforov 2003).

In this thesis, we will focus on model-based approaches. Usually, three families of model-based approaches are considered (Chen and Patton 2012; Travé-Massuyès 2014b):

- parameter estimation focused on parameters representing physical features of the diagnosed process (Pouliezos *et al.* 1989; Zhou *et al.* 2015),
- state estimation, where internal state variables are estimated by state observers (Frank 1994; Lan and Patton 2016),
- parity space, where elimination of unknown variables is used (Chow and Willsky 1984; Wang *et al.* 2015).

Model-based approaches assume that a fault occurrence causes a discrepancy between the observed behavior of the system and the behavior of the model of this system. The measure of this discrepancy is a called residual. In order to calculate residuals in the online mode, only known variables are used, i.e., model parameters, measurements, control signals and process outputs. An equation that uses exclusively known variables is called the computational form of a residual. In contrast, a residual equation in the internal form incorporates faults and their influence on the process. In practice, it is often difficult to obtain the internal form of a residual because it requires the exact mathematical model of a part of the diagnosed process.

Two main scientific communities are developing model-based diagnosis—the fault detection and isolation (FDI) community, which originally comes from the automatic control field and the diagnosis (DX) community, which has emerged from the computer science and artificial intelligence fields (Travé-Massuyès 2014b).

1.3 Fault detection and isolation

In FDI methodology, three main stages of fault diagnosis are considered (Isermann and Balle 1997):

1. fault detection, where residuals are used to determine the presence of a fault,
2. fault isolation, where the kind and location of the detected fault is determined,
3. fault identification, where the size and evolution of the fault in time are estimated.

Most of the methods and approaches presented in this thesis are focused only on the first two stages.

In FDI, residuals are frequently generated by means of Analytical Redundancy Relations (ARRs) (Chow and Willsky 1984; Sanchez *et al.* 2015; Staroswiecki and Comtet-Varga 2001). ARRs are relations without unknown state variables. In practice, we frequently obtain residuals from balance equations or by comparison of outputs of models with their corresponding measured process outputs. Some models include information about how faults influence the process. Other models are limited only to the fault-free mode.

Fault isolability in a diagnostic system is closely related to available diagnostic signals, i.e., outputs of detection algorithms (Korbicz *et al.* 2004; Kościelny 2001) and the form of notation used for describing the relation between diagnostic signal values and faults. Three types of values of diagnostic signals may be distinguished: binary, multi-valued (e.g., three-valued: $-1, 0, +1$)

and continuous. There is also an approach which uses fuzzy diagnostic signals (Bartyś 2013; Kościelny and Syfert 2006; Kościelny *et al.* 1999). It is beyond the scope of this thesis, however.

To formulate a diagnosis, knowledge of the relationship between faults and diagnostic signal values is necessary. A symptom is the appearance of any value of a diagnostic signal that indicates the presence of a fault. It is usually assumed that value 0 corresponds to the faultless state and other values indicate faults. For a given fault, the vector of characteristic values of diagnostic signals is called a signature. The relation between values of diagnostic signals and faults will be referred to as the fault–symptoms relation. This relationship takes various forms depending on the available diagnostic signals. In the case of binary diagnostic signals, the incidence matrix (binary diagnostic matrix—BDM) is primarily used (Chen and Patton 1999; Gertler 1998; Isermann 2006; Korbicz *et al.* 2004; Patton *et al.* 2000b). Another approach is the Fault Information System (FIS) (Kościelny 1999; Kościelny *et al.* 2006), which assumes the usage of multi-valued diagnostic signals. Other forms of notation can be derived from BDM (Korbicz *et al.* 2004), such as logic functions, IF–THEN rules, and fault trees. Similar rules can be derived from FIS. Those derivative forms of notation are alternative representations and do not influence fault isolability because they do not provide any additional information about faults (Korbicz *et al.* 2004). The relation between faults and continuous diagnostic signals is described by regions in residual space (Isermann 2006; Korbicz *et al.* 2004), vectors of directions in residual space (Chen and Patton 1999; Gertler 1998; Patton *et al.* 2000b) and sequences of symptoms (Kościelny *et al.* 2013). Fuzzy reasoning about the fault–symptoms relation was also studied, e.g. (Kościelny *et al.* 2008; Patton *et al.* 2000a).

The above relations are determined based on:

1. process modeling built on derived residual equations in the internal form (Gertler 1998),
2. learning, i.e., identifying regions in the diagnostic signal space that correspond to individual faults, processing data captured during process states with faults (Koivo 1994; Patton *et al.* 1999),
3. expert knowledge (Kościelny 1999; Syfert and Kościelny 2009).

The model of a diagnosed process, making use of the relation between inputs, outputs, and faults (and possibly disturbances and measurement noise) is available only in the first case. Also, the majority of publications in the field of fault isolation relate only to the first case. Isolability definitions given in survey studies (Basseville 1997; Basseville 2001; Ding 2008) also refer to the first case. They assume knowledge of the process model incorporating information about faults.

1.3.1 Fault isolability

Faults are isolable if it is possible to explicitly determine which faults have occurred. Fault isolability has usually been defined in the context of the adopted diagnostic method. Most often, fault isolability was studied in the case of binary diagnostic signals (BDM, incidence matrix, structure matrix) derived from the structure of linear equations of residuals in the internal form. The following definitions (1.3.1 – 1.3.6) were given by Gertler in (Gertler 1998):

Definition 1.3.1. *The structure matrix of a residual set expresses the cause–effect relationship between faults and disturbances as inputs and residuals as outputs. A "1" in the intersection means that the fault/disturbance does affect the residual while "0" means it does not.*

Definition 1.3.2. *A fault or disturbance is undetectable in a residual structure if its column in the structure matrix contains only "0" elements.*

Definition 1.3.3. *Two faults or disturbances are indistinguishable if their respective columns in the structure matrix are identical.*

Definition 1.3.4. *A structure is weakly isolating if all columns in the structure matrix are different and nonzero.*

Definition 1.3.5. *A structure is unidirectionally strongly isolating if it is weakly isolating and if no column in the structure matrix can be obtained from any other column by turning an arbitrary number of "1"s into "0"s or by turning an arbitrary number of "0"s into "1"s.*

Definition 1.3.6. *A structure is bidirectionally strongly isolating of degree 1 if it is weakly isolating and if no column can be obtained from another column by changing any single element. Similarly, a structure is bidirectionally strongly isolating of degree k if no column can be obtained from any other column by changing up to k elements.*

This definition means that in a bidirectionally strongly isolating structure each pair of columns differs in at least $k + 1$ positions.

In a unidirectionally strongly isolating structure, each pair of faults differs in at least two entries. Firstly, where "1" is in the first column and "0" in the other one in the same row and secondly, where "0" is in the first column and "1" in the other one.

In a weakly isolating structure, all faults are detectable (not undetectable) and mutually distinguishable (not indistinguishable).

Example 1.3.1.

Following Gertler's example (Gertler 1998), consider the following structures of residual sets:

$$M_1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \quad M_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \tag{1.1}$$

M_1 and M_2 are unidirectionally strongly isolating because any column cannot be turned into another column by changing "0"s into "1"s or "1"s into "0"s. They are also bidirectionally strongly isolating of degree 1 since changing only one element cannot turn any column into another one.

M_3 is bidirectionally strongly isolating of degree 1, but it is not unidirectionally strongly isolating because the second and third columns can be obtained from the first column by turning "1"s into "0"s.

Faults isolability, obtained on the basis of the structure of residuals derived from expert knowledge, was studied in (Korbicz *et al.* 2004; Kościelny *et al.* 2006). The definitions for BDM are equivalent to those given above.

In the case of multi-valued and continuous diagnostic signals, it is not always possible to determine subsets of unisolable faults. Two faults can be isolable for some particular values of diagnostic signals and not isolable for other values. Therefore, a definition of conditional isolability was introduced in (Korbicz *et al.* 2004; Kościelny *et al.* 2006).

1.4 Consistency-based diagnosis

There is a parallel branch of fault diagnosis developed by the computer science and artificial intelligence community called DX. The DX approach relies upon methods derived from Reiter's consistency-based approaches (Reiter 1987) and later extended in (De Kleer and Williams 1987; De Kleer *et al.* 1992). These methods use a logic-based diagnosis process to explain the differences between the observed and the correct behavior of the system.

De Kleer and Williams proposed in (De Kleer and Williams 1987) the general diagnostic engine (GDE). In GDE, the diagnostic process is an iterative loop of behavior prediction, conflict detection, candidate generation and refinement. Conflicts are sets of components that cannot be all fault-free because assumptions about their proper behavior are inconsistent with current observations and system description. Diagnosis candidates are hitting-sets of the conflict sets. Usually, the set of diagnostic candidates is characterized by the set of minimal diagnosis candidates. It should be noted that diagnosis is defined as a set of faulty components and this implies that multiple faults are also considered.

1.4.1 Comparison of DX and FDI methods

In FDI, the relations between diagnostic signals and faults are calculated offline, during the design phase of a diagnostic system. When determining the diagnosis, only values of residuals are computed and evaluated in the online mode. This approach is very time efficient, especially for time–critical processes. In Reiter's original approach, not only are current observations calculated, but the whole logical inference process is performed in the online mode. Such an approach requires much more computational power and time. Therefore, it may be unacceptable in the case of complex industrial systems. Moreover, the diagnosis of dynamic systems requires additional extensions to the original approach (Lamperti and Zanella 2013). In the study (Pulido and González 2004), the concept of Possible Conflicts (PC) was proposed. The conflicts can be calculated in the offline mode. A similar approach was introduced in (Górny and Ligęza 2002) as the Potential Conflict Structure (PCS), which is a local subgraph sufficient for determination of a potential conflict.

In DX approaches, multiple faults are handled in a natural way. Unfortunately, this may lead to a combinational explosion of the number of diagnoses. In FDI approaches, the number of faults is typically limited to single or double faults. There are some works that try to solve the multiple faults problem by means of FDI methods (Bartyś 2014; Bartyś 2015; Bartyś 2016b).

In recent years, substantial effort has been made to combine FDI and DX methodologies (Biswas *et al.* 2004). In (Cordier *et al.* 2004) the concept of conflict was compared to Analytical Redundancy Relations (ARR), which is the concept underlying FDI. It was proven that under specific conditions both approaches yield identical results. A comprehensive survey of this area of research can be found in (Travé-Massuyès 2014b).

ARR completeness

In (Cordier *et al.* 2004), bridging techniques are based on a concept of Analytical Redundancy Relation (ARR), which is a consistency relation used in FDI. The set of all ARRs leads to the signature matrix (FS), which is equivalent to BDM used in this thesis. The definition of ARR support as well as two of its properties, ARR-d-completeness and ARR-i-completeness, were proposed to establish a correspondence between the DX and FDI approaches (Cordier *et al.* 2004).

Definition 1.4.1. *The ARR support ARR_i is the set of components (columns of a signature matrix, faults) with a nonzero element in the row corresponding to this ARR_i.*

Definition 1.4.2. *ARR-d-completeness property: The set E of ARRs is said to be d-complete if:*

- *E is finite;*
- *for any set of observations OBS, if $SM \cup OBS \models \perp$, then $\exists ARR_i \in E$ such that $\{ARR_i\} \cup OBS \models \perp$.*

Definition 1.4.3. *ARR-i-completeness property: A set E of ARRs is said to be i-complete if:*

- *E is finite;*
- *for any subset C of the set of components $COMPS$, $C \subseteq COMPS$ and for any OBS if $SM(C) \cup OBS \models \perp$, then $\exists ARR_i \in E$ such that the support of ARR_i is included in C and $\{ARR_i\} \cup OBS \models \perp$.*

SM is the system model in the FDI approach. The restriction of the system model to the subset of components $C \subseteq COMPS$ is denoted by $SM(C)$.

ARR-d-completeness is related to fault detectability. ARR-i-completeness means that all multiple faults are isolable. In the paper (Cordier *et al.* 2004), the System Representation Equivalence (SRE) property was presented. SRE defines conditions for which the FDI (SM) and DX (SD) models representing the system are equivalent. It was also proven that when the SRE property is met, the following statements are true:

1. Given an ARR, ARR_i violated by OBS, the support of ARR_i is a Reiter's conflict.
2. If E is a d-complete set of ARRs, then if there exists a Reiter's conflict for (SD, $COMPS$, OBS), there exists an ARR $\text{ARR}_i \in E$ violated by OBS.
3. If E is i-complete, then given a Reiter's conflict C for $(SD, COMPS, OBS)$, there exists an ARR $\text{ARR}_i \in E$ violated by OBS, whose support is included in C.

Exoneration assumption

In FDI, an exoneration assumption is commonly accepted:

Definition 1.4.4. *ARR-exoneration (Travé-Massuyès 2014a): given OBS, any component (column of the signature matrix) in the support of an ARR satisfied by OBS is exonerated, i.e., considered normal.*

Therefore, to point out a fault in diagnosis, all of its symptoms must come into existence. This assumption is not always satisfied. Due to the dynamics of symptoms, the symptoms may not appear simultaneously or may not even appear at all.

There is an analogous definition in the DX approach:

Definition 1.4.5. *Component-exoneration: given OBS and $c \in COMP$, if $SM(c) \cup OBS$ is consistent, then c is exonerated, i.e., considered normal.*

The DX approaches are usually applied without any exoneration assumptions.

It was shown (Travé-Massuyès 2014a) that without exoneration assumptions and when the SRE property is met, diagnoses obtained with DX and FDI are equivalent.

1.5 Other definitions

Other important definitions are related to structural system description. Structural detectability and isolability are widely used in diagnosability analysis (Düştegör *et al.* 2006; Frisk *et al.* 2012) and sensor placement algorithms (Krysander and Frisk 2008). The main advantage of these definitions is their close relationship to dedicated computational schemes. On the other hand, they are limited to structural system description. Multi-valued diagnostic signals and symptom sequences cannot be handled.

A definition of isolability was also proposed in (Travé-Massuyès *et al.* 2006). It is related to the definitions that refer to residual space (Isermann 2006). It should be noted that in this approach, conditional isolability is not symmetric. This definition allows the consideration of multi-valued diagnostic signals but not symptom sequences.

The presented definitions are not complete as distinct faults may be characterized by a particular sequence of symptom occurrence.

Fault isolability is a major issue in fault diagnosis, but no definition exists which covers all of the different approaches.

1.6 Metrics of fault isolability

When designing a diagnostic system, it is essential to formulate all requirements precisely. This allows the designer to develop a diagnostic system with sufficient efficiency and required accuracy.

Fault isolability is one of the elementary attributes characterizing the quality of a diagnostic process. Fault isolability is also a feature reflecting the effectiveness of a fault isolation process. Due to this, it may be used for comparing the effectiveness of different methods of fault detection and isolation.

Basic definitions of isolability given in Section 1.3.1 are applicable only when comparing a pair of signatures of faults. Therefore, they are not sufficient to compare different diagnostic systems, especially when different methods of fault detection and isolation are used. The process of decision making during the design phase is therefore hindered.

There are many methods of determining the set of measurement signals necessary for diagnostic purposes. They usually consider the required minimum fault detectability and isolability of the diagnostic system (Krysander *et al.* 2008; Yassine *et al.* 2008). Some of the proposed methods also maximize the designed fault isolability, i.e. (Spanache *et al.* 2004). Regardless of the chosen method, simple, qualitative methods of analysis of fault isolability are insufficient. A generalized, quantitative method of fault isolability analysis is required.

There are quantitative methods of determining fault isolability for pairs of faults. In the simplest case, the Hamming distance between binary fault signatures can be used (Staroswiecki *et al.* 2000). It is calculated as the total number of differences between two fault signatures. Another, more advanced solution is the Isolability Ratio (Khorasgani *et al.* 2014). It is defined as the ratio of the effects of one fault to the sum of effects of another fault and uncertainties. Other examples are the distinguishability (Eriksson *et al.* 2013) and the expected distinguishability (Jung *et al.* 2015). Both use the Kullback-Leibler divergence. Unfortunately, those measures can only be used when analyzing a pair of faults. It is clear that to analyze a whole diagnostic system other methods are required.

1.6.1 Diagnosability degree

Diagnosability degree is one of the most commonly used measures of fault isolability (Krysander *et al.* 2008; Rouissi and Hoblos 2013; Spanache *et al.* 2004; Yassine *et al.* 2008). It was defined in (Travé-Massuyès *et al.* 2001). The value of this measure is calculated in two steps:

1. The set of all considered faults is divided into disjoint subsets of unisolable faults. Those sets are called D-classes.

2. The number of D-classes denoted as D_c is divided by the number of all considered faults. The obtained ratio is called diagnosability degree:

$$\frac{D_c}{card(F)}. \tag{1.2}$$

In the special case when all faults are isolable, the number of D-classes is equal to the number of all considered faults. The diagnosability degree is then equal to 1.

If all faults are unisolable, then the diagnosability degree is equal to $^1/_{card(F)}$.

Example 1.6.1.

Let us consider the example of a binary diagnostic matrix presented in Tab. 1.1.

Table 1.1: Example of a binary diagnostic matrix.

	f_1	f_2	f_3	f_4
s_1	1	1	1	
s_2		1	1	1
s_3				1

Using the definition of weak fault isolability for binary diagnostic matrix, the following three sets of unisolable faults can be distinguished: $\{f_1\}$, $\{f_2, f_3\}$, $\{f_4\}$. *The diagnosability degree is then equal to* $^3/_4$.

The diagnosability degree is a simple measure, easy to calculate and interpret. However, it has some disadvantages. Fault diagnosability for each pair of faults is defined only for binary diagnostic signals and does not apply to multi-valued diagnostic signals, for example the Fault Information System FIS (Kościelny and Zakroczymski 2001).

Moreover, it does not reflect the difference between weakly or strongly isolating structures. The problem was partially solved in the study (Kilic 2008), where the fuzzy diagnosability degree was proposed. However, this solution was designed only for Discrete Event Systems (DES) and cannot be easily extended to other types of diagnostic systems.

1.6.2 Diagnosis accuracy

In (Kościelny 2001), a simple fault isolability measure was defined. It is called the diagnosis accuracy. It is calculated as the reciprocal of the average number of faults in a diagnosis.

$$\left[\frac{\sum_{d_i \in D} card(d_i)}{card(D)}\right]^{-1}; \; card(D) \neq 0, \tag{1.3}$$

where $d_i \in D$ denotes the i^{th} diagnosis from the set of all elementary diagnoses D. An elementary diagnosis is a set of faults that are unisolable (Bartyś *et al.* 2006).

When all diagnoses point out only single faults, then the value of diagnosis accuracy is equal to 1. In the general case, the value of diagnosis accuracy is in the range $(0, 1]$.

The value of this measure depends on the chosen diagnosis approach. For example, in the case of a binary diagnostic matrix, diagnosis accuracy delivers different results based on whether the exoneration assumption is used or not. This difference results from ignoring information regarding the lack of some symptoms. Nevertheless, in some cases, such an approach is justified, because, after the occurrence of a fault, symptoms do not appear simultaneously. Different interim diagnoses are generated by the appearance of successive symptoms. The value of diagnosis accuracy also depends on the exoneration assumption.

Example 1.6.2.

Let us return to the example of BDM given in Tab. 1.1. Three possible diagnoses can be drawn from this table with the exoneration assumption: $\{f_1\}$, $\{f_2, f_3\}$, $\{f_4\}$. *The average number of faults in a diagnosis equals* $4/3$, *and the diagnosis accuracy is equal to* $3/4$. *In the case of binary diagnostic signals and with the exoneration assumption, the value of diagnosis accuracy is always equal to the diagnosability degree, because possible diagnoses are identical with D-classes.*

Without the exoneration assumption, there are also three possible diagnoses: $\{f_1, f_2, f_3\}$, $\{f_2, f_3\}$, $\{f_4\}$. *The diagnosis accuracy is then equal to* $[6/3]^{-1} = 1/2$.

An alternative definition was given in (Bartyś *et al.* 2006) as the theoretical mean diagnosis accuracy. It is defined as

$$\frac{1}{card(D)} \sum_{d_i \in D} \frac{1}{card(d_i)};\ card(D) \neq 0. \tag{1.4}$$

The definition (1.4) is sensitive to the number of faults in elementary diagnoses. Both definitions yield identical results if all elementary diagnoses contain an equal number of faults. However, if the number of faults is not equal for each diagnosis then (1.4) would give a numerically higher result than (1.3).

Example 1.6.3.

Two examples of BDMs are presented in Tab. 1.2.

For the case (a) $d_1 = \{f_1, f_2\}$ *and* $d_2 = \{f_3, f_4\}$. *Both definitions of diagnosis accuracy give identical results.*

$$\left[\frac{\sum_{d_i \in D} card(d_i)}{card(D)}\right]^{-1} = \frac{1}{card(D)} \sum_{d_i \in D} \frac{1}{card(d_i)} = \frac{1}{2}. \tag{1.5}$$

Table 1.2: Comparison of two definitions of diagnosis accuracy.

(a)

	f_1	f_2	f_3	f_4
s_1	1	1		
s_2			1	1

(b)

	f_1	f_2	f_3	f_4
s_1	1	1	1	
s_2				1

For the case (b) $d_1 = \{f_1, f_2, f_3\}$ *and* $d_2 = \{f_4\}$. *Diagnosis accuracy given by* (1.3) *is still equal to* $1/2$. *However, diagnosis accuracy given by* (1.4) *is equal to:*

$$\frac{1}{card(D)} \sum_{d_i \in D} \frac{1}{card(d_i)} = \frac{2}{3} > \frac{1}{2}. \tag{1.6}$$

An average number of faults in a diagnosis can be difficult to calculate in the case of complex diagnostic systems with multi-valued diagnostic signals or taking into account sequences of symptoms. In this case, the number of possible diagnoses grows exponentially with the number of faults.

1.6.3 Isolability index

The isolability index is often used as the fault isolability measure (Sarrate *et al.* 2012b; Sarrate *et al.* 2014). It is defined as the number of ordered pairs of isolable faults. The maximum number of the isolability index depends on the number of considered faults. If the isolability relation is asymmetric, then the maximal isolability index is equal to $K(K-1)$, where K is the number of considered faults. Similarly to the diagnosability degree, the isolability index is only defined for binary diagnostic signals.

Example 1.6.4.

Following the example given in Tab. 1.1, with the exoneration assumption, there are ten ordered pairs of faults if the first fault is isolable from the second: (f_1, f_2), (f_1, f_3), (f_1, f_4), (f_2, f_1), (f_2, f_4), (f_3, f_1), (f_3, f_4), (f_4, f_1), (f_4, f_2), (f_4, f_3). *It is worth noticing that, assuming exoneration, isolability is symmetric.*

Without the exoneration assumption, there are eight such pairs: (f_1, f_2), (f_1, f_3), (f_1, f_4), (f_2, f_4), (f_3, f_4), (f_4, f_1), (f_4, f_2), (f_4, f_3).

1.7 Solving the optimal sensor placement problem

Increasing the number of sensors is one of the simplest methods of increasing detectability and isolability of faults (Kościelny *et al.* 2006). There are different criteria and constraints formulated for selecting additional sensors. In general, this class of problems is called optimal sensor placement problems.

The binary relation between faults and diagnostic signals is the most widely used form of notation in FDI systems used for optimal sensor placement. In recent years, numerous papers were devoted to different problems related to this issue.

In (Kościelny *et al.* 2006), the effects of a reduction of the sensor set on the fault distinguishability are analyzed. Two definitions for the quantitative evaluation of changes of fault isolability and fault detectability were formulated. In (Travé-Massuyès *et al.* 2006), a method of searching for the optimal sensor set based on ARRs is proposed. First, all ARRs are found under the assumption that all sensor candidates are installed. Then, a sensor set, called Minimal Additional Sensor Set (MASS), is selected, which minimizes the cost while satisfying detectability and isolability requirements. However, this solution is computationally expensive. The diagnosability degree (1.2) was used as a measure of isolability. In general, the addition of new sensors entails expanding the set of faults (sensor faults). The diagnosability degree may then be nonmonotonic with respect to the cardinality of the set of sensors. It occurs when adding a new sensor does not increase the number of D-classes. In (Travé-Massuyès *et al.* 2006), the authors recommend that designers should assume that new sensors are sufficiently reliable and do not introduce additional faults.

A modified approach was proposed in (Rosich *et al.* 2007). The incremental approach was proposed in order to avoid an exponential explosion of a problem complexity. Instead of computing the complete set of ARRs, they are generated iteratively. The algorithm minimizes the cost of new sensors while satisfying the predefined FDI specifications.

In (Sarrate *et al.* 2007), the Binary Integer Programming is used to find the optimal sensor set using a set of all possible Minimal Structurally Overdetermined (MSO) sets. FDI requirements were ensured with nonlinear constraints. The resulting problem is computationally difficult to solve. This method was further improved in (Nejjari *et al.* 2010) and (Rosich *et al.* 2009). The FDI requirements were specified as linear constraints. The cost function was also linear, so the problem belonged to the Binary Integer Linear Programming (BILP) class. It can be efficiently solved with the branch-and-bound algorithm with a standard Linear Programming (LP) solver. Those methods were thoroughly compared in (Sarrate *et al.* 2012a).

Budgetary constraints with the isolability measure were analyzed in (Sarrate *et al.* 2012b) using a structural system model. The proposed method was applied to a Fuel Cell Stack System. The branch-and-bound algorithm was used to obtain the optimal solution.

In (Patan and Uciński 2008) the Fisher information was used to propose a method of solving the optimal sensor placement problem for distributed parameter systems with cost constraints. Again, the branch-and-bound algorithm was used to solve the problem.

1.8 Problem formulation

Various metrics of fault isolability are known. However, there is no general metric. Current definitions are insufficient because their applicability is limited. They are restricted to some types of diagnostic signals (e.g. diagnosability degree), or they are difficult to calculate for complex diagnostic structures (e.g. diagnosis accuracy). Therefore, the following problems arise:

- What universal properties of diagnostic structures should be extracted to formulate a generalized metric of fault isolability? How should such a metric be formally defined?
- How to quantitatively analyze and differentiate weak and uni- or bidirectional strong isolability? How do different exoneration assumptions affect metrics of isolability?
- How to use the proposed generalized metric of isolability in practical applications, including formulating and solving the optimal sensor placement problem with different constraints?

The main aim of this thesis is to propose a generalized metric of fault isolability.

1.9 Organization of the thesis

In Chapter 2, formal definitions of isolability for different types of diagnostic signals are formulated. Chapter 3 presents the new metric of isolability. In Chapter 4, the new measure of isolability is compared with other known metrics of isolability. In Chapter 5, different optimal sensor placement problems using the new metric are formulated. Chapter 6 summarizes the thesis.

Chapter 2

Isolability definitions for different types of diagnostic signals

2.1 Introduction

Fault isolability is usually defined in the context of the adopted diagnostic method, in particular, the form of notation of the faults–symptoms relation. Most often, fault isolability was analyzed in the case of a BDM (incidence matrix) derived from the structure of linear equations of residuals in the internal form (Gertler 1997; Gertler 1998).

Fault isolability, obtained on the basis of BDM and FIS derived from expert knowledge was analyzed in the studies (Korbicz *et al.* 2004; Kościelny *et al.* 2006). In the case of multi-valued and continuous diagnostic signals, it is not always possible to determine subsets of unisolable faults. Then, conditional isolability is considered. Similarly, when analyzing isolability in residual space, faults with partially overlapping regions in residual space are considered conditionally isolable (Korbicz *et al.* 2004).

A definition of isolability was proposed in (Travé-Massuyès *et al.* 2006). It is related to the definitions using residual space (Isermann 2006). It should be noted that conditional isolability defined in (Travé-Massuyès *et al.* 2006) is not symmetric. This definition makes it possible to consider multi-valued diagnostic signals, but it is not useful in the case of symptom sequences.

2.2 Fault isolability based on binary diagnostic matrix

2.2.1 Binary diagnostic matrix

In the simplest case diagnostic signals are binary. There are many notations of the relation between binary diagnostic signals and faults. The structure called Binary Diagnostic Matrix is one of the most widely used structures. BDM is easily understood by industry experts and engineers. It facilitates the cooperation and knowledge transfer between experts and designers of FDI systems. Please note, that all definitions given in this section can be easily extended for other binary forms of notation of diagnostic relations.

The following set of linear equations defining the dependence of process outputs on process inputs and faults is often used in fault diagnostics (Gertler 1991; Gertler 1998):

$$\boldsymbol{y}(s) = \boldsymbol{G}(s)\boldsymbol{u}(s) + \boldsymbol{H}(s)\hat{\boldsymbol{f}}(s), \tag{2.1}$$

where:

$\boldsymbol{y}$ – vector of outputs,

$\boldsymbol{u}$ – vector of inputs,

$\hat{\boldsymbol{f}}$ – vector of magnitudes of faults,

$\boldsymbol{G}(s)$ – matrix of input–output transfer functions,

$\boldsymbol{H}(s)$ – matrix of faults–output transfer functions.

This model is valid for linear systems. It can also describe nonlinear systems in the vicinity of a selected operating point. This point usually corresponds to the nominal or average operating conditions of the system.

A Binary Diagnostic Matrix is widely used for description of the faults–symptoms relation. It can be obtained by many methods, e.g., by applying models considering influence of faults, structural analysis or using expert knowledge. It was named in (Gertler 1998) as a structure matrix of residual sets.

In the ideal case when the matrix $\boldsymbol{H}(s)$ is known then the elements of BDM can be obtained in the following way:

$$v_{j,k} = \begin{cases} 0 & \text{if } H_{j,k}(s) = 0 \\ 1 & \text{if } H_{j,k}(s) \neq 0 \end{cases}. \tag{2.2}$$

Therefore, a BDM defines a relation between faults and binary diagnostic signals. An example of BDM is presented in Tab. 2.1.

Table 2.1: Example of binary diagnostic matrix.

	f_1	f_2	f_3
s_1	1	0	1
s_2	1	1	0
s_3	0	1	1

BDM is a form of notation of a relationship specified as a subset of the Cartesian product of diagnostic signal sets $S = \{s_j : j = 1, 2, \ldots, J\}$ and faults $F = \{f_k : k = 1, 2, \ldots, K\}$:

$$R_{F,S} \subseteq S \times F. \tag{2.3}$$

A fault signature in BDM is defined as a vector of values of diagnostic signals corresponding to a given fault. Therefore, the columns of the binary diagnostic matrix constitute the signatures of the corresponding faults.

$$\boldsymbol{\phi}(f_k) = \begin{bmatrix} v_{1,k} & v_{2,k} & \ldots & v_{J,k} \end{bmatrix}^T. \tag{2.4}$$

In different forms of notation of diagnostic relations, the fault signature is usually defined in a similar way.

BDM can be expressed in an alternative way using logic functions and logic rules. One rule is associated with each fault f_k:

$$If\ (s_1 = v_{1,k}) \wedge \cdots \wedge (s_j = v_{j,k})\ \wedge (s_J = v_{J,k})\ then\ f_k. \tag{2.5}$$

The rows of BDM can be represented by the following rules:

$$If\ (s_j = 1)\ then\ f_a \vee \cdots \vee f_b, \text{ where: } f_a, \ldots, f_b \in \{f_k : v_{j,k} = 1\}. \tag{2.6}$$

Example 2.2.1.

The following rules can be obtained from the BDM presented in Tab. 2.1,

1. $If\ (s_1 = 1) \wedge (s_2 = 1) \wedge (s_3 = 0)\ then\ f_1$
2. $If\ (s_1 = 0) \wedge (s_2 = 1) \wedge (s_3 = 1)\ then\ f_2$
3. $If\ (s_1 = 1) \wedge (s_2 = 0) \wedge (s_3 = 1)\ then\ f_3$

The following rules can be given for the rows of this BDM.

1. $If\ (s_1 = 1)\ then\ f_1 \vee f_3.$
2. $If\ (s_2 = 1)\ then\ f_1 \vee f_2.$
3. $If\ (s_3 = 1)\ then\ f_2 \vee f_3.$

2.2.2 Definitions of isolability

The definitions of unisolability and isolability of faults with the exoneration assumption based on BDM (Kościelny *et al.* 2016) are given below.

Definition 2.2.1. *Faults $f_k, f_m \in F$ are unisolable with the exoneration assumption based on BDM iff their signatures are identical.*

$$f_k R_{U|BDM} f_m \Leftrightarrow \underset{s_j \in S}{\forall} [v_{j,k} = v_{j,m}] \tag{2.7}$$

$R_{U|BDM}$ denotes the unisolability relation based on BDM. Subsequent isolability and unisolability relations, for other forms of notation, are denoted in a similar way.

Relation $R_{U|BDM}$ can be used to split the set of faults F into subsets of unisolable faults.

Definition 2.2.2. *Faults $f_k, f_m \in F$ are isolable in BDM under the exoneration assumption iff their signatures are different.*

$$f_k R_{I|BDM} f_m \Leftrightarrow \underset{s_j \in S}{\exists} [v_{j,k} \neq v_{j,m}]. \tag{2.8}$$

This definition is analogous to the definition of a weakly isolating structure (Definition 1.3.4), i.e., to a structure where residual response to every fault is nonzero and different (Gertler 1998; Gertler 2000).

Theoretically, the maximum number of faults that can be isolated on the basis of the set of J binary diagnostic signals is equal to $(2^J - 1)$. In a faultless state, the values of all diagnostic signals are equal to zero.

The basic approach to increasing the fault isolability is based on generation of secondary residuals. The method of designing secondary residuals depends on whether the internal form of primary residuals is known (Gertler 1998).

Example 2.2.2.

In this example, we will analyze the two tank system presented in Fig. 2.1. This system can be

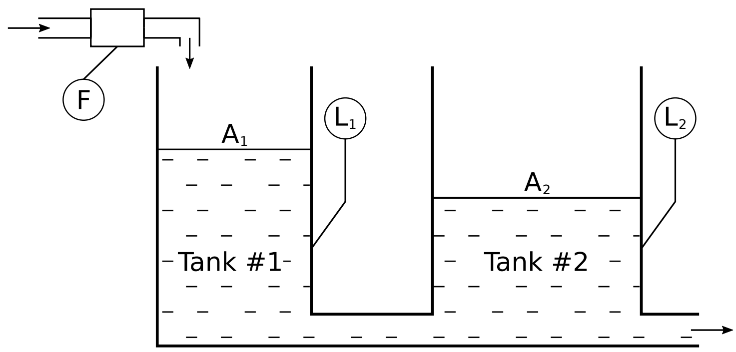

Figure 2.1: Two tank system. F – input flow rate, L_1, L_2 – water levels in the tanks, A_1, A_2 – cross sections of tanks #1 and #2.

described by the following balance equations:

$$A_1 \frac{dL_1}{dt} = F - \alpha_{1,2} S_{1,2} \sqrt{2g\,(L_1 - L_2)}, \tag{2.9}$$

$$A_2 \frac{dL_2}{dt} = \alpha_{1,2} S_{1,2} \sqrt{2g\,(L_1 - L_2)} \; - \; \alpha_2 S_2 \sqrt{2gL_2}\,, \tag{2.10}$$

where $\alpha_{1,2}$ is the discharge coefficient and $S_{1,2}$ denotes the cross section of pipes between tanks #1 and #2.

The considered faults are presented in Tab. 2.2.

Table 2.2: List of the considered faults in the two tanks system.

Fault	Description
f_1	Clog between tanks #1 and #2
f_2	Clog in the outlet of tank #2
f_3	Leakage from tank #1
f_4	Leakage from tank #2
f_5	Faulty measurement of input flow (F)
f_6	Faulty measurement of level in tank #1 (L_1)
f_7	Faulty measurement of level in tank #2 (L_2)

We can subtract both sides of equations (2.9) *and* (2.10) *and extend them with information about faults and obtain the following residuals:*

$$r_1 = F + \hat{f}_5 - \hat{f}_3 - A_1 \frac{d(L_1 + \hat{f}_6)}{dt} - \alpha_{1,2}(S_{1,2} - \hat{f}_1)\sqrt{2g\left(L_1 + \hat{f}_6 - \left(L_2 + \hat{f}_7\right)\right)}, \tag{2.11}$$

$$r_2 = \alpha_{1,2}\left(S_{1,2} - \hat{f}_1\right)\sqrt{2g\left(L_1 + \hat{f}_6 - \left(L_2 + \hat{f}_7\right)\right)} + \\ - \alpha_2(S_2 - \hat{f}_2)\sqrt{2g\left(L_2 + \hat{f}_7\right)} - \hat{f}_4 - A_2\frac{d(L_2 + \hat{f}_7)}{dt}. \tag{2.12}$$

The notation $\hat{f}_k$ denotes the magnitude of the fault f_k.

We can obtain the third residual by combining r_1 and r_2:

$$r_3 = r_1 + r_2 = F + \hat{f}_5 - \alpha_2(S_2 - \hat{f}_2)\sqrt{2g\left(L_2 + \hat{f}_7\right)} + \\ - \hat{f}_3 - \hat{f}_4 - A_2\frac{d(L_2 + \hat{f}_7)}{dt} - A_1\frac{d(L_1 + \hat{f}_6)}{dt}. \tag{2.13}$$

Binary diagnostic signals can be derived from residuals by thresholding. If a value of a residual is smaller than some predetermined threshold, then the value of the corresponding diagnostic signal is equal to 0. Otherwise, it is equal to 1. The binary diagnostic matrix describing the two tank system is shown in Tab. 2.3.

Using Tab. 2.3 and Definitions 2.2.1 and 2.2.2, the following subsets of unisolable faults can be determined: $\{f_1\}$, $\{f_2, f_4\}$, $\{f_3, f_5\}$, $\{f_6, f_7\}$.

Table 2.3: Binary diagnostic matrix of the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
s_1	1	0	1	0	1	1	1
s_2	1	1	0	1	0	1	1
s_3	0	1	1	1	1	1	1

2.3 Fault isolability in an information system

2.3.1 Fault information system FIS

The fault information system FIS is a multi-valued extension of the structure of residual sets. FIS has been defined in (Korbicz *et al.* 2004; Kościelny 1999) as an information system:

$$FIS = \langle F, S, V_S, q \rangle, \tag{2.14}$$

where:

F - set of faults,

S- set of diagnostic signals,

$V_S = \bigcup_{s_j \in S} V_j$ - set of values of diagnostic signals,

V_j - set of values of diagnostic signal s_j,

q - function:

$$q : F \times S \rightarrow V_S, \tag{2.15}$$

associating each element of the Cartesian product with a subset of values of the diagnostic signal s_j characteristic of an occurrence of fault f_k:

$$q(f_k, s_j) \equiv V_{j,k} \subset V_j. \tag{2.16}$$

For example, $V_j = \{-1, 0, +1\}$ and $V_{j,k} = \{-1, +1\} \subset V_j$.

We assume that the zero value of the diagnostic signal, $s_j = 0$, corresponds to a fault-free state The other values of diagnostic signals are symptoms of faults.

FIS assigns a subset of diagnostic signals to each fault. In fact, the BDM is a special case of FIS. If the set of values of all diagnostic signals is equal to $V_s = \{0, 1\}$ and $V_{j,k}$ is a single element set, then FIS is identical with BDM. FIS has the following features:

- each diagnostic signal has its own individual set of values;
- the set V_j of values of the j^{th} diagnostic signal is finite and contains at least 2 elements;
- every set $V_{j,k}$ in FIS is a subset of the set V_j of values of the j^{th} diagnostic signal.

The column of FIS generalizes the signature defined by the formula (2.4).

$$\mathbf{\Phi}(f_k) = \begin{bmatrix} V_{1,k} & V_{2,k} & \dots & V_{J,k} \end{bmatrix}^T. \tag{2.17}$$

This signature can be rewritten in the form:

$$If\ (s_1 \in V_{1,k}) \wedge \cdots \wedge (s_j \in V_{j,k}) \wedge (s_J \in V_{J,k})\ then\ f_k. \tag{2.18}$$

We can also define the rules corresponding to the rows of FIS. The number of rules corresponding to one row of FIS is equal to the number of values of diagnostic signal $v \in V_j$ different from zero.

$$If\ (s_j = v_{j,k} \neq 0) \text{ then } f_a \vee \cdots \vee f_b, \tag{2.19}$$

where: $f_a, \dots, f_b \in \{f_k : v_{j,k} \neq 0\}$.

The specific vector of values of diagnostic signals is called an alternative signature (Bartyś 2013). In the case of binary diagnostic signals, each fault is associated with exclusively one alternative signature, while in the case of multi-valued diagnostic signals there might be multiple alternative signatures.

2.3.2 Definition of isolability

In FIS, a fault can be indicated by multiple values of a diagnostic signal. Therefore, in some cases it is not possible to unambiguously determine if a pair of faults is isolable or not. In the studies (Kościelny *et al.* 2006) and (Kościelny *et al.* 2016), the following definitions of unconditional and conditional unisolability in FIS were given.

Definition 2.3.1. *Faults $f_k, f_m \in F$ are unisolable (unconditionally unisolable) in FIS iff their signatures are identical.*

$$f_k R_{U|FIS} f_m \Leftrightarrow \underset{s_j \in S}{\forall} V_{j,k} = V_{j,m}. \tag{2.20}$$

The conditional unisolability is possible in the case of FIS and other notations of multi-valued forms of the diagnostic signals–faults relation. The faults can be unisolable only for some, but not for all values of diagnostic signals.

Definition 2.3.2. *Faults $f_k, f_m \in F$ are conditionally unisolable in FIS iff for each diagnostic signal, the intersection of subsets of its values corresponding to faults f_k and f_m is nonzero and those faults are not unconditionally unisolable.*

$$f_k R_{CU|FIS} f_m \Leftrightarrow \underset{s_j \in S}{\forall} V_{j,k} \cap V_{j,m} \neq \emptyset \wedge \underset{s_j \in S}{\exists} V_{j,k} \neq V_{j,m}. \tag{2.21}$$

Definition 2.3.3. *Faults $f_k, f_m \in F$ are unconditionally isolable in FIS iff there is a diagnostic signal for which subsets of values corresponding to those faults are disjoint:*

$$f_k R_{I|FIS} f_m \Leftrightarrow \underset{s_j \in S}{\exists} V_{j,k} \cap V_{j,m} = \emptyset. \tag{2.22}$$

Since BDM is a special case of FIS, Definitions 2.3.1 and 2.3.3 can also be used for BDM. If all diagnostic signals are binary, then conditional isolability is not possible.

The maximum number of faults that can be isolated on the basis of a set of tri-valued diagnostic signals is $3^J - 1$. The maximum number of isolable faults does not exceed the product of the power of sets V_j for all diagnostic signals: $\prod_{s_j \in S} |V_j|$. In practice, the number of isolated faults is usually much smaller.

Multi-valued evaluation of residual values may lead to an increased fault isolability in comparison with binary evaluation. Usually, additional knowledge of the sign of a residual in three-valued evaluation increases fault isolability.

Example 2.3.1.

FIS presented in Tab. 2.4 can be obtained for the two tank system described in Section 2.2.2 by additionally considering the direction of the change of residuals.

It was assumed that component faults $f_1 - f_4$ can only have positive values while sensor faults (f_5, f_6 and f_7) can change in the positive as well as in the negative direction. Applying

Table 2.4: FIS for the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	V_j
s_1	+1	0	-1	0	-1,+1	-1,+1	-1,+1	{-1,0,+1}
s_2	-1	+1	0	-1	0	-1,+1	-1,+1	{-1,0,+1}
s_3	0	+1	-1	-1	-1,+1	-1,+1	-1,+1	{-1,0,+1}

Definitions 2.3.1, 2.3.2 and 2.3.3 results in the following conclusions (Tab. 2.5): {f_6, f_7} are unconditionally unisolable, {f_3, f_5} are conditionally isolable. All other faults are unconditionally isolable.

Table 2.5: Distinguishability structure of pairs of faults for the FIS in Tab. 2.4.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
f_1							
f_2	+						
f_3	+	+					
f_4	+	+	+				
f_5	+	+	±	+			
f_6	+	+	+	+	+		
f_7	+	+	+	+	+	−	

\+ unconditionally isolable pairs of faults

± conditionally isolable pairs of faults

− unisolable pairs of faults

2.4 Fault isolability based on directional residuals

2.4.1 Vectors of fault directions in residual space

This method of fault isolation is referred to as directional residuals. In order to isolate faults, a set of residuals is designed in such a way that the occurrence of faults is characterized by a particular direction in the space of residuals (called parity space). Therefore, each fault corresponds to an individually designed directional vector (Chen and Patton 1999; Gertler 1998). This approach is illustrated in Fig. 2.2.

Primary directional residuals are derived from the equations of residuals given in the internal form (2.2), by replacing the transfer functions $H_{j,k}(s)$ with gains $c_{j,k}$ of the particular residuals:

$$c_{j,k} = \begin{cases} 0 & \text{if } H_{j,k}(s) = 0 \\ \lim\limits_{s \to 0} H_{j,k}(s) & \text{if } H_{j,k}(s) \neq 0 \end{cases}. \tag{2.23}$$

Then the vector of gains of residuals corresponding to a given fault $f_k,\ k = 1, \ldots, K$ defines the characteristic direction $\boldsymbol{v}_k = [c_{1,k}, \ldots, c_{j,k}, \ldots, c_{J,k}]$ in the parity space. A fault, after the transient state, manifests itself in this direction. This is the basis of the directional residual method (Chen and Patton 1999; Gertler 1998; Patton *et al.* 2000b).

Fault isolation occurs after assessment of the coincidence of the direction of the residual vector with the direction specific to individual faults as in Fig. 2.2.

Vectors of fault directions in the parity space are designed on the basis of a system model affected by faults. In theory, vectors can also be determined with machine learning techniques. However, this requires data from an object affected by faults and is hardly ever feasible. In addition, expert knowledge is usually not sufficient for determination of directional vectors for individual faults.

2.4.2 Definition of isolability

The following definition of fault isolability can be formulated based on the vectors of fault directions in the parity space:

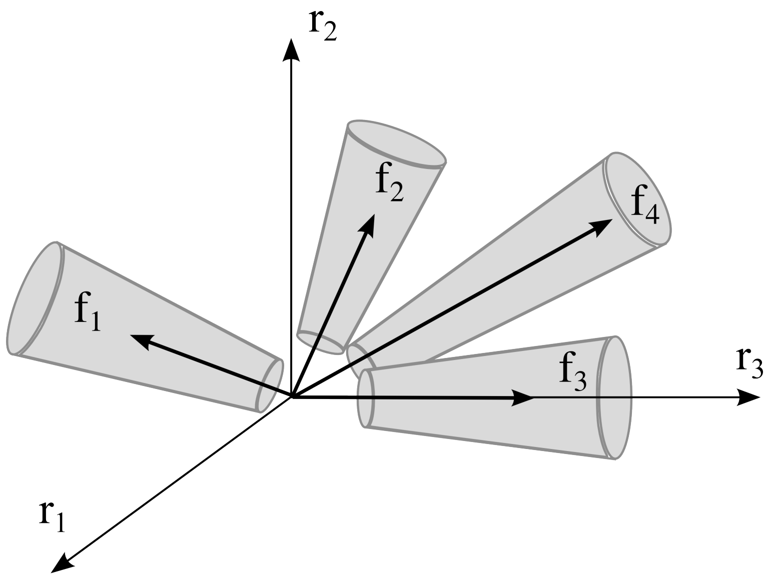

Figure 2.2: Characteristic directions of faults in the parity space.

Definition 2.4.1. *Faults are isolable if their directions in the parity space are different.*

Such a definition is not very useful in practice because a small deviation in the directions makes fault isolation problematic.

In practice, it is often not necessary to examine directions of all residuals. A sufficient condition of isolability for any pair of faults is that corresponding directions in the plane defined by any two residuals differ by more than a predetermined angle.

Directional fault vectors f_k and f_m in the plane defined by the residuals r_j and r_p have the form: $\boldsymbol{v}_k = [c_{j,k}, c_{p,k}]$ and $\boldsymbol{v}_m = [c_{j,m}, c_{p,m}]$. The angle between them is determined by the formula:

$$\alpha = \arccos \frac{c_{j,k} c_{j,m} + c_{p,k} c_{p,m}}{|\boldsymbol{v}_k| \, |\boldsymbol{v}_m|}. \tag{2.24}$$

Example 2.4.1.

Let us continue the example from Section 2.2.2. An alternative form of residuals can be obtained by linearization and Laplace transformation:

$$
\begin{aligned}
r_1(s) &= -L_1(s) + \frac{k_1}{T_1 s + 1} F(s) + \frac{k_2}{T_1 s + 1} L_2(s) = \\
&= -\hat{f}_6(s) + \frac{k_1}{T_1 s + 1}\hat{f}_5(s) + \frac{k_2}{T_1 s + 1}\hat{f}_7(s) + \frac{k_3}{T_1 s + 1}\hat{f}_1(s) - \frac{k_4}{T_1 s + 1}\hat{f}_3(s),
\end{aligned}
\quad (2.25)
$$

$$
\begin{aligned}
r_2(s) &= -L_2(s) + \frac{k_5}{T_2 s + 1} L_1(s) = \\
&= -\hat{f}_7(s) + \frac{k_5}{T_2 s + 1}\hat{f}_6(s) - \frac{k_6}{T_2 s + 1}\hat{f}_1(s) + \frac{k_7}{T_2 s + 1}\hat{f}_2(s) - \frac{k_8}{T_2 s + 1}\hat{f}_4(s),
\end{aligned}
\quad (2.26)
$$

$$
\begin{aligned}
r_3(s) &= -L_2(s) - \frac{k_9 s}{T_2 s + 1} L_1(s) + \frac{k_{10}}{T_2 s + 1} F(s) = \\
&= -\hat{f}_7(s) - \frac{k_9 s}{T_2 s + 1}\hat{f}_6(s) + \frac{k_7}{T_2 s + 1}\hat{f}_2(s) - \frac{k_8}{T_2 s + 1}\hat{f}_4(s) + \frac{k_{10}}{T_2 s + 1}\hat{f}_5(s) - \frac{k_{11}}{T_2 s + 1}\hat{f}_3(s).
\end{aligned}
\quad (2.27)
$$

The direction corresponding to a fault is given in the parity space spanned by primary residuals. Consequently, the fault directions can be derived. These directions are shown in Tab. 2.6.

Table 2.6: Fault directions in the three-dimensional space of residuals for the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
r_1	k_3	0	$-k_4$	0	$\pm k_1$	± 1	$\pm k_2$
r_2	$-k_6$	k_7	0	$-k_8$	0	$\pm k_5$	± 1
r_3	0	k_7	$-k_{11}$	$-k_8$	$\pm k_{10}$	$\pm k_9$	± 1

Each fault has a different direction in parity space. For example faults f_2 and f_4, which were unisolable using BDM, can be easily isolated as they have opposite directions.

2.5 Fault isolability based on sequential residuals

The order of emergence of symptoms is important information which can be useful for diagnostic inference in the diagnostic process. Different time instances of emerging symptoms can be used for isolation of faults in linear systems. For linear systems, the sequence of emergence of symptoms can be determined for each fault on the basis of transmittance $H_{j,k}$:

$$H_{j,k}(s) = y_j(s)/\hat{f}_k(s);\ k = 1, \ldots, K, \tag{2.28}$$

where $y_j(s)$ is the j^{th} system output.

The order of emergence of symptoms depends on the dynamic properties of the system, the dynamics of the fault (abrupt, incipient, etc.) and dynamic parameters of the fault detection algorithm.

Let us assume occurrence of a single fault f_k. The residual equation becomes:

$$r_j(s)|_{f_k} = H_{j,k}(s)\hat{f}_k(s);\ f_m = 0;\ m = 1, 2, \ldots, K,\ m \neq k. \tag{2.29}$$

The residual can be transformed into the time domain by applying the inverse Laplace transform:

$$r_j\,(t)\Big|_{f_k} = L^{-1}\ r_j\,(s)\Big|_{f_k} = L^{-1}[H_{j,k}(s)\hat{f}_k(s)]. \tag{2.30}$$

The time after which the k^{th} fault will affect the j^{th} diagnostic signal can be easily determined. Let us assume that the function $\hat{f}_k(t)$ is a step function and the threshold value for the j^{th} residual is A_j. Thresholds should be calculated for all residuals that are sensitive to fault f_k. For $\hat{f}_k(t)$ different from a step function, the times of emergence of symptoms will be different, but their order should not change for a given A_j (Syfert and Kościelny 2009). Let us denote by $es_{j,p}(f_k)$ the elementary sequence, i.e., a sequence of two symptoms j and p for the fault f_k. The sequences are equal if the order of symptoms in all sequences is identical.

Sequences of some pairs of symptoms can be recognized using expert knowledge (Syfert and Kościelny 2009) or by means of causal graphs GP (Sztyber *et al.* 2015; Sztyber 2017). Graph GP is a convenient method of describing the structure of models. It makes it possible to determine the sensitivity of the chosen model structures to faults. A simple example of such a graph is shown in Fig. 2.3. Faults f_1 and f_2 are isolable because the elementary sequences are different $es_{1,2}(f_1) = < s_1,\ s_2 > \neq es_{1,2}(f_2) = < s_2,\ s_1 >$.

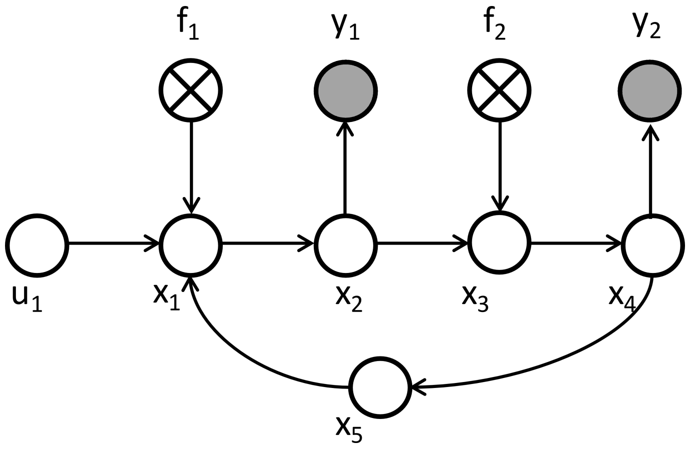

Figure 2.3: Example of a graph GP (Sztyber 2017); x_i - process variables, u_1 - input, y_i - outputs, $s_1(y_1, u_1)$, $s_2(y_2, u_1)$ - diagnostic signals: $es_{1,2}(f_1) = < s_1,\ s_2 >$, $es_{1,2}(f_2) = < s_2,\ s_1 >$.

2.5.1 Definition of fault isolability

The order of emergence of symptoms brings new, useful information apart from the values of diagnostic signals. Therefore, an appropriate extension to the definitions of fault isolability have been formulated (Kościelny *et al.* 2016):

Definition 2.5.1. *Faults $f_k, f_m \in F$ are unisolable (unconditionally unisolable) on the basis of elementary sequences of symptoms iff their elementary sequences of symptoms are identical.*

$$f_k R_{U|SEQ} f_m \Leftrightarrow \underset{s_j, s_p \in S}{\forall} es_{j,p}(f_k) = es_{j,p}(f_m). \tag{2.31}$$

In many cases, elementary sequences make it possible to isolate faults, which are unisolable on the basis of analysis of values of diagnostic signals.

Definition 2.5.2. *Faults $f_k, f_m \in F$ are unconditionally isolable on the basis of elementary sequences of symptoms iff there exists at least one different elementary sequence of symptoms for these faults.*

$$f_k R_{I|SEQ} f_m \Leftrightarrow \underset{s_j, s_p \in S}{\exists} [es_{j,p}(f_k) \neq es_{j,p}(f_m)]. \tag{2.32}$$

Definitions 2.5.1 and 2.5.2 require knowledge about both sequences $es_{j,p}(f_k)$ and $es_{j,p}(f_m)$.

Definition 2.5.3. *Faults $f_k, f_m \in F$ are conditionally unisolable on the basis of elementary sequences of symptoms iff they are not unconditionally isolable and there exists a pair of symptoms for which the elementary sequence of symptoms can be determined for only one of those faults.*

If the internal form of residuals is known, one can design (Kościelny *et al.* 2013) pairs of secondary residuals with different sequences for certain pairs of faults. In addition, these residuals can have an arbitrary delays $\tau_{k,j,p}$. Such a sequence can be written as:

$$es^d_{j,p}(f_k, \tau_{k,j,p}) =< s_j, \tau_{k,j,p}, s_p > . \tag{2.33}$$

This sequence will be called a delay-designed elementary sequence.

Consequently, we can obtain a pair of secondary residuals for any fault f_k (Kościelny *et al.* 2013). The same pair of primary residuals may be used for generating additional pairs of secondary residuals for two or more faults that are detectable by primary residuals. In such a case, different time delays between symptoms should be selected. Particularly, it can be assumed that $\tau_{k,j,p} = 0$. It means that the symptoms of signals s_j and s_p are appearing simultaneously.

Analogous definitions of unisolability and isolability of faults apply to the delay-designed elementary sequence $es^d_{j,p}(f_k, \tau_{k,j,p})$ and to the sequence $es_{j,p}(f_k)$ (Def. 2.5.1 and 2.5.2).

Definition 2.5.4. *Faults $f_k, f_m \in F$ are unisolable on the basis of delay-designed elementary sequences of symptoms $es^d_{j,p}(f_k, \tau_{k,j,p})$ iff the corresponding delay-designed elementary sequences of symptoms are identical.*

$$f_k R_{U|DSEQ} f_m \Leftrightarrow \underset{s_j, s_p \in S}{\forall} es^d_{j,p}(f_k, \tau_{k,j,p}) = es^d_{j,p}(f_m, \tau_{m,j,p}). \tag{2.34}$$

Definition 2.5.5. *Faults $f_k, f_m \in F$ are isolable on the basis of delay-designed elementary sequences of symptoms $es^d_{j,p}(f_k, \tau_{k,j,p})$ iff there exists at least one different delay-designed elementary sequence for these faults.*

$$f_k R_{I|DSEQ} f_m \Leftrightarrow \underset{s_j, s_p \in S}{\exists} es^d_{j,p}(f_k, \tau_{k,j,p}) \neq es^d_{j,p}(f_m, \tau_{m,j,p}). \tag{2.35}$$

It is worth mentioning that (2.35) holds when $\tau_{k,j,p} \neq \tau_{m,j,p}$ and the order of symptoms is the same or when the order of symptoms is different.

There is a significant difference between Definitions 2.5.2 and 2.5.5. In the case of Definition 2.5.5, faults with the identical order of a pair of symptoms but different delays are also isolable. However, knowledge of the internal form of residuals is required to generate secondary residuals.

Example 2.5.1.

Applying the method of sequential residuals for the two tank system given in Section 2.2.2 with a known internal form of primary residuals results in the sequences of symptoms presented in Tab. 2.7. Numbers indicate the order of symptoms. The appearance of two identical numbers

Table 2.7: Example of sequences of symptoms. The numbers indicate the order of symptoms.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
s_1	1	-	1	-	1	1	2
s_2	2	1	-	1	-	2	1
s_3	-	1	2	1	2	2	1

in a signature means that it is impossible to determine which symptom will appear first. In this example, it is assumed that the time constant of the second tank is greater than the time constant of the first tank ($T_1 < T_2$).

Using Definitions 2.5.1 and 2.5.2, the following sets of unisolable faults can be distinguished: $\{f_1\}$, $\{f_2, f_4\}$, $\{f_3, f_5\}$, $\{f_6\}$, $\{f_7\}$. The pair of faults $\{f_6, f_7\}$ which were unisolable by means of BDM with primary residuals can be isolated using sequences of symptoms.

It is possible to further improve isolability by designing secondary sequential residuals. In the discussed two tank system, it is beneficial to introduce the following residuals:

$$r_{1,3/5} = \frac{k_{10}}{(T_2 s + 1)} r_1 = \frac{k_{10}}{(T_2 s + 1)} \left(-\hat{f}_6(s) + \frac{k_1}{T_1 s + 1} \hat{f}_5(s) + \right. \\ \left. + \frac{k_2}{T_1 s + 1} \hat{f}_7(s) + \frac{k_3}{T_1 s + 1} \hat{f}_1(s) - \frac{k_4}{T_1 s + 1} \hat{f}_3(s) \right), \quad (2.36)$$

$$r_{3,1/5} = \frac{k_1}{(T_1 s + 1)} r_3 e^{-\tau_{3,1} s} = \frac{k_1 e^{-\tau_{3,1} s}}{(T_1 s + 1)} \left(-\hat{f}_7(s) - \frac{k_9 s}{T_2 s + 1} \hat{f}_6(s) + \right.$$
$$\left. + \frac{k_7}{T_2 s + 1} \hat{f}_2(s) - \frac{k_8}{T_2 s + 1} \hat{f}_4(s) + \frac{k_{10}}{T_2 s + 1} \hat{f}_5(s) - \frac{k_{11}}{T_2 s + 1} \hat{f}_3(s) \right). \quad (2.37)$$

Residuals $r_{1,3/5}$ and $r_{3,1/5}$ make it possible to isolate f_5 from f_3. If fault f_5 occurs, then both residuals will be delayed by $\tau_{3,1}$.

2.6 Functional diagnosability and detectability

Analytical Redundancy Relations (ARR) were first proposed by Chow and Willsky in (Chow and Willsky 1984). They combine input and output signals of a diagnosed process and their derivatives into a set of relations. They can be used for determining the set of diagnostic signals and their relationships with faults. The functional diagnosability definition for nonlinear dynamical systems based on ARR was proposed in (Verdière *et al.* 2015).

The following, nonlinear dynamic parametric models are considered:

$$\begin{cases} \dot{\boldsymbol{x}}(t, \boldsymbol{p}, \boldsymbol{f}) = \boldsymbol{g}(\boldsymbol{x}(t, \boldsymbol{p}), \boldsymbol{u}(t), \boldsymbol{f}, \boldsymbol{\epsilon}, \boldsymbol{p}), \\ \boldsymbol{y}(t, \boldsymbol{p}, \boldsymbol{f}) = \boldsymbol{h}(\boldsymbol{x}(t, \boldsymbol{p}), \boldsymbol{u}(t), \boldsymbol{f}, \boldsymbol{\epsilon}, \boldsymbol{p}), \\ \boldsymbol{x}(t_0, \boldsymbol{p}, \boldsymbol{f}) = \boldsymbol{x_0}, \\ t_0 \leq t \leq T. \end{cases} \quad (2.38)$$

Where:

$\boldsymbol{x}(t, \boldsymbol{p}, \boldsymbol{f}) \in \mathbb{R}^N$ - vector of state variables,

$\boldsymbol{y}(t, \boldsymbol{p}, \boldsymbol{f}) \in \mathbb{R}^M$ - vector of process outputs,

$\boldsymbol{u}(t) \in \mathbb{R}^R$ - vector of process inputs,

$\boldsymbol{f}$ - vector of faults,

$\boldsymbol{\epsilon}$ - vector of stochastic noise in the system,

$\boldsymbol{p}$ - vector of process parameters.

Let us consider the set of ARRs:

$$w_i(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{f}, \bar{\boldsymbol{\epsilon}}, \boldsymbol{p}) = 0, \quad i = 1, \ldots, M \quad (2.39)$$

where:

w_i - $i^{th} ARR$,

$\bar{\boldsymbol{\vartheta}}$ denotes some vector $\boldsymbol{\vartheta}$ and its time derivatives up to some unspecified order (Staroswiecki and Comtet-Varga 2001).

Any ARR equation can be divided into deterministic and stochastic parts. The stochastic part is difficult to model. Therefore, very often only the deterministic part $w_{d,i}$ is used for fault detection (Staroswiecki and Comtet-Varga 2001; Verdière *et al.* 2015):

$$w_i(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{f}, \bar{\boldsymbol{\epsilon}}, \boldsymbol{p}) = w_{d,i}(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{f}, \boldsymbol{p}) = 0. \tag{2.40}$$

The $w_{d,i}$ can be further decomposed:

$$w_{d,i}(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{f}, \boldsymbol{p}) = w_{0,i}(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{p}) - w_{1,i}(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{f}, \boldsymbol{p}). \tag{2.41}$$

Combining (2.40) with (2.41), we can see that $w_{0,i}(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{p}) = w_{1,i}(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{f}, \boldsymbol{p})$. The fault-free term, $w_{0,i}$, is also known as the computational form of a residual. The $w_{1,i}$ part depends on faults and is known as the internal form.

Following the method presented in (Verdière *et al.* 2015), let us denote by $\boldsymbol{f}_{[k]}$ the fault vector, where all components except f_k are equal to 0. $\boldsymbol{f}_{[k]}$ can be then understood as a fault vector resulting from the single fault f_k.

Finally, the functional fault signature can be defined (Verdière *et al.* 2015).

Definition 2.6.1. *The functional fault signature is a function $FSig$ which associates the vector $\left(w_{1,i}\left(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{f}_{[k]}, \boldsymbol{p}\right)\right)_{i=1,\ldots,M}$ to a fault f_k.*

From this definition, we can see that the functional fault signature $\boldsymbol{FSig}(f_k)$ is a vector consisting of the internal form of residuals calculated under the assumption of the occurrence of a single fault f_k. Let us also denote by $FSig^{(i)}(f_k) = w_{1,i}\left(\bar{\boldsymbol{y}}, \bar{\boldsymbol{u}}, \boldsymbol{f}_{[k]}, \boldsymbol{p}\right)$ the i^{th} component of $\boldsymbol{FSig}(f_k)$. The functional signatures can be collected as columns of a Functional Signature Matrix (Verdière *et al.* 2015).

Functional diagnosability has very good isolability properties. Unfortunately, it requires knowledge about the internal form of residuals. This is a strong requirement that often cannot be met in practice.

2.6.1 Functional diagnosability definitions

The following set of isolability definitions of functional diagnosis was given in (Verdière *et al.* 2015):

Definition 2.6.2. *Two faults f_k and f_m are input-strongly functionally isolable if for all inputs u, there exists at least one index i and a finite time $t_1 \in (t_0, T]$ such that for all $t \in [t_0, t_1]$, $FSig^{(i)}(f_k) \neq FSig^{(i)}(f_m)$.*

$$f_k R_{SI|FUN} f_m \Leftrightarrow \underset{u \in \mathbb{R}^R}{\forall} \; \underset{i=1,\ldots,M}{\exists} \; \underset{t_1 \in (t_0,T]}{\exists} \; \underset{t \in [t_0,t_1]}{\forall} FSig^{(i)}(f_k) \neq FSig^{(i)}(f_m). \tag{2.42}$$

When all the faults are input-strongly functionally isolable, the model is said to be input-strongly functionally diagnosable.

Definition 2.6.3. *Two faults f_k and f_m are input-weakly functionally isolable if there exists an input u and there also exists at least one index i and a finite time $t_1 \in (t_0, T]$ such that for all $t \in [t_0, t_1]$, $FSig^{(i)}(f_k) \neq FSig^{(i)}(f_m)$.*

$$f_k R_{WI|FUN} f_m \Leftrightarrow \underset{u \in \mathbb{R}^R}{\exists} \; \underset{i=1,\ldots,M}{\exists} \; \underset{t_1 \in (t_0,T]}{\exists} \; \underset{t \in [t_0,t_1]}{\forall} FSig^{(i)}(f_k) \neq FSig^{(i)}(f_m). \tag{2.43}$$

When all the faults are input-weakly functionally isolable, the model is said to be input-weakly functionally diagnosable.

If the model is uncontrolled, the definition needs to be modified:

Definition 2.6.4. *Two faults f_k and f_m are functionally isolable if there exists at least one index i and a finite time $t_1 \in (t_0, T]$ such that for all $t \in [t_0, t_1]$, $FSig^{(i)}(f_k) \neq FSig^{(i)}(f_m)$.*

$$f_k R_{I|FUN} f_m \Leftrightarrow \underset{i=1,\ldots,M}{\exists} \; \underset{t_1 \in (t_0,T]}{\exists} \; \underset{t \in [t_0,t_1]}{\forall} FSig^{(i)}(f_k) \neq FSig^{(i)}(f_m). \tag{2.44}$$

When all the faults are functionally isolable, the model is said to be functionally diagnosable.

When stochastic behavior is included in an uncontrolled model, then the following definition can be given:

Definition 2.6.5. *If two faults f_k and f_m act on the same residual j, they are said to be ϵ-functionally isolable if there exists a time interval $[t_1, t_2] \subseteq [t_0, T]$ such that for all $t \in [t_1, t_2]$:*

$$\delta(FSig^j(f_k), FSig^j(f_m)) > \epsilon, \tag{2.45}$$

where δ is a distance in $\mathbb{R}$.

Analogous definitions can be given for controlled models.

Example 2.6.1.

Let us use the internal form of residual r_1 and r_2 given in the example in Section 2.2.2:

$$r_1 = F + \hat{f}_5 - \hat{f}_3 - A_1 \frac{d(L_1 + \hat{f}_6)}{dt} - \alpha_{1,2}(S_{1,2} - \hat{f}_1)\sqrt{2g\left(L_1 + \hat{f}_6 - \left(L_2 + \hat{f}_7\right)\right)}, \tag{2.46}$$

$$\begin{aligned} r_2 = \alpha_{1,2}\left(S_{1,2} - \hat{f}_1\right)\sqrt{2g\left(L_1 + \hat{f}_6 - \left(L_2 + \hat{f}_7\right)\right)} + \\ - \alpha_2(S_2 - \hat{f}_2)\sqrt{2g\left(L_2 + \hat{f}_7\right)} - \hat{f}_4 - A_2 \frac{d(L_2 + \hat{f}_7)}{dt}. \end{aligned} \tag{2.47}$$

The following functional signatures can be obtained from these residuals:

$$\begin{aligned} FSig^1(f_1) &= \hat{f}_1\sqrt{2g\,(L_1 - L_2)} \\ FSig^2(f_1) &= -\hat{f}_1\sqrt{2g\,(L_1 - L_2)} \end{aligned} \tag{2.48}$$

$$\begin{aligned} FSig^1(f_2) &= 0 \\ FSig^2(f_2) &= \hat{f}_2\sqrt{2gL_2} \end{aligned} \tag{2.49}$$

$$\begin{aligned} FSig^1(f_3) &= -\hat{f}_3 \\ FSig^2(f_3) &= 0 \end{aligned} \tag{2.50}$$

$$\begin{aligned} FSig^1(f_4) &= 0 \\ FSig^2(f_4) &= -\hat{f}_4 \end{aligned} \tag{2.51}$$

$$\begin{aligned} FSig^1(f_5) &= \hat{f}_5 \\ FSig^2(f_5) &= 0 \end{aligned} \tag{2.52}$$

$$FSig^1(f_6) = \alpha_{1,2}S_{1,2}\left(\sqrt{2g\,(L_1 - L_2)} - \sqrt{2g\left(L_1 + \hat{f}_6 - L_2\right)}\right) + A_1\frac{d\hat{f}_6}{dt}$$

$$FSig^2(f_6) = -\alpha_{1,2}S_{1,2}\left(\sqrt{2g\,(L_1 - L_2)} - \sqrt{2g\left(L_1 + \hat{f}_6 - L_2\right)}\right) \tag{2.53}$$

$$FSig^1(f_7) = \alpha_{1,2}S_{1,2}\left(\sqrt{2g\,(L_1 - L_2)} - \sqrt{2g\left(L_1 - L_2 - \hat{f}_7\right)}\right)$$

$$FSig^2(f_7) = -\alpha_{1,2}S_{1,2}\left(\sqrt{2g\,(L_1 - L_2)} - \sqrt{2g\left(L_1 - L_2 - \hat{f}_7\right)}\right) + \alpha_2 S_2\left(\sqrt{2g\,(L_2)} - \sqrt{2g\left(L_2 + \hat{f}_7\right)}\right) - A_2\frac{d\hat{f}_7}{dt} \tag{2.54}$$

All faults have different functional signatures. Therefore, they are input-weakly isolable according to Definition 2.6.3. Moreover, during normal operation of the system, i.e., when the input flow $F > 0$, $L_1 > L_2 > 0$, *faults are also input-strongly isolable. If these conditions are not satisfied, then some faults will become undetectable, e.g., when* $L_1 = L_2$ *then* $FSig^1(f_1) = FSig^2(f_1) = 0$ *and, therefore,* f_1 *will be unisolable.*

2.7 Generalization of definitions of fault isolability

The definitions of fault isolability based on the values of the diagnostic signals (residuals) can be combined with the definitions based on the sequences of symptoms. The condition of fault unisolability while using two or more fault–symptoms relations simultaneously may be formed as a conjunction of conditions of unisolability for the individual forms of notation of fault–symptoms relations. However, the condition of isolability is a logic alternative of the corresponding conditions.

One can distinguish two main cases regarding knowledge about a diagnosed system. The first case is when models of fault influence on the system are known, i.e., when the internal form of residuals is known. The second case is more common in practice and relates to the case when only the computational form of residuals is known. For both of these cases, fault unisolability and isolability conditions can be formulated by using two or more forms of notation for the fault–symptoms relation simultaneously (Kościelny *et al.* 2016). In particular, it is appropriate to combine signature-based methods specified by the model values of diagnostic signals (BDM, FIS) with signatures in the form of sequences of symptoms.

Below, generalized definitions of isolability are presented for the case where both forms of residuals, internal and computational, are known.

Definition 2.7.1. *Faults $f_k, f_m \in F$ are unisolable (unconditionally unisolable) based on delay-designed sequences of symptoms and FIS iff the conditions, which are the conjunction of conditions* (2.20) *and* (2.34)*, are met:*

$$f_k R_{U|FIS,DSEQ} f_m \Leftrightarrow [\underset{s_j \in S}{\forall} V_{j,k} = V_{j,m}] \wedge [\underset{s_j,s_p \in S}{\forall} es^d_{j,p}(f_k, \tau_{k,j,p}) = es^d_{j,p}(f_m, \tau_{m,j,p})]. \tag{2.55}$$

Definition 2.7.2. *Faults f_k, $f_m \in F$ are isolable (unconditionally isolable) based on delay-designed sequences of symptoms and FIS iff the conditions, which are the alternative of conditions* (2.22) *and* (2.35)*, are met:*

$$f_k R_{I|FIS,DSEQ} f_m \Leftrightarrow [\underset{s_j \in S}{\exists} V_{j,k} \cap V_{j,m} = \emptyset] \vee [\underset{s_j,s_p \in S}{\exists} es^d_{j,p}(f_k, \tau_{k,j,p}) \neq es^d_{j,p}(f_m, \tau_{m,j,p})]. \tag{2.56}$$

For a similar case, when only residuals in the computational form are known, the following definitions can be given:

Definition 2.7.3. *Faults $f_k, f_m \in F$ are unisolable (unconditionally unisolable) on the basis of FIS and sequences of symptoms iff signatures and sequences of those faults are identical. Thus the conditions* (2.20) *and* (2.31) *are met:*

$$f_k R_{U|FIS,SEQ} f_m \Leftrightarrow [\underset{s_j \in S}{\forall} V_{j,k} = V_{j,m}] \wedge [\underset{s_j,s_p \in S}{\forall} es_{j,p}(f_k) = es_{j,p}(f_m)]. \tag{2.57}$$

Definition 2.7.4. *Faults $f_k, f_m \in F$ are (unconditionally) isolable on the basis of FIS and sequences of symptoms iff there is a diagnostic signal, for which subsets of values corresponding to those faults are disjoint or their elementary sequences of symptoms are different. This means combining conditions* (2.22) *and* (2.32).

$$f_k R_{I|FIS,SEQ} f_m \Leftrightarrow$$

$$[\underset{s_j \in S}{\exists} V_{j,k} \cap V_{j,m} = \emptyset] \vee \{\underset{s_j,s_p \in S}{\exists} [es_{j,p}(f_k) =< s_j, s_p >] \wedge [es_{j,p}(f_m) =< s_p, s_j >]\}. \tag{2.58}$$

With the above definitions, using knowledge (usually incomplete) about sequences of symptoms makes it possible to increase fault isolability in comparison with reasoning only on the basis of signatures, defined by exemplary values of diagnostic signals.

2.8 Summary

The set of formal definitions of fault isolability and unisolability on the basis of binary diagnostic matrix, knowledge of the sequence of symptoms and FIS information system have been presented in this chapter. These faults–symptoms relations are designed based on system modeling, taking into account fault influence and expert knowledge.

A primary method of increasing fault isolability is increasing the number of measured signals. This makes it possible to generate additional residuals. However, it is not always possible or economically justified.

Another possibility is to design secondary residuals. This is promising when the internal form of residuals is known and when models with fault influence are available. There are also some limited possibilities to generate secondary residuals when only the computational form is known. An alternative or complementary method of increasing fault isolability can be the application of multi-valued or continuous residual evaluation instead of a binary one. Using FIS with tri-valued residuals instead of binary is particularly purposeful. This approach can be used in case where system models including fault influence are known, as well as when expert knowledge is available.

An additional possibility to increase fault isolability results from making use of the knowledge of sequences of symptoms. A different order of symptoms, or even the same order, but with a different delay, makes it possible to isolate faults, which are unisolable when analyzing only the values of diagnostic signals. Knowledge of the sequence of symptoms is complementary to the knowledge of the signatures of faults from BDM or FIS and to the knowledge of the directions of faults in the residual space.

Chapter 3

New metric of fault isolability

3.1 Introduction

The metrics of fault isolability presented in previous chapters have substantial limitations. For example, the diagnosability degree does not refer to weak and strong isolability. Both diagnosability degree and isolability index cannot be used with multi-valued diagnostic signals and sequences of symptoms. Diagnosis accuracy is very complex and computationally expensive, and therefore it is cumbersome in some applications, e.g., optimal sensor placement. Consequently, there is a need for a new universal metric of isolability. The following requirements should be essential for this metric:

- It should be possible to use the proposed metric with different types of diagnostic signals, i.e., binary, multi-valued, continuous, and sequences of symptoms.
- The metric should reflect weak and strong isolability.
- The metric of isolability should be applicable to formulating and solving an optimal sensor placement problem.

3.1.1 Bidirectional and unidirectional fault isolability

The exoneration assumption described in Section 1.4.1 states that if for any fault, some of its symptoms are not present, then the fault is exonerated. However, in practice, symptoms usually do not appear simultaneously. Sometimes, when the magnitude of a fault is too small, some symptoms may even not appear at all. This is closely connected with the sensitivity of diagnostic signals to the magnitude of the fault. This can often lead to incorrect diagnoses or unisolable faults.

Example 3.1.1.

Let us analyze the example of a diagnostic system with two weakly isolated faults presented in Tab. 3.1.

Table 3.1: Example of weak isolability.

	f_1	f_2
s_1	1	1
s_2		1

After the occurrence of the fault f_2, there are two possible sequences of symptoms. If the exoneration assumption is taken, then diagnosis would be performed in the following way:

- *The first symptom is $s_1 = 1$. The diagnosis is then $\{f_1\}$, which is incorrect. After some time, the symptom $s_2 = 1$ appears and the diagnosis changes to the correct one: $\{f_2\}$.*
- *The first symptom is $s_2 = 1$. It can only be interpreted as an unknown fault under the exoneration assumption. After the appearance of the second symptom, $s_1 = 1$, the diagnosis points to $\{f_2\}$.*

Without the exoneration assumption some diagnoses are different:

- *After the appearance of $s_1 = 1$ the diagnosis would be $\{f_1, f_2\}$. Then it would be reduced to $\{f_2\}$ after the appearance of $s_2 = 1$.*
- *If $s_2 = 1$ appears first then the diagnosis would be $\{f_2\}$ and it would remain so regardless of the later emergence of $s_1 = 1$.*

After the occurrence of the fault f_1, the only symptom is $s_1 = 1$. With the exoneration assumption, the diagnosis is $\{f_1\}$. This is a more accurate diagnosis than the diagnosis without the exoneration assumption, which is $\{f_1, f_2\}$.

From this example, we can see that without the exoneration assumption diagnosis includes all faults that can explain current symptoms. Therefore, diagnoses can be less precise, but they do not exclude the real fault even in the case when some of its symptoms did not appear.

Strong isolability is required to improve the robustness of a diagnostic system. Two basic definitions were given in Section 1.3.1: bidirectional and unidirectional strong isolability. Let us analyze how those definitions affect the diagnoses with and without the exoneration assumption.

Bidirectional strong isolability

Example 3.1.2.

An extension of the previous example is presented in Tab. 3.2. This is a diagnostic matrix with bidirectional strong isolability with degree 1. However, it is not unidirectionally strongly isolable.

Table 3.2: Example of a BDM with bidirectional strong isolability degree 1.

	f_1	f_2
s_1	1	1
s_2		1
s_3		1

Let us assume that symptoms do not appear simultaneously. After the occurrence of the fault f_2*, with the exoneration assumption, the following scenarios are possible:*

- *If* $s_1 = 1$ *is the first symptom, then:* $\{f_1\} \rightarrow$ *unknown fault* $\rightarrow \{f_2\}$.
- *Otherwise, if the first symptom is* $s_2 = 1$ *or* $s_3 = 1$*: unknown fault* $\rightarrow \{f_2\}$.

Without the exoneration assumption:

- *If* $s_1 = 1$ *is the first symptom, then:* $\{f_1, f_2\} \rightarrow \{f_2\}$.
- *If the first symptom is different, the diagnosis is* $\{f_2\}$ *from the beginning.*

After the occurrence of the fault f_1*, the diagnosis with the exoneration assumption is* $\{f_1\}$*. Without this assumption it is* $\{f_1, f_2\}$*. This is the same as for the previously analyzed, weakly isolating system.*

A new diagnostic signal added to increase the degree of bidirectional strong isolability without introducing unidirectional strong isolability would not affect those results. That is because this additional signal does not affect the set of possible diagnoses.

Without the exoneration assumption, possible scenarios of evolution of diagnoses are identical for the weakly isolating structure and for the bidirectionally strongly isolating diagnostic structure. That is because additional information about lack of some symptoms, available in a bidirectionally strongly isolating diagnostic structure, is ignored without the exoneration assumption. Moreover, for bidirectional fault isolability with the exoneration assumption, there are more possible combinations of symptoms resulting in an unidentified fault, because the number of possible combinations of symptoms increases with the number of diagnostic signals, but the number of signatures of faults remains constant.

Unidirectional strong isolability

Example 3.1.3.
Let us analyze the diagnostic matrix presented in Tab. 3.3. It is an extension of the diagnostic structure presented in Tab. 3.1. The additional diagnostic signal s_3 provides the unidirectional strong isolability property. It is also bidirectionally strongly isolating structure with degree 1.

Table 3.3: Example of a BDM with unidirectional strong isolability.

	f_1	f_2
s_1	1	1
s_2		1
s_3	1	

After the occurrence of the fault f_2, with the exoneration assumption, there is one possible scenario: unknown fault $\rightarrow$ $\{f_2\}$. Without the exoneration assumption, there are two possible scenarios:

- $\{f_1, f_2\} \rightarrow \{f_2\}$ *if* $s_1 = 1$ *is the first symptom,*
- $\{f_2\}$ *if* $s_2 = 1$ *is the first symptom.*

The case of the occurrence of the fault f_1 is similar. After the occurrence of the fault f_1, with the exoneration assumption, there is also only one possible scenario: unknown fault $\rightarrow \{f_1\}$. Without the exoneration assumption, there are two possible scenarios:

- $\{f_1, f_2\} \rightarrow \{f_1\}$ *if the first symptom is* $s_1 = 1$,
- $\{f_1\}$ *if* $s_3 = 1$ *is the first symptom.*

The unidirectional strong isolability makes it possible to reduce false diagnoses when analyzing symptoms with the exoneration assumption. It comes from the definition of the unidirectionally strongly isolating structure (1.3.5), that there is no set of all symptoms of a fault that is also a subset of symptoms of another fault:

$$\nexists (f_k, f_m \in F, k \neq m) \wedge \left(\{v_{j,k} : v_{j,k} \neq 0\} \subseteq \{v_{j,m} : v_{j,m} \neq 0\}\right) . \tag{3.1}$$

Under the exoneration assumption, all symptoms need to appear in order to indicate that a particular fault occurred. Therefore, regardless of the order of appearance of symptoms, there are no false diagnoses.

When comparing bi- and unidirectional strong isolability one can see that unidirectional strong isolability provides qualitative improvement of diagnoses. The bidirectional strong isolability by itself does not improve the precision of final diagnoses. It provides redundancy in diagnostic signals, improving robustness of a diagnostic inference. Although bidirectional strong isolability is desired, it does not improve isolability properties. It is especially visible in the case of the diagnostic reasoning without the exoneration assumption. Therefore it is more convenient for practical applications, due to natural dynamic properties of diagnostic signals.

Therefore, this work is focused on developing isolability metrics that particularly emphasize unidirectional strong isolability.

3.1.2 General qualitative definitions of isolability

The definitions of weak and unidirectional strong isolability given by Gertler are applicable only for binary diagnostic signals. Therefore, there is a need to formulate more general qualitative definitions.

Definition 3.1.1. *A fault f_k is excluded by a signature ϕ if the fault f_k cannot explain the signature ϕ.*

Using this definition, we generalize the Gertler's definitions of isolability recalled in Section 1.3.1. The following general definitions can be formulated:

Definition 3.1.2. *Faults $f_k, f_m \in F$ are weakly isolated iff each alternative fault signature $\phi(f_k)$ excludes the fault f_m, or each alternative fault signature $\phi(f_m)$ excludes the fault f_k.*

Definition 3.1.3. *Faults $f_k, f_m \in F$ are unidirectionally strongly isolated iff each alternative fault signature $\phi(f_k)$ excludes the fault f_m, and each alternative fault signature $\phi(f_m)$ excludes the fault f_k.*

Definition 3.1.4. *Faults $f_k, f_m \in F$ are conditionally weakly isolated iff they are not weakly isolated and there exists an alternative fault signature $\phi_i(f_k)$ that excludes the fault f_m, or there exists an alternative fault signature $\phi(f_m)$ that excludes the fault f_k.*

Definition 3.1.5. *Faults $f_k, f_m \in F$ are conditionally unidirectionally strongly isolated iff they are not strongly isolated, and there exists an alternative fault signature $\phi_i(f_k)$ that excludes the fault f_m, and there exists an alternative fault signature $\phi(f_m)$ that excludes the fault f_k.*

3.2 Proposition of a new measure of isolability

The definitions given in the previous section provide a background for a qualitative evaluation of isolability. In the given context, a quantitative approach should be introduced in order to propose a new measure of isolability. The following algorithm is proposed for calculating this measure:

1. For every ordered pair of faults (f_k, f_m) calculate the value which characterizes how often alternative signatures of the fault f_k exclude the fault f_m. Normalize this value to the range $[0, 1]$.

2. Calculate the sum of values obtained in the previous step.

3. Normalize the obtained result by a factor $\frac{1}{K(K-1)}$, where $K = card(F)$. Thanks to that, the final value of measure of isolability does not depend on the number of faults.

3.2.1 Binary diagnostic signals

An implementation of the measure of isolability for a BDM was proposed in (Rostek 2014). The value of this measure is calculated in two steps:

1. Calculate the value of the following discrete function for all possible ordered pairs of faults:
$$D\colon F \times F \to \{0, 1\}, \tag{3.2}$$
where: F is the set of faults and $f_k \in F,\ i = 1\ \dots\ K$ are particular faults. It is assumed that the value $D(f_k, f_m) = 1$ when the appearance of all symptoms of the fault f_k excludes the fault f_m. If this is not true, then $D(f_k, f_m) = 0$.

2. Calculate the value of the measure as:
$$\psi = \frac{1}{(K-1)K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m \neq k}}^{K} D(f_k, f_m). \tag{3.3}$$

This definition makes it possible to distinguish unidirectional strong isolability from weak isolability. If $D(f_k, f_m) = 1 \ \vee\ D(f_m, f_k) = 1$ then the signature of the fault f_k excludes the fault f_m or the signature of the fault f_m excludes the fault f_k. Therefore, according to Definition 3.1.2, the faults are weakly isolable. Moreover, if $D(f_k, f_m) = 1 \ \wedge\ D(f_m, f_k) = 1$, then the signature of the fault f_k excludes the fault f_m and the signature of the fault f_m excludes the fault f_k. Thus, faults f_k and f_m are unidirectionally strongly isolable as defined in 3.1.3.

Example 3.2.1.
Let us analyze the example of BDM presented in Tab. 3.4. The results of the first step of calculation (3.2) *are presented in Tab. 3.5. The value of the measure of isolability, after applying the formula* (3.3)*, is equal to* 0.82.

The value of the presented measure of isolability can be interpreted as a mean fraction of all diagnoses that can be excluded, after the occurrence of a single fault. The measure of isolability takes the maximal value when all pairs of faults are unidirectionally strongly isolable. In such a case, each single fault signature excludes $(K - 1)$ other faults (the fault does not exclude itself).

Table 3.4: Example of BMD.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
s_1	1	1	1		1			
s_2		1	1	1	1			
s_3				1	1	1		1
s_4						1	1	1
s_5							1	1

Table 3.5: Values of $D(f_k, f_m)$ for the example of BDM presented in Tab. 3.4.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
f_1				1		1	1	1
f_2	1			1		1	1	1
f_3	1			1		1	1	1
f_4	1	1	1			1	1	1
f_5	1	1	1	1		1	1	1
f_6	1	1	1	1	1		1	
f_7	1	1	1	1	1	1		
f_8	1	1	1	1	1	1	1	

Then $\sum_{k=1}^{K}\sum_{\substack{m=1\\m\neq k}}^{K} D\left(f_k, f_m\right) = (K-1)\,K$ and the value of the measure of isolability is equal to $\frac{(K-1)K}{(K-1)K} = 1$.

This formula ((3.2) and (3.3)) can be used for any approach to fault isolation that is based on binary diagnostic signals. The only necessary requirement for using the proposed metric is implementation of a method of determining whether a signature excludes other faults or not.

3.2.2 Multi-valued diagnostic signals

In the case of multi-valued diagnostic signals, the conditional isolability metric was proposed in (Rostek 2016). The first step of calculation of the value of the proposed metric (3.2) needs to be slightly modified in order to take into account conditional isolability. Instead of assigning exclusively values 0 or 1 to each ordered pair of faults, the $D\left(f_k, f_m\right)$ can take any value from the range [0, 1]. Let $D\left(f_k, f_m\right)$:

$$D\left(f_k, f_m\right) = \frac{card\left(\{\boldsymbol{\phi} : \boldsymbol{\phi} \in \Phi(f_k) \wedge \boldsymbol{\phi} \text{ excludes } f_m\}\right)}{card\left(\Phi(f_k)\right)}, \tag{3.4}$$

where: $\Phi(f_k)$ is the set of all alternative signatures of the fault f_k.

The formula (3.4) generalize the definition (3.2). It can be understood as a fraction of all alternative signatures of f_k that exclude f_m. In the case of binary diagnostic signals, there is always only one alternative signature $\boldsymbol{\phi}(f_k)$. The value of $D\left(f_k, f_m\right)$ is then equal to 0 or 1. Consequently, in the case of binary diagnostic signals, the formula (3.4) is equivalent to (3.2).

Example 3.2.2.

An example of the FIS system is presented in Tab. 3.6. Results of the first step of calculation of the measure of isolability for this FIS are shown in Tab. 3.7. The resulting value of the measure of isolability ψ *is equal to* $\frac{1}{(K-1)K}\sum_{k=1}^{K}\sum_{\substack{m=1\\m\neq k}}^{K} D\left(f_k, f_m\right) = \frac{2}{3}$.

3.2.3 Continuous diagnostic signals

In the case of continuous diagnostic signals, the definition of the measure of isolability is similar to the definition for multi-valued diagnostic signals (3.4). The only difference is in calculation

Table 3.6: Example of FIS.

	f_1	f_2	f_3	V_i
s_1	-1	$-1, +1$	$+1$	$-1, 0, +1$
s_2	0	$0, 1$	$1, 2$	$0, 1, 2$
s_3	$+1$	$+1$	$-1, +1$	$-1, 0, +1$

Table 3.7: Values of $D(f_k, f_m)$ for the FIS presented in Tab. 3.6.

	f_1	f_2	f_3
f_1	0	0	1
f_2	3/4	0	2/4
f_3	1	3/4	0

of the value $D(f_k, f_m)$. It can be defined as:

$$D(f_k, f_m) = \frac{\text{volume of}\,(\Omega(f_k) \setminus \Omega(f_m))}{\text{volume of}\,\Omega(f_k)}, \tag{3.5}$$

where: $\Omega(f_k)$ denotes the area in residual space characteristic to the fault f_k.

Such a definition makes it possible to properly determine isolability as shown schematically in Fig. 3.1. In this case, for faults f_1 and f_2, the values of both r_1 and r_2 are overlapping. When

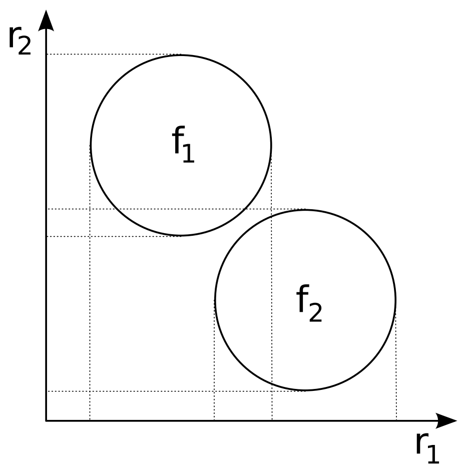

Figure 3.1: Two faults isolable in residual space.

analyzing only the values of single residuals, these faults are conditionally isolable. However, these two faults are strongly isolable in residual space.

3.2.4 Sequences of symptoms

To calculate the measure of isolability for a diagnostic inference approach with sequential symptoms, the binary diagnostic structure can be transformed into a multi-valued (FIS) structure.

The transformation procedure is as follows:

1. For each unique elementary sequence $es_{j,p}$ assign a new diagnostic signal.
2. For each fault, determine the set of values of this diagnostic signal:
   - $\{0\}$ – if both diagnostic signals in the elementary sequence $es_{j,p}$ are not sensitive to this particular fault,
   - $\{-1, +1\}$ – if both diagnostic signals are sensitive to this particular fault, but the order of appearance of symptoms of s_j and s_p is not determined,
   - $\{-1\}$ or $\{+1\}$ – There are two cases:
     (a) both diagnostic signals are sensitive to the given fault and the order of the appearance of symptoms is determined. The values $+1$ and -1 can be applied in any way. The only requirement is to apply them consequently. If for one fault the sequence $es_{j,p} =< s_j, s_p >$ was assigned the value $+1$, then the same sequence for other faults should also have assigned the value $+1$. Consequently, the the value -1 should be assigned to the reversed sequence $es_{j,p} =< s_p, s_j >$.
     (b) only one diagnostic signal is sensitive to the given fault. The value $+1$ or -1 should be applied as if the sensitive signal was first in the sequence.

The resulting multi-valued structure has equivalent isolability properties to the original structure with sequential symptoms. If exclusion properties of any pair of faults are identical for both original and transformed diagnostic structures then the transformations proposed above make it possible to properly reflect isolability properties of the original diagnostic structure without the exoneration assumption. Fault f_k excludes fault f_m on the basis of sequences of symptoms if $es_{j,p}(f_k) \neq es_{j,p}(f_m)$, or if $es_{j,p}(f_m) =< s_j, s_p >$ and only s_p is sensitive to f_k. After the transformation described above, the value of equivalent, multi-valued diagnostic signal is different for both faults. Moreover, if an elementary symptom sequence is determined for only one fault, e.g., $es_{j,p}(f_k) =< s_j, s_p >$, but both s_j and s_p are sensitive to f_k and f_m then f_m conditionally excludes f_k. After the transformation, the values of the new, multi-valued diagnostic signal are $\{+1\}$ for f_k and $\{-1, +1\}$ for f_m. In all these cases exclusion properties are maintained after transformation.

Example 3.2.3.

Let us analyze the example of a diagnostic system presented in Tab. 3.8. In this example, there are five faults and two diagnostic signals.

Table 3.8: Example of sequential residuals.

	f_1	f_2	f_3	f_4	f_5
s_1	1	1	1	1	
s_2	1	1	1		1
$es_{1,2}$	$< s_1, s_2 >$	$< s_2, s_1 >$		$< s_1, - >$	$< s_2, - >$

It is not possible to isolate faults f_1, f_2 and f_3 in this structure using only the values of diagnostic signals. The value of the proposed measure of isolability for this BDM is $\psi = 0.40$. However, from Tab. 3.8, elementary sequences of symptoms of two faults are known. For the fault f_1, the diagnostic signal s_1 always appears before the signal s_2. For the fault f_2, the order is reversed and the signal s_2 appears before s_1. For the other faults, the order of symptoms is not determined.

To calculate the value of the metric of isolability a new diagnostic signal, s_3, needs to be added. The value $+1$ of the signal s_3, corresponds to the elementary sequence $es_{1,2} =< s_1, s_2 >$. Similarly, the value -1 corresponds to the elementary sequence $es_{1,2} =< s_2, s_1 >$. The FIS that was derived with this method is presented in Tab. 3.9.

The faults f_1 and f_2 are unidirectionally strongly isolable, and $D(f_1, f_2) = D(f_2, f_1) = 1$. Similarly, pairs of faults $\{f_1, f_5\}$ and $\{f_2, f_4\}$ are also unidirectionally strongly isolable. Pairs $\{f_1, f_3\}$ and $\{f_2, f_3\}$ are conditionally weakly isolable. Pairs $\{f_1, f_4\}$, $\{f_1, f_5\}$, $\{f_3, f_4\}$, and

Table 3.9: FIS derived from the example of sequential residuals.

	f_1	f_2	f_3	f_4	f_5
s_1	$+1$	$+1$	$+1$	$+1$	
s_2	$+1$	$+1$	$+1$		$+1$
s_3	$+1$	-1	$-1, +1$	$+1$	-1

$\{f_3, f_5\}$ *are unconditionally weakly isolable.*

The value of the measure of isolability of this FIS is equal to $\psi = \frac{1}{(K-1)K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m \neq k}}^{K} D(f_k, f_m) = \frac{13}{20} = 0.65$. *Consequently, for the original diagnostic system with sequential residuals, the value of the measure of isolability is also equal to* 0.65.

3.3 Properties of the proposed metric

3.3.1 Examples of elementary types of diagnostic structures

Unisolable diagnostic systems

Example 3.3.1.

An example of a diagnostic structure with three unisolable faults is presented in Tab. 3.10.

Table 3.10: Example of an unisolable diagnostic structure.

	f_1	f_2	f_3
s_1	1	1	1

In this example, all faults are detectable, but no fault signature excludes any other fault. For each pair of faults (f_k, f_m) *in this system* $D(f_k, f_m) = 0$. *Therefore:*

$$\psi = \frac{1}{(K-1)K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m \neq k}}^{K} D(f_k, f_m) = 0 \tag{3.6}$$

This result does not depend on K and $\psi = 0$ for any diagnostic system that is not able to isolate faults.

Weakly isolating diagnostic structure

Example 3.3.2.

In Tab. 3.11 an example of a weakly isolable diagnostic structure is presented.

Table 3.11: Example of a weakly isolating diagnostic structure.

	f_1	f_2	f_3
s_1	1	1	1
s_2		1	1
s_3			1

Each pair of faults is weakly isolable from others. Therefore, for each pair of faults $(f_k, f_m;\ k \neq m)$: $D(f_k, f_m) + D(f_m, f_k) = 1$. Then:

$$\psi = \frac{1}{(K-1)K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m \neq k}}^{K} D(f_k, f_m) = \frac{\frac{1}{2}(K-1)K}{(K-1)K} = \frac{1}{2} \tag{3.7}$$

Generally, the value of the isolability measure is equal to $1/2$ for each diagnostic structure in which each pair of faults is weakly isolable.

Unidirectionally strongly isolating diagnostic structure

Example 3.3.3.

An example of a unidirectionally strongly isolating diagnostic structure is presented in Tab. 3.12. For each pair of faults in this structure both $D(f_k, f_m) = 1$ and $D(f_m, f_k) = 1$. Then:

$$\psi = \frac{1}{(K-1)K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m \neq k}}^{K} D(f_k, f_m) = \frac{(K-1)K}{(K-1)K} = 1. \tag{3.8}$$

The proposed measure of isolability distinguishes between weak and unidirectional strong isolability. In the case of the diagnosability degree that was discussed in Section 1.6.1, the

Table 3.12: Example of a unidirectionally strongly isolating diagnostic structure.

	f_1	f_2	f_3
s_1	1		
s_2		1	
s_3			1

diagnostic structures that are presented in Tables 3.11 and 3.12 have equal diagnosability degrees.

3.4 Summary

In this chapter, the new metric of fault isolability was proposed. It meets the requirements for a general metric of isolability.

This metric is useful for different types of diagnostic structures, including those based on binary, multi-valued and continuous diagnostic signals. It can also be calculated for fault isolation approaches based on using sequences of symptoms.

The metric is calculated based on using ordered pairs of faults. Therefore, it is sensitive to whether f_k is isolable from f_m or f_m is isolable from f_k. Consequently, weak and strong unidirectional isolability influence the value of the metric.

The proposed metric of isolability is also useful in formulation of linear optimal sensor placement problems.

Metrics of fault isolability can be used in the design phase of a diagnostic system and as a performance index in solving optimal sensor placement problems. In addition, the availability of sensors changes during the operation of a industrial process, e.g., due to fault occurrences or ongoing maintenance. Metrics of fault isolability can be used in on-line monitoring in diagnostic systems as an indicator of current fault isolation capabilities.

Chapter 4

Comparison of isolability metrics

4.1 Introduction

In this chapter, the measure ψ of isolability proposed in this thesis will be compared with metrics of isolability discussed in Chapter 1. The example of a two tank system, introduced in Section 2.2.2, will be used to demonstrate their similarities and differences. Different types of diagnostic signals will be analyzed. The impact of the exoneration assumption on values of metrics of fault isolability will be also discussed.

4.2 Binary diagnostic matrix

Tab. 4.1 recalls the binary diagnostic matrix of the two tank system from Fig. 2.1. Only the primary residuals are taken into account.

Table 4.1: BDM for the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
s_1	1		1		1	1	1
s_2	1	1		1		1	1
s_3		1	1	1	1	1	1

4.2.1 Diagnosability degree

In order to calculate the diagnosability degree, faults should be divided into disjoint subsets of unisolable faults called D-classes. The following D-classes can be determined from Tab. 4.1 and applying Definition 2.2.2: $\{f_1\}$, $\{f_2, f_4\}$, $\{f_3, f_5\}$, $\{f_6, f_7\}$. Therefore, the diagnosability degree is equal to $\frac{4}{7} \approx 0.57$

4.2.2 Diagnosis accuracy

A calculation of diagnosis accuracy requires determination of possible diagnoses. The diagnosis accuracy can be calculated for two cases.

1. With the exoneration assumption, there are four possible diagnoses: $\{f_1\}, \{f_2, f_4\}, \{f_3, f_5\}$, and $\{f_6, f_7\}$. The diagnosis accuracy is then equal to: $(\frac{7}{4})^{-1} \approx 0.57$.
2. Without the exoneration assumption, there are also four diagnoses: $\{f_1, f_6, f_7\}, \{f_2, f_4, f_6, f_7\}$, $\{f_3, f_5, f_6, f_7\}$, and $\{f_6, f_7\}$. The diagnosis accuracy is equal to $(\frac{13}{4})^{-1} \approx 0.31$.

4.2.3 Isolability Index

To find the value of isolability index, pairs of isolable faults need to be determined. Isolability index, similarly to diagnosis accuracy, can be calculated in two variants:

1. With the exoneration assumption, isolability index is equal to 36, because there are 36 ordered pairs of isolable faults. They are listed in Tab. 4.2a.
2. Without the exoneration assumption there are 26 such pairs. They are listed in Tab. 4.2b.

4.2.4 Measure of isolability

The first step of calculation of the measure of isolability is to determine values of $D(f_k, f_m)$. The results are presented in Table 4.3. Afterwards, using these values, the value of the measure of isolability is calculated: $\psi = \frac{1}{(K-1)K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m\neq k}}^{K} D(f_k, f_m) = \frac{26}{42} \approx 0.62$.

This result means that after the occurrence of a single fault, on average 62% of all possible faults are rejected from diagnosis.

Table 4.2: Ordered pairs of faults that are isolable using BDM for the two tank system:

(a) with the exoneration assumption,

f_k	faults isolable from f_k
f_1	$f_2, f_3, f_4, f_5, f_6, f_7$
f_2	f_1, f_3, f_5, f_6, f_7
f_3	f_1, f_2, f_4, f_6, f_7
f_4	f_1, f_3, f_5, f_6, f_7
f_5	f_1, f_2, f_4, f_6, f_7
f_6	f_1, f_2, f_3, f_4, f_5
f_7	f_1, f_2, f_3, f_4, f_5

(b) without the exoneration assumption.

f_k	faults isolable from f_k
f_1	f_2, f_3, f_4, f_5
f_2	f_1, f_3, f_5
f_3	f_1, f_2, f_4
f_4	f_1, f_3, f_5
f_5	f_1, f_2, f_4
f_6	f_1, f_2, f_3, f_4, f_5
f_7	f_1, f_2, f_3, f_4, f_5

Table 4.3: Values $D(f_k, f_m)$ for the BDM for the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
f_1		1	1	1	1		
f_2	1		1		1		
f_3	1	1		1			
f_4	1		1		1		
f_5	1	1		1			
f_6	1	1	1	1	1		
f_7	1	1	1	1	1		

4.3 Fault Information System

Tab. 4.4 recalls FIS for the two tank system. The FIS uses multi-valued diagnostic signals. Positive and negative values represent directions of the change of values of diagnostic signals.

Table 4.4: Example of the FIS for the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
s_1	$+1$	0	-1	0	$-1, +1$	$-1, +1$	$-1, +1$
s_2	-1	$+1$	0	-1	0	$-1, +1$	$-1, +1$
s_3	0	$+1$	-1	-1	$-1, +1$	$-1, +1$	$-1, +1$

4.3.1 Diagnosability degree and isolability index

In this example, some pairs of faults are only conditionally isolable, e.g., some alternative signatures of the fault f_6 exclude the fault f_3, and some do not. Therefore, it is impossible to divide faults into disjoint D-classes. Consequently, the diagnosability degree and isolability index cannot be calculated.

4.3.2 Diagnosis accuracy

The diagnosis accuracy should once more be analyzed in two variants:

1. With the exoneration assumption, there are six possible diagnoses: $\{f_1\}$, $\{f_2\}$, $\{f_3, f_5\}$, $\{f_4\}$, $\{f_5\}$, $\{f_6, f_7\}$. The diagnosis accuracy is equal to: $\left(\frac{8}{6}\right)^{-1} = 0.75$.

2. Without the exoneration assumption, there are also six diagnoses: $\{f_1, f_6, f_7\}$, $\{f_2, f_6, f_7\}$, $\{f_3, f_5, f_6, f_7\}$, $\{f_4, f_6, f_7\}$, $\{f_5, f_6, f_7\}$, $\{f_6, f_7\}$. The value of diagnostic accuracy is $\left(\frac{18}{6}\right)^{-1} \approx 0.33$

The diagnosis accuracy can be used with multi-valued diagnostic signals. However, the same weights are given to all possible diagnoses. In the above example, three out of four alternative signatures of f_5 are isolable from the fault f_3. However, it is not reflected in the value of diagnosis accuracy. Moreover, the diagnoses $\{f_5\}$ and $\{f_3, f_5\}$ have the same weights.

4.3.3 Measure of isolability

The results of the first step of calculation of $D(f_k, f_m)$ are presented in Tab. 4.5. The measure

Table 4.5: The first step of calculation of the measure of isolability for the FIS for the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
f_1		1	1	1	1		
f_2	1		1	1	1		
f_3	1	1		1			
f_4	1	1	1		1		
f_5	1	1	3/4	1			
f_6	1	1	1	1	1		
f_7	1	1	1	1	1		

of isolability for this system ψ is equal to $\frac{28.75}{42} \approx 0.685$. This indicates that the FIS structure has better isolating properties than the BDM structure.

4.4 Sequential residuals and BDM

Measures of isolability are applicable for diagnostic systems which make use of the knowledge of sequences of symptoms. For the example of the two tank system, the sequencesof symptoms are presented in Tab. 4.6.

4.4.1 Diagnosability degree

In case of sequences of symptoms, faults may be conditionally isolable, regardless of the type of diagnostic signals. Generally, diagnosability degree and isolability index are not suitable for this class of diagnostic systems. However, in the studied example, there is no conditionally

Table 4.6: Sequences of symptoms for the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
s_1	1	–	1	–	1	1	1
s_2	1	1	–	1	–	1	1
s_3	–	1	1	1	1	1	1
$es_{1,2}$	$< s_1, s_2 >$					$< s_1, s_2 >$	$< s_2, s_1 >$
$es_{1,3}$			$< s_1, s_3 >$		$< s_1, s_3 >$	$< s_1, s_3 >$	$< s_3, s_1 >$

isolable pair of faults. There are five D-classes of faults: $\{f_1\}$, $\{f_2, f_4\}$, $\{f_3, f_5\}$, $\{f_6\}$, $\{f_7\}$. The diagnosability degree is equal to $\frac{5}{7} \approx 0.71$.

4.4.2 Diagnosis accuracy

1. With the exoneration assumption, there are five possible diagnoses after the occurrence of a single fault: $\{f_1\}$, $\{f_2, f_4\}$, $\{f_3, f_5\}$, $\{f_6\}$, $\{f_7\}$. The value of this measure is: $\left(\frac{7}{5}\right)^{-1} \approx 0.71$.

2. Without the exoneration assumption, diagnoses are: $\{f_1, f_6\}$, $\{f_2, f_4, f_7\}$, $\{f_3, f_5, f_6\}$, $\{f_6\}$, $\{f_7\}$. The diagnosis accuracy is equal to: $\left(\frac{10}{5}\right)^{-1} = 0.5$.

4.4.3 Isolability index

The ordered pairs of isolable faults used in calculation of the isolability index are listed in Tab. 4.7. The value of the isolability index is equal to 38 with the exoneration assumption. Without this assumption, the index is equal to 33.

4.4.4 Measure of isolability

In order to calculate the measure of isolability, the method presented in Section 3.2.4 should be used. The sequences of symptoms $es_{1,2}$ and $es_{1,3}$ from Tab. 4.6 are replaced with multi-valued diagnostic signals $s_{1,2}$ and $s_{1,3}$ respectively. Tab. 4.8 combines both the original binary

Table 4.7: Ordered pairs of faults that are isolable using the BDM and sequences of symptoms for the two tank system.

(a) With the exoneration assumption.

f_k	faults isolable from f_k
f_1	$f_2, f_3, f_4, f_5, f_6, f_7$
f_2	f_1, f_3, f_5, f_6, f_7
f_3	f_1, f_2, f_4, f_6, f_7
f_4	f_1, f_3, f_5, f_6, f_7
f_5	f_1, f_2, f_4, f_6, f_7
f_6	$f_1, f_2, f_3, f_4, f_5, f_7$
f_7	$f_1, f_2, f_3, f_4, f_5, f_6$

(b) Without the exoneration assumption.

f_k	faults isolable from f_k
f_1	f_2, f_3, f_4, f_5, f_7
f_2	f_1, f_3, f_5, f_6
f_3	f_1, f_2, f_4, f_7
f_4	f_1, f_3, f_5, f_6
f_5	f_1, f_2, f_4, f_7
f_6	$f_1, f_2, f_3, f_4, f_5, f_7$
f_7	$f_1, f_2, f_3, f_4, f_5, f_6$

diagnostic signals and the new multi-valued signals.

Table 4.8: BDM for the two tank system with multi-valued diagnostic signals obtained from known sequences of symptoms.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
s_1	1	–	1	–	1	1	1
s_2	1	1	–	1	–	1	1
s_3	–	1	1	1	1	1	1
$s_{1,2}$	+1	−1	+1	−1	+1	+1	−1
$s_{1,3}$	+1	−1	+1	−1	+1	+1	−1

Results of the first step of calculation are presented in Tab. 4.9. The measure of isolability can be calculated using information from Tab. 4.9:

$$\psi = \frac{1}{(K-1)\,K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m\neq k}}^{K} D\left(f_k, f_m\right) = \frac{33}{42} \approx 0.79.$$

Table 4.9: Values of $D(f_k, f_m)$ for sequential residuals and the BDM for the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
f_1		1	1	1	1		1
f_2	1		1		1	1	
f_3	1	1		1			1
f_4	1		1		1	1	
f_5	1	1		1			1
f_6	1	1	1	1	1		1
f_7	1	1	1	1	1	1	

4.5 Sequential residuals and FIS

for isolation purposes, the sequences of symptoms can be also combined with multi-valued diagnostic signals. Tab. 4.10 presents such a structure on the example of the two tank system. Similarly to the previous example, the elementary sequences of symptoms were replaced with multi-valued diagnostic signals.

Table 4.10: FIS of the two tank system with information about sequences of symptoms.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
s_1	+1	0	−1	0	−1, +1	−1, +1	−1, +1
s_2	−1	+1	0	−1	0	−1, +1	−1, +1
s_3	0	+1	−1	−1	−1, +1	−1, +1	−1, +1
$s_{1,2}$	+1	−1	+1	−1	+1	+1	−1
$s_{1,3}$	+1	−1	+1	−1	+1	+1	−1

4.5.1 Diagnosability degree and isolability index

The diagnosability degree and isolability index cannot be calculated because some faults are conditionally isolable.

4.5.2 Diagnosis accuracy

1. With the exoneration assumption, there are seven diagnoses: $\{f_1\}$, $\{f_2\}$, $\{f_3, f_5\}$, $\{f_4\}$, $\{f_5\}$, $\{f_6\}$, $\{f_7\}$. The value of the diagnosis accuracy is equal to: $\left(\frac{8}{7}\right)^{-1} = 0.875$

2. Without the exoneration assumption, there are the following diagnoses: $\{f_1, f_6\}$, $\{f_2, f_7\}$, $\{f_3, f_5, f_6\}$, $\{f_4, f_7\}$, $\{f_5, f_6\}$, $\{f_6\}$, $\{f_7\}$, and the diagnosis accuracy is $\left(\frac{13}{7}\right)^{-1} \approx 0.54$

4.5.3 Measure of isolability

Tab. 4.11 lists the values of $D(f_k, f_m)$ needed for calculation of the measure of isolability. Using

Table 4.11: Values of $D(f_k, f_m)$ for sequential residuals and the FIS for the two tank system.

	f_1	f_2	f_3	f_4	f_5	f_6	f_7
f_1		1	1	1	1		1
f_2	1		1	1	1	1	
f_3	1	1		1			1
f_4	1	1	1		1	1	
f_5	1	1	3/4	1			1
f_6	1	1	1	1	1		1
f_7	1	1	1	1	1	1	

Tab. 4.11, the value of the measure of isolability ψ can be calculated from Tab. 4.11:

$$\psi = \frac{1}{(K-1)K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m \neq k}}^{K} D(f_k, f_m) = \frac{35.75}{42} \approx 0.85.$$

4.6 Comparison and summary

Tab. 4.12 summarizes the values of different metrics of isolability for different types of diagnostic signals in case of the two tank system.

Table 4.12: Comparison of different metrics of isolability.

			Sequential residuals:	
	BDM	FIS	+ BDM	+ FIS
Diagnosability degree	0.57	–	0.71	–
Diagnosis accuracy with exoneration	0.57	0.75	0.71	0.875
Diagnosis accuracy without exoneration	0.31	0.33	0.5	0.54
Isolability index with exoneration	36	–	38	–
Isolability index without exoneration	26	–	33	–
Measure of isolability ψ	0.62	0.685	0.79	0.85

Values of different measures of isolability have different interpretations and cannot be directly compared with each other.

It is interesting to notice that, for the diagnosis accuracy with the exoneration assumption, the value of the metric for FIS is greater than for BDM with sequential residuals. Without this assumption, BDM with sequential residuals has a higher score. Therefore, when using the diagnosis accuracy during the design of a diagnostic system, it should be carefully analyzed whether the exoneration assumption is appropriate, because it can significantly affect the designing process. The exoneration assumption was thoroughly analyzed in Section 3.1.1.

In all studied cases, FIS with sequential residuals is the diagnostic structure with the best isolability properties. It shows that it is very beneficial to extend binary diagnostic signals and introduce a multi-valued evaluation of residuals and the sequences of symptoms.

It is not possible to differentiate between weak and unidirectional strong isolability using the diagnosability degree. All other metrics include this information.

Only diagnosis accuracy and the proposed measure of isolability can be used for multi-valued diagnostic signals and sequences of symptoms. However, the calculation of diagnosis accuracy requires additional assumptions, e.g., with or without the exoneration assumption, or whether or not diagnoses caused by multiple faults are included. Moreover, in case of conditional isolability, diagnosis accuracy considers both possibilities (isolable or unisolable) with the same weight, regardless of the actual distribution of alternative signatures into these two categories.

The isolability index and the new measure of isolability make it possible to formulate the optimal sensor placement problem as a binary integer linear programming problem, which can be efficiently solved using existing solvers.

Chapter 5

Optimal sensor placement problem

5.1 Introduction

Using the proposed measure of isolability (3.3), it is possible to construct an optimization problem for searching for the set of sensors providing the best isolability:

$$\begin{aligned} &\underset{x}{\text{maximize}} \quad \frac{1}{(K-1)K}\sum_{k=1}^{K}\sum_{\substack{m=1\\m\neq k}}^{K} D(f_k, f_m), \\ &x_i \in \{0, 1\} \end{aligned} \tag{5.1}$$

where: x_i is a decision variable and $x_i = 1$ when the i^{th} sensor is chosen. If the set of faults is constant, then the normalization factor $\frac{1}{(K-1)K}$ does not change the results of optimization. Consequently, this normalization only needs to be used when new faults, related to new sensors, are considered. However, then the optimal sensor placement problem becomes much more difficult to solve. Therefore, in this chapter, a constant number of faults is assumed.

5.2 Problem formulation for BDM

In this section, only binary diagnostic signals are analyzed. If a fault f_k is unisolable from f_m, then $D(f_k, f_m)$ is equal to 0. Otherwise, it is equal to 1. The value of $D(f_k, f_m)$ can be calculated in the following way:

$$x_{D_{k,m}} = D(f_k, f_m) = \underset{x_{s_j}}{max} \left\{ x_{s_j} : v_{j,k} \neq 0 \wedge v_{j,m} = 0 \right\}. \tag{5.2}$$

This formula states that $D(f_k, f_m)$ is equal to 1 if at least one diagnostic signal s_j is sensitive to the fault f_k and not sensitive to the fault f_m. The shorthand notation $x_{D_{k,m}}$ will be used instead of $D(f_k, f_m)$ as a variable in the description of an optimal sensor placement problem.

Similarly, the variable x_{s_j} can be expressed as:

$$0 \leq x_{s_j} \leq \underset{x_i}{min} \left\{ x_i : x_i \text{ is necessary to calculate } s_j \right\}, \tag{5.3}$$

where: x_i is the decision variable, which indicates that i^{th} sensor is available, as in the original problem formulation (5.1). If even one of the sensors necessary for the diagnostic signal s_j is unavailable, then this signal cannot be used. The inequality relation ≤ is used, because, even if all required sensors are available, the diagnostic signal may be not of interest, e.g., due to too high cost of development of necessary models.

Example 5.2.1.

In Tab. 5.1 an example of a simple BDM is presented. There are three faults and three diagnostic signals. Each diagnostic signal requires two sensors to be available.

Table 5.1: Simple BDM and sensor requirements for diagnostic signals.

	f_1	f_2	f_3
$s_1(x_1, x_2)$	1	1	1
$s_2(x_1, x_3)$		1	1
$s_3(x_2, x_3)$			1

The following equations can be constructed:

$$
\begin{aligned}
x_{s_1} &\leq min\{x_1, x_2\},\\
x_{s_2} &\leq min\{x_1, x_3\},\\
x_{s_3} &\leq min\{x_2, x_3\},\\
x_{D_{2,1}} &= D(f_2, f_1) = max\{x_{s_2}\} = x_{s_2},\\
x_{D_{3,1}} &= D(f_3, f_1) = max\{x_{s_2}, x_{s_3}\},\\
x_{D_{3,2}} &= D(f_3, f_2) = max\{x_{s_3}\} = x_{s_3},\\
x_{s_1}, x_{s_2}, x_{s_3} &\geq 0,\\
x_i, s_{s_j} &\in \{0, 1\}, \quad i, j = 1 \ldots 3.
\end{aligned}
\tag{5.4}
$$

The pairs of faults for which $D(f_k, f_m) = 0$ *were omitted.*

The objective function $\underset{x}{maximize}\frac{1}{6}\sum_{k=1}^{K}\sum_{\substack{m=1\\m\neq k}}^{K} x_{D_{k,m}}$ *with constraints* (5.4) *is a difficult, constrained, non-linear optimisation problem.*

5.2.1 Additional constraints

Fault detectability

Generally, it is not possible to determine which faults will be detectable before solving the basic optimal sensor placement problem (5.1). In practice, detectability of the most important faults is often required. In a special case, this requirement can refer to all faults.

The detectability of a given fault can be interpreted as the possibility to distinguish this fault from the state without faults. To satisfy detectability requirements, an additional constraint can be added to (5.1) in the following way:

$$D(f_k, \text{faultless state}) = \underset{x_{s_j}}{max}\left\{x_{s_j} : v_{j,k} \neq 0\right\} = 1. \tag{5.5}$$

This ensures that there is at least one signal sensitive to fault f_k.

If the problem with this additional constraint becomes infeasible, then it is impossible to meet the detectability requirements.

Example 5.2.2.

The detectability requirements for the problem introduced in Example 5.2.1 can be formulated in the following way:

$$\begin{aligned} f_1: &\quad max\{x_{s_1}\} = x_{s_1} = 1, \\ f_2: &\quad max\{x_{s_1}, x_{s_2}\} = 1, \\ f_3: &\quad max\{x_{s_1}, x_{s_2}, x_{s_3}\} = 1. \end{aligned} \tag{5.6}$$

Isolability constraints

For some critical subset of faults it may be beneficial to require the solution of the optimal sensor placement problem to isolate these faults. These requirements can be fulfilled by introducing additional equality constraints. For example, if it is important that a fault f_k is isolable from a fault f_m, then the following constraint should be added:

$$x_{D_{k,m}} = 1. \tag{5.7}$$

If unidirectional strong isolability is desired, then two constraints should be added:

$$\begin{aligned} x_{D_{k,m}} &= 1, \\ x_{D_{m,k}} &= 1. \end{aligned} \tag{5.8}$$

If the isolability requirements cannot be satisfied, then the constrained problem will be infeasible.

Example 5.2.3.

For the diagnostic system introduced in Example 5.2.1, it is required that the fault f_3 is isolable from both f_1 and f_2. Then the following constraints should be added:

$$\begin{aligned} x_{D_{3,1}} &= 1, \\ x_{D_{3,2}} &= 1. \end{aligned} \tag{5.9}$$

Budgetary constraints

The costs of implementation of a diagnostic system need to be taken into consideration when dealing with a real implementation. These costs include:

- investments costs – new equipment, its installation, and commissioning, e.g., sensors, communication modules etc.,
- development costs of the diagnostic system, e.g., designing and identification of models, necessary experiments etc.

Typically, there are some budgetary limits. These limits reduces the set of affordable solutions of the optimal sensor placement problem. Some solutions with better isolability might be too expensive for implementation.

A budgetary constraint of an optimization problem can be expressed by:

$$\boldsymbol{c}^T \boldsymbol{x} \leq b, \tag{5.10}$$

where: $\boldsymbol{c}$ is the cost vector, $\boldsymbol{x}$ is the vector of decision variables for sensors and diagnostic signals, and b is the available budget.

The total cost of implementation of a diagnostic system can be used as a secondary objective function in the case where multiple solutions provide identical isolability. This can be realized by solving a modified optimization problem where objective function

$$\underset{x}{\text{maximize}} \frac{1}{(K-1)K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m \neq k}}^{K} x_{D_{k,m}}, \tag{5.11}$$

is replaced with

$$\underset{x}{\text{minimize}}\; \boldsymbol{c}^T \boldsymbol{x}. \tag{5.12}$$

To retain isolability properties, the following constraint should be added:

$$\frac{1}{(K-1)K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m \neq k}}^{K} x_{D_{k,m}} \geq G, \tag{5.13}$$

where G is the desired value of the isolability measure.

The cost of a diagnostic system can also be understood as the number of new sensors to be installed.

Example 5.2.4.

Tab. 5.2 presents an example of costs of sensors and diagnostic signals considered in Example 5.2.1.

Table 5.2: Costs of sensors and implementation of diagnostic signals for the simple example.

x_1	x_2	x_3	s_1	s_2	s_3
1	2	3	4	2	3

The budgetary constraint has the following form:

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix} \leq b. \tag{5.14}$$

5.3 Problem formulation for FIS

In the case of FIS, the calculation of $D(f_k, f_m)$ according to the formula (5.2) is insufficient. For a given fault, there are possible multiple alternative signatures with different isolability properties. This results in conditional isolability. The set of diagnostic signals sensitive to a fault f_k and excluding f_m depends on alternative signatures of the fault f_k. The formula (3.4) should be expressed in terms of decision variables x_{s_j} in order to use $D(f_k, f_m)$ for the multi-valued diagnostic signal:

$$x_{D_{k,m}} = \frac{1}{card\left(\Phi\left(f_k\right)\right)} \sum_{\boldsymbol{\phi} \in \Phi(f_k)} max\left\{x_{s_j} : s_j\left(\boldsymbol{\phi}\right) \neq 0 \wedge s_j\left(\boldsymbol{\phi}\right) \notin s_j\left(f_m\right)\right\}, \tag{5.15}$$

where: $\Phi\left(f_k\right)$ is the set of alternative signatures of f_k, $s_j\left(\boldsymbol{\phi}\right)$ is the value of s_j for the alternative signature $\boldsymbol{\phi}$, and $s_j\left(f_m\right)$ is the set of possible values of s_j for the fault f_m. In the case of multi-valued diagnostic signals, $x_{D_{k,m}}$ is not binary.

Example 5.3.1.

In Tab. 5.3 an example of a simple FIS is presented.

Table 5.3: Simple FIS with two faults and two diagnostic signals.

	f_1	f_2
s_1	$-1, +1$	-1
s_2	$-1, +1$	-1

The $x_{D_{2,1}} = 0$ *because the signature of the fault* f_2 *does not exclude the fault* f_1. *For the fault* f_1 *there are four alternative signatures:*

- $[-1, -1]^T$ *– does not exclude* f_2*,*
- $[+1, -1]^T$ *– excludes* f_2 *if* s_1 *is available,*
- $[-1, +1]^T$ *– excludes* f_2 *if* s_2 *is available,*
- $[+1, +1]^T$ *– excludes* f_2 *if any diagnostic signal is available.*

Therefore, the optimisation problem for this system can be formulated in the following way:

$$\begin{aligned} \underset{x}{maximize} \quad & \frac{1}{2}\left(x_{D_{1,2}} + x_{D_{2,1}}\right), \\ s.t. \quad & x_{D_{1,2}} = \frac{1}{4}\left(max\left\{x_{s_1}\right\} + max\left\{x_{s_2}\right\} + max\left\{x_{s_1}, x_{s_2}\right\}\right), \\ & x_{D_{2,1}} = 0, \\ & x_{s_1}, x_{s_2} \in \{0, 1\}. \end{aligned} \tag{5.16}$$

For simplicity, sensor constraints were omitted in this example.

If both x_{s_1} *and* x_{s_2} *are available then* $x_{D_{1,2}}$ *is equal to* $\frac{3}{4}$. *If only one diagnostic signal is available then* $x_{D_{1,2}}$ *is equal to* $\frac{1}{2}$*, and if neither of them is available then* $x_{D_{1,2}} = 0$.

Multi-valued diagnostic signals that were obtained from sequences of symptoms (Section 3.2.4) instead of sensors should require diagnostic signals that are used to calculate their values. For example, if s_3 is calculated using the order of appearance of symptoms of s_1 and s_2, then $x_{s_3} = min\{x_{s_1}, x_{s_2}\}$.

5.4 Problem linearization

The original problem (5.1) with constraints (5.4) is non-linear. There are techniques that will allow us to solve this problem easily, by constructing an equivalent, higher dimensional, linear optimisation problem. To simplify non-linear constraints, two corollaries will be exploited:

Corollary 5.4.1.

The problem $\underset{x}{maximize}\ min\{x_1, \dots, x_k\}$ *has the same optimal solution as linear, constrained problem (Boyd and Vandenberghe 2004):*

$$\begin{aligned} &\underset{x}{maximize} && x_{k+1}, \\ &s.t. && x_{k+1} \le x_1, \\ & && \vdots \\ & && x_{k+1} \le x_k. \end{aligned} \tag{5.17}$$

The solution of (5.17) is the biggest lower bound (infimum) of a set $\{x_1, \dots, x_k\}$. For finite sets it is always equal to minimum, which is the solution of the original problem.

Corollary 5.4.2.

Binary Integer Linear Programming BILP problem $\underset{x}{maximize}\ max\{x_1, \dots, x_k\}\ s.t. x_i \in \{0, 1\}$ *has the same optimal solution as:*

$$\begin{aligned} &\underset{x}{maximize} && min\{x_1 + x_2 + \dots + x_k, 1\} \\ &s.t. && x_i \in \{0, 1\}. \end{aligned} \tag{5.18}$$

When x_i is binary, i.e. $x_i \in \{0, 1\}$, $\max\{x_1, \dots, x_k\} = 0$ iff $x_1 = x_2 = \dots = x_k = 0$. In such a case $x_1 + x_2 + \dots + x_k = 0$. Otherwise, $x_1 + x_2 + \dots + x_k \ge 1$ and $\min\{x_1 + x_2 + \dots + x_k, 1\} = 1$.

Corollary 5.4.1 makes it possible to linearize requirements of sensors for diagnostic signals. Corollary 5.4.2 followed by 5.4.1 can be used to formulate a linear equivalent of the calculation of $x_{D_{k,m}}$. Using Corollaries 5.4.1 and 5.4.2 it is possible to construct a higher dimensional, linear equivalent of (5.4) by substituting *min* and *max* functions with new, constrained variables. Some of the new equality constraints that result from Corollary 5.4.2 may be repeated multiple times.

Duplicates of constraints should be removed to reduce the number of constraints in the final problem.

If multi-valued diagnostic signals are used, then $x_{D_{k,m}}$ may be not binary and the optimization problem becomes Mixed Integer Linear Programming MILP, which is a generalization of BILP. These substitutions can be easily automated and used in a tool that builds a BILP (or MILP) problem using only BDM (or FIS) and a list of sensor requirements as inputs.

Example 5.4.1.

Using techniques presented in this section, the example presented in (5.4) *can be transformed into a BILP problem. In this example, there is no need to introduce new control variables. The following BILP problem can be directly obtained:*

$$\begin{aligned}
\underset{x}{maximize} \quad & \frac{1}{6}\sum_{k=1}^{K}\sum_{\substack{m=1\\m\neq k}}^{K} x_{D_{k,m}} \\
s.t. \quad & \boldsymbol{c}^T\boldsymbol{x} \le b, \\
& x_{s_1} \le x_1, \\
& x_{s_1} \le x_2, \\
& x_{s_2} \le x_1, \\
& x_{s_2} \le x_3, \\
& x_{s_3} \le x_2, \\
& x_{s_3} \le x_3, \\
& x_{D_{2,1}} = x_{s_2}, \\
& x_{D_{3,1}} \le x_{s_2} + x_{s_3}, \\
& x_{D_{3,2}} = x_{s_3}, \\
& x_{s_1}, x_{s_2}, x_{s_3} \ge 0, \\
& x_{D_{3,1}} \in \{0, 1\}, \\
& x_i, x_{s_j} \in \{0, 1\}, \quad i, j = 1 \ldots 3.
\end{aligned} \tag{5.19}$$

5.5 Solving ILP problem with the branch-and-bound algorithm

An Integer Linear Programming ILP problem can be solved using a family of methods called branch-and-bound. These methods work by analyzing the lower or upper bounds of subsets (branches) of possible solutions. If these bounds indicate that the analyzed subset cannot contain a better solution than the current best, then the whole subset is discarded (bounding).

An example of the branch-and-bound method developed for the purpose of this thesis is presented as Algorithm 1.

In order to solve the original problem this algorithm solves a series of relaxed LP problems. A constraint $\boldsymbol{x} \in \{0, 1\}$ is replaced with $0 \leq \boldsymbol{x} \leq 1$. In Algorithm 1, constraints are represented as four matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{A}_{eq}, \boldsymbol{B}_{eq}$: $\boldsymbol{A}x \leq \boldsymbol{B}$ for inequality constraints and $\boldsymbol{A}_{eq}\boldsymbol{x} = \boldsymbol{B}_{eq}$ for equality constraints. If a solution of a relaxed problem is not feasible, i.e., some integer variable x_i is not an integer, then two new LP problems (nodes) are created. One of the nodes has the constraint $x_i = 0$ and the other $x_i = 1$. Both problems are added to the node queue. This operation is called branching. Usually, there is more than one infeasible (non-integer) variable. There are many heuristics focused on choosing on which one to branch. In the Algorithm 1 the most infeasible variable is chosen. Variable feasibility is calculated as $abs(0.5 - \boldsymbol{x})$. The variable closest to 0.5 is the most infeasible.

The algorithm retains the current best integer solution. The solution to a relaxed problem is always equal or better than the solution to the original problem. Therefore, if the optimal solution to the relaxed problem is worse than the current best, then such node is discarded. A node is also discarded if, after branching, the relaxed LP problem is no longer feasible. These operations are called bounding.

Algorithm 1 considers one more case. $C(x)$ is the total cost of solution x. If the solution for the current node is an integer and it has an equal value of the objective function as the best solution, then total costs are compared and the cheaper solution is selected. This ensures that the cheapest solution offering optimal isolability is always selected.

Algorithm 1 $x = branchAndBound(f, A, b, Aeq, beq, x_b)$

```
if relaxed problem is infeasible then return 0
end if
x ← solveLP(f,A,b,Aeq,beq)
if f^T x < f^T x_b then return 0                               ▷ Bounding
else
    if isInteger(x) then
        if f^T x = f^T x_b then
            if C(x) < C(x_B) then return x
            else
                return 0
            end if
        else
            return x
        end if
    else                                                       ▷ Branching
        i ← chooseInfisible(x)
        [Aeq_1 Beq_1] ← addConstraint(0, i, Aeq, beq)
        [Aeq_2 Beq_2] ← addConstraint(1, i, Aeq, beq)
        x_1 ← branchAndBound(f, A, b, Aeq_1, beq_1, x_B)
        if f^T x_1 ≥ f^T x_B then x_B ← x_1
        end if
        x_2 ← branchAndBound(f, A, b, Aeq_2, beq_2, x_B)
        if f^T x_2 ≥ f^T x_B then x_B ← x_2
        end if
        return x_B
    end if
end if
```

5.6 Optimal sensor placement problem for an electro–pneumatic actuator

To demonstrate an example of the optimal sensor placement formulation, an electro–pneumatic valve actuator will be discussed. Fig. 5.1 illustrates the actuator (Bartyś 2016a). It consists

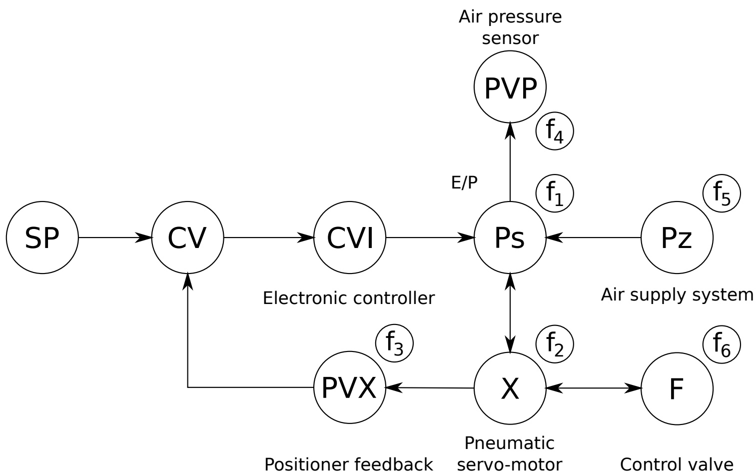

Figure 5.1: Causal graph of the electro-pneumatic actuator (Bartyś 2016a). SP – set point, CV – control value, CVI – control value of the electro-pneumatic transducer, PVP – pressure measurement in servo–motor chamber, PVX – stem displacement, F – flow rate.

of an electronic controller, an electro–pneumatic converter, a servo–motor, a control valve, and an electro–mechanical stem position feedback. The list of available measurements include the control value CV, the control value of the electro–pneumatic transducer CVI, the stem displacement measurement X, and the pressure in the chamber of the servo–motor. Tab. 5.4 lists the considered faults. Tab. 5.5 specifies the considered binary diagnostic matrix.

Using Tab. 5.5, the maximum value of the metric of isolability ψ can be calculated as:

$$\psi = \frac{1}{(K-1)\,K} \sum_{k=1}^{K} \sum_{\substack{m=1 \\ m\neq k}}^{K} D\left(f_k, f_m\right) = \frac{9}{30} = 0.3. \tag{5.20}$$

To find the diagnostic structure with $\psi = 0.3$ and the minimum number of required sensors the optimal sensor placement problem should be formulated as (5.21).

Table 5.4: List of actuator faults.

Fault	Faulty component
f_1	E/P transducer
f_2	Pneumatic servomotor
f_3	Position feedback
f_4	Pressure sensor fault
f_5	Supply air pressure
f_6	Control valve

Table 5.5: List of considered diagnostic signals for the electro–pneumatic actuator.

Signal	Residual	f_1	f_2	f_3	f_4	f_5	f_6
s_1	$X - f(CV)$	1	1	1		1	1
s_2	$X - f(CVI)$	1	1	1		1	1
s_3	$X - f(P_s)$		1	1		1	1
s_4	$P_s - f(CV)$	1	1	1	1	1	1
s_5	$P_s - f(CVI)$	1	1	1	1	1	1

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & x_{CV} + x_{CVI} + x_{Ps} + x_X \\
\text{s.t.} \quad & \frac{1}{30}\sum_{k=1}^{K}\sum_{\substack{m=1\\ m\neq k}}^{K} x_{D_{k,m}} = 0.3, \\
& x_{s_1} \leq x_X, \\
& x_{s_1} \leq x_{CV}, \\
& x_{s_2} \leq x_X, \\
& x_{s_2} \leq x_{CVI}, \\
& x_{s_3} \leq x_X, \\
& x_{s_3} \leq x_{Ps}, \\
& x_{s_4} \leq x_{Ps}, \\
& x_{s_4} \leq x_{CV}, \\
& x_{s_5} \leq x_{Ps}, \\
& x_{s_5} \leq x_{CVI}, \\
& x_{D_{2,1}} \leq x_{s_3}, \\
& x_{D_{3,1}} \leq x_{s_3}, \\
& x_{D_{5,1}} \leq x_{s_3}, \\
& x_{D_{6,1}} \leq x_{s_3}, \\
& x_{D_{1,4}} \leq x_{s_1} + x_{s_2}, \\
& x_{D_{2,4}} \leq x_{s_1} + x_{s_2} + x_{s_3}, \\
& x_{D_{3,4}} \leq x_{s_1} + x_{s_2} + x_{s_3}, \\
& x_{D_{5,4}} \leq x_{s_1} + x_{s_2} + x_{s_3}, \\
& x_{D_{6,4}} \leq x_{s_1} + x_{s_2} + x_{s_3}, \\
& x_{CV}, x_{CVI}, x_X, x_{Ps}, x_{s_j}, x_{D_{k,m}} \in \{0,1\}, \quad k, m = 1\ldots6,\ j = 1\ldots5.
\end{aligned}
\tag{5.21}
$$

To ensure that all faults are detectable the following constraints should be added:

$$\begin{aligned}
f_1: &\quad x_{s_1} + x_{s_2} + x_{s_4} + x_{s_5} \geq 1, \\
f_2: &\quad x_{s_1} + x_{s_2} + x_{s_3} + x_{s_4} + x_{s_5} \geq 1, \\
f_3: &\quad x_{s_1} + x_{s_2} + x_{s_3} + x_{s_4} + x_{s_5} \geq 1, \\
f_4: &\quad x_{s_4} + x_{s_5} \geq 1, \\
f_5: &\quad x_{s_1} + x_{s_2} + x_{s_3} + x_{s_4} + x_{s_5} \geq 1, \\
f_6: &\quad x_{s_1} + x_{s_2} + x_{s_3} + x_{s_4} + x_{s_5} \geq 1.
\end{aligned} \tag{5.22}$$

The problem (5.21) with (5.22) was solved using a Coin–or branch–and–cut (Cbc) solver and a PuLP modeler. The following solution was returned by the solver: $x_{CV} = 1.0$, $x_{CVI} = 0.0$, $x_{Ps} = 1.0$, $x_X = 1.0$. Consequently, the optimal sensor set for given constraints is $\{CV, Ps, X\}$ and the resulting BDM is presented in Tab. 5.6.

Table 5.6: Optimal BDM for the electro–pneumatic actuator.

	f_1	f_2	f_3	f_4	f_5	f_6
s_1	1	1	1		1	1
s_3		1	1		1	1
s_4	1	1	1	1	1	1

In Tab. 5.7 the values of different metrics of isolability for the resulting BDM are presented.

5.7 Binary optimal sensor placement problem with budgetary constraints for Fuel Cell Stack System

The Fuel Cell Stack System will be used as an example of an optimal sensor placement problem. Fuel cells are electrochemical devices that convert chemical energy from a gas fuel into electricity. In this example PEMFC (Polymer Electrolyte Membrane Fuel Cell) is analysed. Hydrogen is supplied to an anode and oxygen to a cathode. As a result of a chemical reaction, water

Table 5.7: Values of metrics of isolability for the electro–pneumatic actuator.

Metric	Value
Diagnosability degree	0.5
Diagnosis accuracy with exoneration	0.5
Diagnosis accuracy without exoneration	0.2
Isolability index with exoneration	18
Isolability index without exoneration	9
Measure of isolability ψ	0.3

and electric energy are produced. A detailed description was given in (Pukrushpan 2003). In (Sarrate *et al.* 2012a) a simplified structural was presented. Variables used in this example are shown in Tab. 5.8. In the same paper, approximate sensor costs were proposed. The considered faults are listed in Tab. 5.9.

The method of optimal sensor placement discussed in this chapter requires information about possible diagnostic signals and new sensors required. One way to obtain the information is to make use of expert knowledge. There is also work on automatic generation of diagnostic structures. The model structures and BDM, which are summarized in Tables 5.10 and 5.11 respectively, were obtained by means of the casual graph method presented in (Sztyber *et al.* 2015; Sztyber 2017). SN_0 denotes that any of following sensors can be used: p_{ca}, W_{cp}, $W_{sm,out}$, ω_{cp}, p_{sm}. Some of these signals require sensors that are already installed. To handle this case, the already measured values were added to the optimization problem with costs equal to 0.

Using the method presented in the previous section, the BILP problem was derived from Tables 5.10 and 5.11. The obtained linear problem has 59 variables and 132 inequality constraints. There are 2^{59} possible solutions.

To solve this problem, Algorithm 1 was used. It was implemented in the Matlab environment. A standard simplex solver was used to solve LP problems. Other, more advanced solvers can also be used, e.g., branch-and-cut.

The optimal values of the isolability measure for different budgetary constraints are presented in Fig. 5.2. The total cost of the best performing FDI system is 7. A further increase of the budget

Table 5.8: Model variables.

Control variables, already measured	
V_{cm}	Compressor voltage
W_{cp}	Air flow through the compressor
I_{st}	Stack current
V_{st}	Stack voltage

Unmeasurable variables	
τ_{cm}	Compressor motor torque
τ_{cp}	Load torque
$W_{v,inj}$	Humidifier injector flow

Possible sensor locations		Costs
ω_{cp}	Compressor angular speed	2
p_{sm}	Supply manifold pressure	1
$W_{sm,out}$	Supply manifold exit flow	5
p_{ca}	Cathode pressure	1
$W_{ca,out}$	Cathode output flow	5
p_{an}	Anode pressure	1
$W_{an,in}$	Anode input flow	5
$W_{rm,out}$	Return manifold exit flow	5

Table 5.9: Fuel Cell Stack System faults.

	Description
$f_{p_{sm}}$	Compressor fault
$f_{W_{sm,out}}$	Supply manifold fault
$f_{W_{rm,out}}$	Return manifold faul
$f_{I_{st}}$	Fuel Cell Stack fault
f_n	Cell fault

Table 5.10: Proposed models for diagnostic signals.

	output	inputs
1	$W_{rm,out}$	V_{cm},I_{st}
2	$W_{rm,out}$	I_{st},SN_0
3	p_{an}	I_{st}
4	p_{ca}	W_{cp}
5	p_{ca}	V_{cm}
6	W_{cp}	V_{cm}
7	W_{cp}	p_{ca},V_{cm}
8	V_{st}	p_{an},I_{st},SN_0
9	V_{st}	p_{an},V_{cm},I_{st}
10	V_{st}	V_{cm},I_{st}
11	V_{st}	I_{st},SN_0

Table 5.11: BDM for the Fuel Cell Stack System.

	$f_{p_{sm}}$	$f_{W_{sm,out}}$	$f_{W_{rm,out}}$	$f_{I_{st}}$	f_n
s_1	1	1	1		
s_2			1		
s_3					1
s_4		1			
s_5	1	1			
s_6	1	1	1		
s_7	1	1			
s_8				1	
s_9	1	1		1	
s_{10}	1	1		1	1
s_{11}				1	1

does not improve isolability (Tab. 5.3). The number of nodes created by the branch-and-bound algorithm is shown in Fig. 5.4. It is worth noticing that even in the worst case the total number of nodes was much smaller than theoretical 2^{59} or even 2^8 (where 8 is the number of new sensors that can be installed).

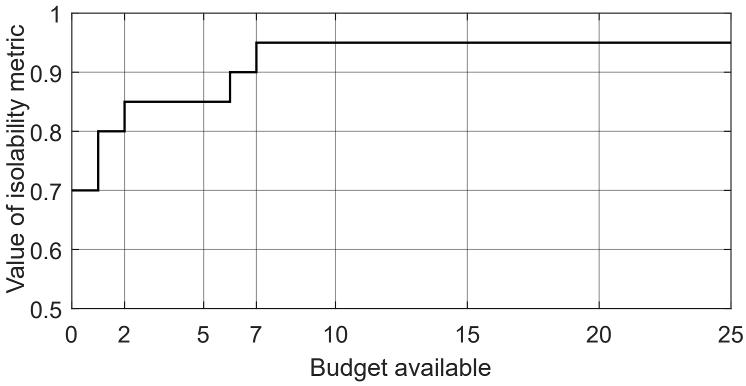

Figure 5.2: Value of the isolability metric of optimal solutions for different budget values.

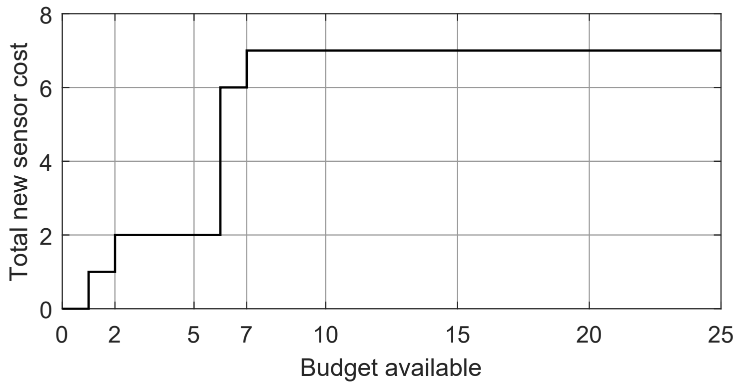

Figure 5.3: Total costs of new sensors of optimal solutions for different budget values.

The obtained results can be compared to those from (Sarrate *et al.* 2012a). There, the optimal solution obtained with three different methods was $S^* = \{p_{ca}, p_{an}\}$ with the total cost $C(S^*) = 2$. This is the same result as obtained in this Section with budgetary constraint $2 \leq B < 6$. In Fig. 5.2 one can see that it is possible to improve the value of the isolability metric. It is because additional measurements make some weakly isolating pairs of faults strongly isolating. With

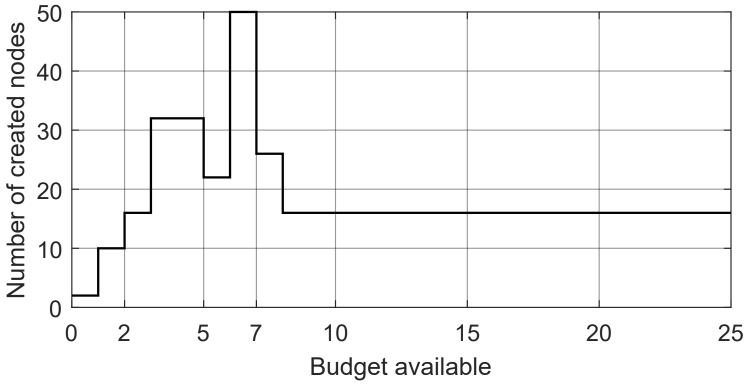

Figure 5.4: Number of nodes created to find optimal solutions for different budget values.

budget $B = 6$ sensors $\{p_{an}, W_{rm,out}\}$ are chosen and with budget $B \geq 7$ sensors $\{p_{ca}, p_{an}, W_{rm,out}\}$. The obtained results depend on a set of new model structures that are considered during the optimization procedure. If this set is incomplete, then results may not be optimal.

5.8 Comparison of optimal sensor placement problems with different metrics of isolability

In this section, the results of an optimal sensor placement problem obtained with the proposed metric of isolability are compared with results obtained with other metrics. Both binary and multi-valued versions of a problem are compared. The computational complexity of different approaches is also analyzed. As an example the Three Tank System is used. It is a popular diagnostic benchmark with a well documented analytic description.

5.8.1 Formulation of Optimization Problem

The optimal sensor set is determined by the following requirements:

1. The preferred solution makes it possible to detect all faults, i.e., for each fault there should be at least one diagnostic signal sensitive to this fault.

2. If many solutions offer the same detectability, the solution with higher isolability metric is chosen.

3. If there are many solutions with the same detectability and isolability, the cheapest one is selected.

The optimization problem was solved with the branch-and-bound algorithm (Sarrate *et al.* 2012b). For some isolability metrics it was impossible to use Algorithm 1, because problems were non-linear. Therefore, a much simpler branch-and-bound algorithm was applied. A simple heuristic was used for determining the upper isolability limit. It estimates maximum isolability of a branch by relaxing the budgetary constraint and choosing all sensors that are not excluded. A discrete decision variable was defined for each sensor with three possible values: positive, negative and unknown. The upper bound of the currently analyzed solution is compared to the current best solution according to the criteria given above.

5.8.2 Three Tank System example

The proposed analysis was performed on the example of a FIS for a system of three tanks (TTS) depicted in Fig. 5.5. Sixteen faults were considered (Tab. 5.12). Seven new sensor locations with cost estimation are proposed in Tab. 5.13. The cost of CV_v is 0, because this signal is already available in the diagnosed system.

By using the method presented in (Sztyber *et al.* 2015), 25 possible diagnostic signals were generated, using sensors listed in Tab. 5.13. Tri-valued diagnostic signals were considered. All budget values lower or equal to the total cost of all proposed sensors were analyzed.

5.8.3 Results of the comparison

Two main cases were compared. In the first one, BDM was used. It was obtained from Tab. 5.14 by replacing non-zero values with 1. In the second case, three-valued diagnostic signals were used as presented in Tab. 5.14. In each case, the optimization process was repeated for each appropriate isolability metric and for each budget value. The optimization results from the first case are presented in Tab. 5.16. Similarly, the results of the second case are presented in Tab. 5.15. The optimal sensor set is presented with the corresponding value of the metric for each combination of metrics and available budgets.

Table 5.12: Considered faults.

Fault	Description
f_1	measurement chain $FI102$ fault
f_2	measurement chain $LI103$ fault
f_3	measurement chain $LI104$ fault
f_4	measurement chain $LIC105$ fault
f_5	measurement chain $XI101$ fault
f_6	control signal fault
f_7	valve actuator fault
f_8	valve fault
f_9	pump fault
f_{10}	low water level fault
f_{11}	clogging between tanks $T1$ and $T2$
f_{12}	clogging between tanks $T2$ and $T3$
f_{13}	clogged outflow from tank $T3$
f_{14}	leak from tank $T1$
f_{15}	leak from tank $T2$
f_{16}	leak from tank $T3$

Table 5.13: Sensor locations and costs.

Symbol	Measurement	Cost
F_1	Inflow	5
P_v	Valve position	2
L_1	Level in T1	1
L_2	Level in T2	1
L_3	Level in T3	1
CV_v	Control signal	0
p_{zp}	Pressure on pump inlet	1
n	Pump rotational speed	2

Table 5.14: Diagnostic signals and their expected directions of change caused by faults.

Signals (sensor sets)	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}
$s_1(F_1, CV_v, p_{zp}, n)$	±					±	±	−	−	−						
$s_2(F_1, P_v, p_{zp}, n)$	±				±			−	−	−						
$s_3(P_v, CV_v)$					±	±	±									
$s_4(F_1, L_3)$	±			±							−	−	+	−	−	−
$s_5(P_v, L_3, p_{zp}, n)$				±	±			−	−	−	−	−	+	−	−	−
$s_6(L_3, CV_v, p_{zp}, n)$				±		±	±	−	−	−	−	−	+	−	−	−
$s_7(L_2, L_3)$			±	±								−	+			−
$s_8(L_1, L_3)$		±		±							−	−	+		−	−
$s_9(F_1, L_2)$	±		±								−	+	+	−	−	−
$s_{10}(P_v, L_2, p_{zp}, n)$			±		±			−	−	−	−	+	+	−	−	−
$s_{11}(L_2, CV_v, p_{zp}, n)$			±			±	±	−	−	−	−	+	+	−	−	−
$s_{12}(F_1, L_2, L_3)$	±		±	±							−	+		−	−	
$s_{13}(L_1, L_2, L_3)$		±	±	±							−	+			−	
$s_{14}(L_2, L_3, CV_v, p_{zp}, n)$			±	±		±	±	−	−	−	−	+		−	−	
$s_{15}(P_v, L_2, L_3, p_{zp}, n)$			±	±	±			−	−	−	−	+		−	−	
$s_{16}(L_1, L_2)$	±		±								−	+	+		−	−
$s_{17}(F_1, L_1)$	±	±									+	+	+	−	−	−
$s_{18}(F_1, L_1, L_2)$	±	±	±								+			−		
$s_{19}(P_v, L_1, L_2, p_{zp}, n)$		±	±		±			−	−	−	+			−		
$s_{20}(L_1, CV_v, p_{zp}, n)$		±				±	±	−	−	−	+	+	+	−	−	−
$s_{21}(P_v, L_1, p_{zp}, n)$		±			±			−	−	−	+	+	+	−	−	−
$s_{22}(F_1, L_1, L_3)$	±	±		±							+	+		−	−	
$s_{23}(L_1, L_3, CV_v, p_{zp}, n)$		±		±		±	±	−	−	−	+	+		−	−	
$s_{24}(P_v, L_1, L_3, p_{zp}, n)$		±		±	±			−	−	−	+	+		−	−	
$s_{25}(L_1, L_2, CV_v, p_{zp}, n)$		±	±			±	±	−	−	−	+			−		

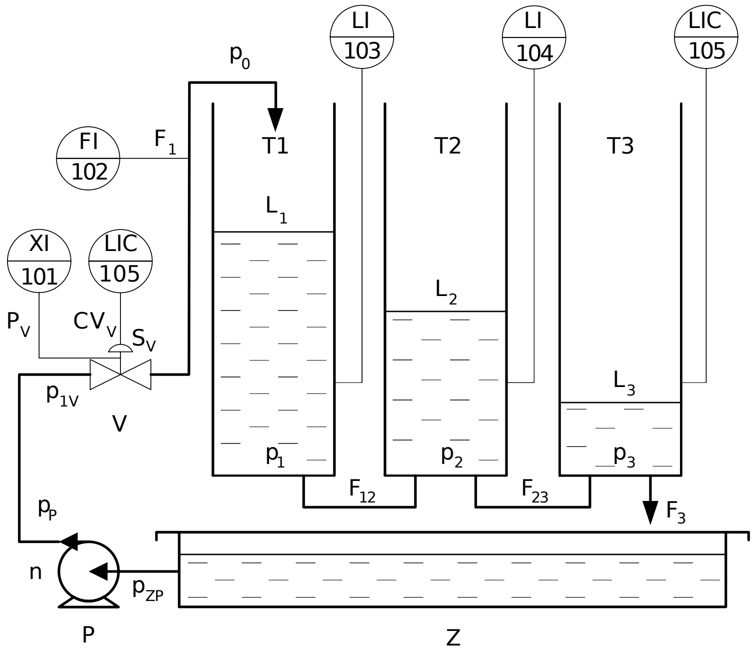

Figure 5.5: Three tank system (Ostasz 2006).

Table 5.15: Results of optimization with various measures of isolability for the set of multi-valued diagnostic signals.

Available budget						
2:	from 3 to 4:	5:	6:	7:	from 8 to 12:	13:
Diagnosis accuracy with exoneration:						
0.17:	0.56:	0.63:	0.67:	0.67:	0.75:	0.81:
$\mathbf{L_1}$ $\mathbf{L_2}$	L_1 L_2 L_3	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_1 L_2 L_3 P_v n p_{zp}	CV_v F_1 L_1 L_2 L_3 P_v n p_{zp}
Diagnosis accuracy without exoneration:						
0.13:	0.31:	0.31:	0.33:	0.34:	0.54:	0.75:
$\mathbf{L_2}$ $\mathbf{L_3}$	L_1 L_2 L_3	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 n p_{zp}	$\mathbf{F_1}$ $\mathbf{L_2}$ $\mathbf{L_3}$	CV_v L_1 L_2 L_3 P_v n p_{zp}	CV_v F_1 L_1 L_2 L_3 P_v n p_{zp}
Isolability Measure:						
0.33:	0.53:	0.73:	0.83:	0.85:	0.93:	0.97:
L_1 L_2	L_1 L_2 L_3	$\mathbf{CV_v}$ $\mathbf{L_1}$ $\mathbf{L_2}$ **n** $\mathbf{p_{zp}}$	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_1 L_2 P_v n p_{zp}	CV_v L_1 L_2 L_3 P_v n p_{zp}	CV_v F_1 L_1 L_2 L_3 P_v n p_{zp}

Table 5.16: Results of optimization with various measures of isolability for the set of binary diagnostic signals.

Available budget						
2:	from 3 to 4:	5:	6:	7:	from 8 to 12:	13:
Diagnosis accuracy with exoneration:						
0.13:	0.50:	0.56:	0.63:	0.63:	0.69:	0.75:
CV_v P_v	L_1 L_2 L_3	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_1 L_2 L_3 P_v n p_{zp}	CV_v F_1 L_1 L_2 L_3 P_v n p_{zp}
Diagnosis accuracy without exoneration:						
0.11:	0.205:	0.214:	0.23:	0.24:	0.39:	0.55:
CV_v P_v	L_1 L_2 L_3	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_2 L_3 P_v n p_{zp}	CV_v L_1 L_2 L_3 P_v n p_{zp}	CV_v F_1 L_1 L_2 L_3 P_v n p_{zp}
Isolability Measure:						
0.26:	0.48:	0.64:	0.73:	0.75:	0.87:	0.93:
L_1 L_2	L_1 L_2 L_3	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_1 L_2 P_v n p_{zp}	CV_v L_1 L_2 L_3 P_v n p_{zp}	CV_v F_1 L_1 L_2 L_3 P_v n p_{zp}
Diagnosability degree:						
0.13:	0.50:	0.56:	0.63:	0.63:	0.69:	0.75:
CV_v P_v	L_1 L_2 L_3	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_1 L_2 L_3 P_v n p_{zp}	CV_v F_1 L_1 L_2 L_3 P_v n p_{zp}
Isolability Index with exoneration:						
126:	194:	218:	218:	218:	224:	230:
L_1 L_2	L_1 L_2 L_3	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 P_v n p_{zp}	CV_v F_1 L_1 L_2 L_3 P_v n p_{zp}
Isolability Index without exoneration:						
63:	115:	154:	175:	179:	208:	223:
L_1 L_2	L_1 L_2 L_3	CV_v L_1 L_2 L_3 P_v	CV_v L_1 L_2 L_3 n p_{zp}	CV_v L_1 L_2 P_v n p_{zp}	CV_v L_1 L_2 L_3 P_v n p_{zp}	CV_v F_1 L_1 L_2 L_3 P_v n p_{zp}

The achieved results are very close despite the used isolability metric. In each case, if a budget is greater than 7, the results are identical for all considered metrics. This is due to the fact that the set of sensors $\{CV_v, L_1, L_2, L_3, P_v, n, p_{zp}\}$ is the cheapest set providing detectability of all considered faults. In each case, when the available budget is equal to 13, all of the proposed sensors are selected.

There are 4 differences in optimization results obtained from multi-valued and binary diagnostic signals. They were marked bold in Tab. 5.15. The sensor sets from Tab. 5.16 that are different than those from Tab. 5.15 were analyzed. The corresponding values of isolability metrics were calculated with FIS. Diagnosis accuracy with exoneration for the set $\{CV_v, P_v\}$ equals 0.13. Diagnosis accuracy without exoneration for the same set is 0.11. Both values are identical as in the case of BDM because the signal s_3 does not provide any new fault isolability in the multi-valued case. The set $\{CV_v, L_2, L_3, P_v, n, p_{zp}\}$ with diagnosis accuracy without exoneration gives the value of the metric equal to 0.32. Finally, the isolability measure of the set $\{CV_v, L_1, L_2, L_3, P_v\}$ is 0.69.

The results obtained with BDM are close to optimal results obtained with FIS. This shows that results of binary optimization can be chosen as a reasonable initial guess for the optimization algorithm used for multi-valued diagnostic signals. It is especially important because computational complexity of calculation of metrics of isolability grows with the increase of the number of alternative fault signatures. In the analyzed example, the optimization procedure for binary diagnostic signals is on average 2.1 times faster than for multi-valued ones.

It is also worth noticing that in all studied cases the value of the isolability measure obtained for FIS is greater than for BDM. Similarly, a greater value of isolability is obtained where exoneration is taken into consideration. The optimization results for budget value 7 with the exoneration assumption are in all cases identical to those with available budget equal to 6, whereas without that assumption the results are different. This implies that smaller differences between solutions are taken into consideration without adoption of the exoneration assumption. When designing a diagnostic system, it is important to check if this assumption can be satisfied. If the expected time between occurrence of the first and last symptom is greater than the time needed to formulate a diagnosis, then the assumption may generate false diagnoses.

5.9 Conclusion

In this chapter the sensor placement problem was addressed. A key contribution is formulation of the objective function based on the proposed metric of fault isolability. A strategy of introducing new variables to obtain BILP and MILP problems was presented. This makes it possible to use efficient solvers to find optimal sets of sensors.

The proposed method can be used for both binary and multi-valued diagnostic signals. The performed tests showed that, for the Three Tank System, the results for both types of diagnostic signals were similar in most cases. Therefore, results from optimal sensor placement are good initial guesses for optimal sensor placement with multi-valued diagnostic signals. It would allow to ignore groups of possible solutions early using the branch-and-bound algorithm and result in a significant decrease in time needed to solve an optimization problem. The following algorithm can be used to efficiently solve optimal sensor placement problems for FIS:

1. Transform FIS into BDM by replacing non-zero values with 1.
2. Solve the optimal sensor placement problem for BDM.
3. Solve the original problem using the solution from step (2) as an initial guess.

Various possible additional constraints were analyzed. They include budgetary constraints and the required detectability and isolability of faults.

Chapter 6

Summary

A novel metric of single fault isolability was introduced in this thesis. This metric has the following features:

- It is applicable to various types of diagnostic signals, e.g., binary, multi-valued, continuous, and sequences of symptoms.
- It makes it possible to distinguish weak and unidirectional strong isolability.
- It can be used to formulate a linear optimal sensor placement problem that can be solved with standard optimization solvers.

There are other known metrics of isolability, but they do not have all features of the proposed metric.

- Diagnosability degree (Travé-Massuyès *et al.* 2001) can be only used with binary diagnostic signals. It does not differentiate between weak and strong isolability and cannot be used as a linear objective function in an optimal sensor placement problem.
- Diagnosis accuracy (Kościelny 2001) can be used with any type of diagnostic signals and distinguish weak and strong isolability. However, it is very computationally complex, and an optimal sensor placement problem formulated with it is not linear.
- Isolability index (Sarrate *et al.* 2014) can be used as a linear objective function or a constraint and considers weak and strong isolability. It cannot, however, be used with diagnostic signals different than binary.

A quantitative, general approach to the analysis of fault isolability makes it possible to compare different methods of fault detection and isolation. Such approach can be used in the design phase of a diagnostic system. In particular, a quantitative analysis of fault isolability is indispensable when formulating and solving the problem of optimal sensor placement for diagnostic purposes.

6.1 Main contributions

The author considers the following contributions of this thesis the most important:

- Gathering formal definitions of fault isolability for various types of FDI systems and the rules of combining these definitions (Chapter 2). The analyzed types of diagnostic structures include BDM, FIS, sequential residuals, functional diagnosability and directional residuals.
- The novel metric of isolability proposed in Chapter 3. It can be used with binary, multi-valued and continuous diagnostic signals. It takes into account unidirectional strong isolability.
- A method of calculation of the metric of isolability for sequences of symptoms by constructing multi-valued diagnostic signals that provide identical isolability properties (Section 3.2.4).
- A method of formulation of an optimal sensor placement problem using the proposed metric of isolability. It can be used with binary (Section 5.2) and multi-valued diagnostic signals and sequences of symptoms (Section 5.3). The obtained optimization problem is a Mixed Integer Linear Programming problem, which can be solved using standard optimization tools and frameworks.
- A method for formulation of linear constraints, which can be added to an optimization problem, including budgetary constraints or detectability and isolability requirements (Section 5.2.1).

6.2 Future work

The optimal sensor placement problem is getting increased attention in the field of diagnostics of processes. It is also an important step towards automated or semi-automated tools supporting the development of diagnostic systems. It can be useful when combined with other approaches that are focused on automatic generation of potential structures of qualitative models, i.e., a graph of a process GP (Sztyber *et al.* 2015; Sztyber 2017) or a bond graph (Biswas *et al.* 2009; Ould-Bouamama *et al.* 2012). However, further work that focuses on the generation of model structures is necessary.

Another unsolved problem is how to efficiently optimize sensor placement in large-scale systems. Such systems can usually be partially decoupled. However, optimal decoupling of such systems is an open question.

The assumption about faultlessness of sensors is obviously unrealistic. An assumption regarding faultless sensors, which is adopted in this thesis, should be relaxed in future works. In general, the methods presented in Chapter 5 are sufficient for a formulation of a sensor placement problem with the assumption of faulty sensors. However, the problem is no longer linear, and it is much more complex to solve.

6.3 Acknowledgements

This work was partially supported by the Warsaw University of Technology, Faculty of Mechatronics Dean's Grants: 504/01536, 504/02141 and 504/02801.

Bibliography

Bartyś, Michał (2013). “Generalized reasoning about faults based on the diagnostic matrix”. In: *International Journal of Applied Mathematics and Computer Science* 23.2, pp. 407–417. doi: 10.2478/amcs-2013-0031.

Bartyś, Michał (2014). “Multiple Fault Isolation Algorithm Based on Binary Diagnostic Matrix”. In: *Intelligent Systems in Technical and Medical Diagnostics*. Springer, pp. 441–452.

Bartyś, Michał (2015). “Diagnosing Multiple Faults with the Dynamic Binary Matrix”. In: *IFAC-PapersOnLine* 48.21, pp. 1297–1302.

Bartyś, Michał (2016a). “Single Fault Isolability Metrics of the Binary Isolating Structures”. In: *Advanced and Intelligent Computations in Diagnosis and Control*. Springer, pp. 61–75.

Bartyś, Michał (2016b). “The weak isolability of the structure of binary residuals of multiple faults”. In: *Advanced Mechatronics Solutions*. Ed. by Ryszard Jabłoński and Tomas Brezina. Springer International Publishing, pp. 3–11. isbn: 978-3-319-23923-1. doi: 10.1007/978-3-319-23923-1_1. url: http://dx.doi.org/10.1007/978-3-319-23923-1_1.

Bartyś, Michał *et al.* (2006). “Introduction to the DAMADICS actuator FDI benchmark study”. In: *Control Engineering Practice* 14.6. A Benchmark Study of Fault Diagnosis for an Industrial ActuatorA Benchmark Study of Fault Diagnosis for an Industrial Actuator, pp. 577 –596. issn: 0967-0661. doi: http://dx.doi.org/10.1016/j.conengprac.2005.06.015.

Basseville, Michèle (1997). “Information criteria for residual generation and fault detection and isolation”. In: *Automatica* 33.5, pp. 783–803.

Basseville, Michèle (2001). “On fault detectability and isolability”. In: *European Journal of Control* 7.6, pp. 625–637.

Biswas, Gautam *et al.* (2004). “Diagnosis of complex systems: Bridging the methodologies of the FDI and DX communities”. In: *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 34.5, pp. 2159–2162.

Biswas, Gautam *et al.* (2009). “Analytic redundancy, possible conflicts, and TCG-based fault signature diagnosis applied to nonlinear dynamic systems”. In: *IFAC Proceedings Volumes* 42.8, pp. 1486–1491.

Blanke, Mogens *et al.* (2006). *Diagnosis and fault-tolerant control*. Vol. 691. Springer.

Boyd, Stephen and Lieven Vandenberghe (2004). *Convex optimization*. Cambridge university press.

Chen, Jie and Ron J Patton (1999). *Robust model-based fault diagnosis for dynamic systems*. Springer Science & Business Media.

Chen, Jie and Ron J Patton (2012). *Robust model-based fault diagnosis for dynamic systems*. Vol. 3. Springer Science & Business Media.

Chow, E. and A.S. Willsky (June 1984). “Analytical redundancy and the design of robust failure detection systems”. In: *Automatic Control, IEEE Transactions on* 29.7, pp. 603–614. ISSN: 0018-9286. DOI: 10.1109/TAC.1984.1103593.

Cordier, Marie-Odile *et al.* (2004). “Conflicts versus analytical redundancy relations: a comparative analysis of the model based diagnosis approach from the artificial intelligence and automatic control perspectives”. In: *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 34.5, pp. 2163–2177.

De Kleer, Johan and Brian C Williams (1987). “Diagnosing multiple faults”. In: *Artificial intelligence* 32.1, pp. 97–130.

De Kleer, Johan, Alan K Mackworth, and Raymond Reiter (1992). “Characterizing diagnoses and systems”. In: *Artificial Intelligence* 56.2, pp. 197–222.

Ding, Steven X (2008). *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Springer Science & Business Media.

Düştegör, Dilek *et al.* (2006). “Structural analysis of fault isolability in the DAMADICS benchmark”. In: *Control Engineering Practice* 14.6, pp. 597–608.

Eriksson, Daniel, Erik Frisk, and Mattias Krysander (2013). “A method for quantitative fault diagnosability analysis of stochastic linear descriptor models”. In: *Automatica* 49.6, pp. 1591–1600.

Fillatre, L. and I. Nikiforov (Feb. 2007). “Non-Bayesian Detection and Detectability of Anomalies From a Few Noisy Tomographic Projections”. In: *Signal Processing, IEEE Transactions on* 55.2, pp. 401–413. ISSN: 1053-587X. DOI: 10.1109/TSP.2006.885693.

Flouladirad, M. and I. Nikiforov (June 2003). “Optimal statistical fault detection with nuisance parameters”. In: *American Control Conference, 2003. Proceedings of the 2003*. Vol. 4, 2997–3002 vol.4. DOI: 10.1109/ACC.2003.1243781.

Frank, Paul M (1994). “On-line fault detection in uncertain nonlinear systems using diagnostic observers: a survey”. In: *International journal of systems science* 25.12, pp. 2129–2154.

Frisk, Erik *et al.* (2012). “Diagnosability analysis considering causal interpretations for differential constraints”. In: *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 42.5, pp. 1216–1229.

Gertler, J (1997). “Fault detection and isolation using parity relations”. In: *Control Engineering Practice* 5.5, pp. 653 –661. ISSN: 0967-0661. DOI: http://dx.doi.org/10.1016/S0967-0661(97)00047-6. URL: http://www.sciencedirect.com/science/article/pii/S0967066197000476.

Gertler, Janos (1991). “Analytical redundancy methods in fault detection and isolation”. In: *Preprints of IFAC/IMACS Symposium on Fault Detection, Supervision and Safety for Technical Processes SAFEPROCESS’91*, pp. 9–21.

Gertler, Janos (1998). *Fault detection and diagnosis in engineering systems*. CRC press.

Gertler, Janos (2000). “Structured Parity Equations in Fault Detection and Isolation”. In: *Issues of fault diagnosis for dynamic systems*. Springer, pp. 285–313.

Górny, Bartłomiej and Antoni Ligęza (2002). “Model-based diagnosis of dynamic systems: systematic conflict generation”. In: *Logical and Computational Aspects of Model-Based Reasoning*. Springer, pp. 273–291.

Isermann, Rolf (2006). *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer Science & Business Media.

Isermann, Rolf and Peter Balle (1997). “Trends in the application of model-based fault detection and diagnosis of technical processes”. In: *Control engineering practice* 5.5, pp. 709–719.

Jung, Daniel, Erik Frisk, and Mattias Krysander (2015). “Quantitative isolability analysis of different fault modes”. In: *IFAC-PapersOnLine, Paris*.

Khorasgani, Hamed *et al.* (2014). “Off-line Robust Residual Selection Using Sensitivity Analysis”. In: *Proceedings of 25th International Workshop on Principles of Diagnosis, Graz*.

Kilic, Erdal (2008). “Diagnosability of fuzzy discrete event systems”. In: *Information Sciences* 178.3, pp. 858–70.

Koivo, H.N. (1994). “Artificial neural networks in fault diagnosis and control”. In: *Control engineering practice* 2.1, pp. 89–101.

Korbicz, Józef *et al.* (2004). *Fault diagnosis: models, artificial intelligence, applications*. Springer Science & Business Media.

Kościelny, Jan M. (1999). “Application of fuzzy logic for fault isolation in a three-tank system”. In: *Proc. 14-th IFAC World Congress, Bejing, China*, pp. 73–78.

Kościelny, Jan M. (2001). *Diagnostyka zautomatyzowanych procesów przemysłowych*. Warszawa: Akadem. Oficyna Wyd. EXIT.

Kościelny, Jan M. and Michał Syfert (2006). “Fuzzy diagnostic reasoning that takes into account the uncertainty of the relation between faults and symptoms”. In: *International Journal of Applied Mathematics and Computer Science* 16.1, p. 27.

Kościelny, Jan M. and K. Zakroczymski (2001). “Fault Isolation Algorithm That Takes Into Account Dynamics of Symptoms Appearances”. In: *Bulletin of the Polish Academy of Sciences. Technical Sciences* 49.2, pp. 323–36.

Kościelny, Jan M., D. Sędziak, and Karol Zakroczymski (1999). “Fuzzy-logic fault isolation in large-scale systems”. In: *International Journal of Applied Mathematics and Computer Science* 9.3, pp. 637–652.

Kościelny, Jan M. *et al.* (2006). “Actuator fault distinguishability study for the DAMADICS benchmark problem”. In: *Control engineering practice* 14.6, pp. 645–652.

Kościelny, Jan M., Michał Bartyś, and Michał Syfert (2008). “Diagnostics of Industrial Processes in Decentralised Structures with Application of Fuzzy Logic”. In: *{IFAC} Proceedings*

Volumes 41.2. 17th {IFAC} World Congress, pp. 6944 –6949. ISSN: 1474-6670. DOI: http://dx.doi.org/10.3182/20080706-5-KR-1001.01177. URL: http://www.sciencedirect.com/science/article/pii/S1474667016400613.

Kościelny, Jan M., Michał Syfert, and Łukasz Tabor (2013). "Application of knowledge about residual dynamics for fault isolation and identification". In: *Control and Fault-Tolerant Systems (SysTol), 2013 Conference on*. IEEE, pp. 275–280.

Kościelny, Jan M. *et al.* (2016). "Fault isolability with different forms of the faults–symptoms relation". In: *International Journal of Applied Mathematics and Computer Science* 26.4.

Krysander, Mattias and Erik Frisk (2008). "Sensor placement for fault diagnosis". In: *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 38.6, pp. 1398–1410.

Krysander, Mattias, J. Aslund, and Mattias Nyberg (2008). "An Efficient Algorithm for Finding Minimal Overconstrained Subsystems for Model-Based Diagnosis". In: *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 38.1, pp. 197–206.

Lamperti, Gianfranco and Marina Zanella (2013). *Diagnosis of active systems: principles and techniques*. Vol. 741. Springer Science & Business Media.

Lan, Jianglin and Ron J Patton (2016). "Integrated fault estimation and fault-tolerant control for uncertain Lipschitz nonlinear systems". In: *International Journal of Robust and Nonlinear Control*.

Nejjari, Fatiha, Ramon Sarrate, and Albert Rosich (2010). "Optimal sensor placement for fuel cell system diagnosis using bilp formulation". In: *Control & Automation (MED), 2010 18th Mediterranean Conference on*. IEEE, pp. 1296–1301.

Ostasz, Andrzej (2006). "Graf przyczynowo-skutkowy procesu i jego zastosowanie w wyznaczaniu zbioru residuow i relacji diagnostycznej". PhD thesis. Warsaw University of Technology.

Ould-Bouamama, B *et al.* (2012). "Bond graphs for the diagnosis of chemical processes". In: *Computers & chemical engineering* 36, pp. 301–324.

Papadopoulos, Yiannis and JA McDermid (2001). "Automated safety monitoring: A review and classification of methods". In: *International Journal of COMADEM* 4.4, pp. 14–32.

Patan, Maciej and Dariusz Uciński (2008). “Configuring a sensor network for fault detection in distributed parameter systems”. In: *International Journal of Applied Mathematics and Computer Science* 18.4, pp. 513–524.

Patton, Ron J., C.J. Lopez-Toribio, and F.J. Uppal (1999). “Artificial intelligence approaches to fault diagnosis for dynamic systems”. In: *International Journal of applied mathematics and computer science* 9.3, pp. 471–518.

Patton, Ron J, Paul M Frank, and Robert N Clark (2000a). *Issues of fault diagnosis for dynamic systems*. Springer Science & Business Media.

Patton, Ron J., Paul M. Frank, and Robert N. Clark (2000b). *Issues of fault diagnosis for dynamic systems*. Springer Science & Business Media.

Pouliezos, A., G. Stavrakakis, and C. Lefas (1989). “Fault detection using parameter estimation”. In: *Quality and Reliability Engineering International* 5.4, pp. 283–290. ISSN: 1099-1638. DOI: 10.1002/qre.4680050407.

Pukrushpan, Jay Tawee (2003). *Modeling and control of fuel cell systems and fuel processors*. University of Michigan.

Pulido, Belarmino and Carlos Alonso González (2004). “Possible conflicts: a compilation technique for consistency-based diagnosis”. In: *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 34.5, pp. 2192–2206.

Reiter, Raymond (1987). “A theory of diagnosis from first principles”. In: *Artificial intelligence* 32.1, pp. 57–95.

Rosich, Albert *et al.* (2007). “Efficient optimal sensor placement for model-based FDI using an incremental algorithm”. In: *Decision and Control, 2007 46th IEEE Conference on*. IEEE, pp. 2590–2595.

Rosich, Albert, Ramon Sarrate, and Fatiha Nejjari (2009). “Optimal sensor placement for fdi using binary integer linear programming”. In: *20th International Workshop on Principles of Diagnosis*.

Rostek, Kornel (2014). “Measure of fault isolability of diagnostic system”. In: *25th International Workshop on Principles of Diagnosis*. URL: http://dx-2014.ist.tugraz.at/papers/DX14_Tue_PM_posters_paper9.pdf.

Rostek, Kornel (2016). “Optimal Sensor Placement for Fault Information System”. In: *Advanced Mechatronics Solutions*. Ed. by Ryszard Jabłoński and Tomas Brezina. Cham: Springer International Publishing, pp. 67–72. ISBN: 978-3-319-23923-1. DOI: 10.1007/978-3-319-23923-1_10.

Rostek, Kornel, Łukasz Morytko, and Anna Jankowska (2015). “Early detection and prediction of leaks in fluidized-bed boilers using artificial neural networks”. In: *Energy* 89, pp. 914 –923. ISSN: 0360-5442. DOI: http://dx.doi.org/10.1016/j.energy.2015.06.042. URL: http://www.sciencedirect.com/science/article/pii/S0360544215007987.

Roth, Matthias (2010). *Identification and Fault Diagnosis of Industrial Closed-loop Discrete Event Systems: Identifikation und Fehlerdiagnose Industrieller Ereignisdiskreter Closed-Loop-Systeme*. Logos Verlag Berlin GmbH.

Rouissi, F. and G. Hoblos (2013). “Fault Tolerance in Wind Turbine Sensor Systems for Diagnosability Properties guarantee”. In: *Conference on Control and Fault-Tolerant Systems (SysTol)*, pp. 523–28.

Sanchez, Hector *et al.* (2015). “Fault diagnosis of an advanced wind turbine benchmark using interval-based ARRs and observers”. In: *IEEE Transactions on Industrial Electronics* 62.6, pp. 3783–3793.

Sarrate, Ramon *et al.* (2007). “Optimal sensor placement for model-based fault detection and isolation”. In: *Decision and Control, 2007 46th IEEE Conference on*. IEEE, pp. 2584–2589.

Sarrate, Ramon, Fatiha Nejjari, and Albert Rosich (2012a). “Model-based Optimal Sensor Placement Approaches to Fuel Cell Stack System Fault Diagnosis”. In: *Fault Detection, Supervision and Safety of Technical Processes, Volume# 8| Part# 1*.

Sarrate, Ramon, Fatiha Nejjari, and Albert Rosich (2012b). “Sensor Placement for Fault Diagnosis Performance Maximization under Budgetary Constraints”. In: *2nd International Conference on Systems and Control*.

Sarrate, Ramon *et al.* (2014). “Sensor placement for leak detection and location in water distribution networks”. In: *Water Science and Technology: Water Supply, 795–803*.

Spanache, Stefan, Teresa Escobet, and Louise Trave-Massuyes (2004). “Sensor Placement Optimisation Using Genetic Algorithms”. In: *Proceedings of the 15th International Workshop on Principles of Diagnosis, DX-04*.

Staroswiecki, Marcel and G Comtet-Varga (2001). “Analytical redundancy relations for fault detection and isolation in algebraic dynamic systems”. In: *Automatica* 37.5, pp. 687–699.

Staroswiecki, Marcel, J Ph Cassar, and Ph Declerck (2000). “A structural framework for the design of FDI system in large scale industrial plants”. In: *Issues of fault diagnosis for dynamic systems*. Springer, pp. 245–283.

Syfert, Michał and Jan Kościelny (2009). “Diagnostic reasoning based on symptom forming sequence”. In: *Fault Detection, Supervision and Safety of Technical Processes*, pp. 89–94.

Sztyber, A., A. Ostasz, and J.M. Kościelny (July 2015). “Graph of a Process - A New Tool for Finding Model Structures in a Model-Based Diagnosis”. In: *Systems, Man, and Cybernetics: Systems, IEEE Transactions on* 45.7, pp. 1004–1017.

Sztyber, Anna (2017). “Sensor Placement for Fault Diagnosis Using Graph of a Process”. In: *Journal of Physics: Conference Series*. Vol. 783. 1. IOP Publishing, p. 012007.

Travé-Massuyès, L (2014a). “Bridges between diagnosis theories from control and AI perspectives”. In: *Intelligent Systems in Technical and Medical Diagnostics*. Springer, pp. 3–28.

Travé-Massuyès, Louise (2014b). “Bridging control and artificial intelligence theories for diagnosis: A survey”. In: *Engineering Applications of Artificial Intelligence* 27, pp. 1–16.

Travé-Massuyès, Louise, Teresa Escobet, and R. Milne (2001). “Model-based diagnosability and sensor placement application to frame 6 gas turbine subsystem”. In: *In 12th Int, Workshop on principles of diagnosis*, pp. 205–12.

Travé-Massuyès, Louise, Teresa Escobet, and Xavier Olive (2006). “Diagnosability analysis based on component-supported analytical redundancy relations”. In: *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on* 36.6, pp. 1146–1160.

Verdière, Nathalie, Carine Jauberthie, and Louise Travé-Massuyès (2015). “Functional diagnosability and detectability of nonlinear models based on analytical redundancy relations”. In: *Journal of Process Control* 35, pp. 1 –10. ISSN: 0959-1524. DOI: `http://dx.doi.org/10.`

1016/j.jprocont.2015.08.001. URL: http://www.sciencedirect.com/science/article/pii/S0959152415001614.

Wang, Rixin, Yao Cheng, and Minqiang Xu (2015). "Analytical redundancy based fault diagnosis scheme for satellite attitude control systems". In: *Journal of the Franklin Institute* 352.5, pp. 1906 –1931. ISSN: 0016-0032. DOI: http://dx.doi.org/10.1016/j.jfranklin.2015.02.003. URL: http://www.sciencedirect.com/science/article/pii/S0016003215000678.

Yassine, Abed, Stéphane Ploix, and Jean-Marie Flaus (2008). "A method for sensor placement taking into account diagnosability criteria". In: *International Journal of Applied Mathematics and Computer Science* 18.4, pp. 497–512.

Yin, S. *et al.* (2015). "An Improved Incremental Learning Approach for KPI Prognosis of Dynamic Fuel Cell System". In: *IEEE Transactions on Cybernetics* PP.99, pp. 1–10. ISSN: 2168-2267.

Yin, S., G. Wang, and H. Gao (July 2016). "Data-Driven Process Monitoring Based on Modified Orthogonal Projections to Latent Structures". In: *IEEE Transactions on Control Systems Technology* 24.4. DOI: 10.1109/TCST.2015.2481318, pp. 1480–1487. ISSN: 1063-6536.

Zhou, Gan *et al.* (2015). "A comparison of state estimation algorithms for hybrid systems". In: *29th European conference on modelling and simulation*, pp. 312–319.

List of Figures

List of Tables