Model-based diagnosis through Structural Analysis and Causal Computation for automotive Polymer Electrolyte Membrane Fuel Cell systems

Pierpaolo Polverino, Erik Frisk, Daniel Jung, Mattias Krysander and Cesare Pianese

Tweet

LINKÖPINGS UNIVERSITET

Model-based diagnosis through Structural Analysis and Causal Computation for automotive Polymer Electrolyte Membrane Fuel Cell systems

Pierpaolo Polverino[a,*], Erik Frisk[b], Daniel Jung[b], Mattias Krysander[b], Cesare Pianese[a]

[a]*Department of Industrial Engineering, University of Salerno, via Giovanni Paolo II 132, 84084, Fisciano (SA), ITALY.*
[b]*Department of Electrical Engineering, Linköping University, S-5183 Linköping, SWEDEN.*

Abstract

The present paper proposes an advanced approach for Polymer Electrolyte Membrane Fuel Cell (PEMFC) systems fault detection and isolation through a model-based diagnostic algorithm. The considered algorithm is developed upon a lumped parameter model simulating a whole PEMFC system oriented towards automotive applications. This model is inspired by other models available in the literature, with further attention to stack thermal dynamics and water management. The developed model is analysed by means of Structural Analysis, to identify the correlations among involved physical variables, defined equations and a set of faults which may occur in the system (related to both auxiliary components malfunctions and stack degradation phenomena). Residual generators are designed by means of Causal Computation analysis and the maximum theoretical fault isolability, achievable with a minimal number of installed sensors, is investigated. The achieved results

---

*Corresponding author
  Email address: `ppolverino@unisa.it` (Pierpaolo Polverino)
  URL: `www.eprolab.unisa.it` (Pierpaolo Polverino)

proved the capability of the algorithm to theoretically detect and isolate almost all faults with the only use of stack voltage and temperature sensors, with significant advantages from an industrial point of view. The effective fault isolability is proved through fault simulations at a specific fault magnitude with an advanced residual evaluation technique, to consider quantitative residual deviations from normal conditions and achieve univocal fault isolation.

1. Introduction

The current advancements in materials design and components manufacturing brought Polymer Electrolyte Membrane Fuel Cells (PEMFCs) among the most suitable power production systems for small and medium size applications, e.g., automotive, backup, portable, etc. [1, 2]. Nevertheless, the limitations for PEMFC mass market deployment are mainly related to their high production and operating costs, as well as limited durability with respect to the reference targets set for the considered applications [3, 4].

To achieve these targets, the current research and industrial efforts are focused on both material and manufacturing processes improvement, and on-field operation reliability and durability increase. On this latter issue, suitable diagnostic algorithms coupled with efficient control strategies can be adopted [5, 6, 7, 8]. The proper combination of diagnostics and control can promptly act to prevent fault worsening, so as to reduce system degradation or avoid any abrupt failure, and thus improve system durability and lifetime [9, 10].

In the available literature, many approaches are proposed to cope with fault

2

diagnosis, and they can be roughly classified in model- and signal-based techniques [6, 11]. On the one hand, model-based diagnosis uses a mathematical model to simulate system variables during normal operation and to generate residuals by comparing the simulated variables with those measured on the system [5, 7, 8, 9, 12, 13, 14]. On the other hand, signal-based approaches directly treat measured signals to extract information and define different patterns representative of variable behaviour during both normal and faulty conditions [15, 16, 17, 18, 19].

Signal-based techniques require limited knowledge about the system under study but large amount of experiments, implying several issues such as infeasibility of the tests, high cost and complexity of the experiments, etc. [6]. On the contrary, model-based approaches are recognized as more robust against changes in system features, thus allowing the application of the same developed algorithm on different technologies/cases with slight modifications [14].

The availability of accurate models does not imply itself optimal fault diagnosis. Indeed, suitable residuals are required to correctly detect fault occurrence and ensure reliable system diagnosis. In many works, residuals are generated by comparing measured variables with those simulated by the model [5, 8, 20, 21]. However, this approach leads the number of residuals to equal the number of measurements. It is widely recognized that robust diagnosis can be achieved by increasing as much as possible the number of residuals, in accordance with the concept of analytical redundancy; this improves information extraction and reduces false alarm and miss detection occurrence. To meet these requirements, the use of Structural Analysis can be advised [20]. Many works available in the literature apply this technique for system diagnosis on, e.g., engines [22, 23], hybrid vehicles [24] and other complex systems [25, 26, 27], but only limited number of papers

3

focus on fuel cells [14, 28].

The most notable work applying Structural Analysis for fault detection on a PEMFC system is that of Rosich et al. [14]. The authors developed residual generators following a model-based approach combined with structural analysis. They use the model developed by [29] to design fault diagnosis structure. Analysing the structural information derived from the model, they manually selected sets of equations sensitive to chosen faults, taking into account the causal computability of the involved variables.

Although interesting and significantly valuable for the research community, the work of Rosich et al. [14] presents some limitations. For instance, the considered model does not take into account membrane water management and stack thermal-dynamics, being this latter assumed much slower than all the other phenomena. This is true during normal operation, when the fuel cell and the other components are working properly; but when dealing with faulty states, e.g., the occurrence of an abnormal behaviour in the fuel cell cooling system, or the worsening of some fuel cell degradation mechanisms, faster dynamics could be induced, driving the system towards abrupt failure. Another aspect relates to the faults and sensors considered in the work: seven faults at component level are investigated and nine sensors are considered (a cost analysis drove towards sensors choice). However, no abnormal events (either faults or degradation mechanisms) which may occur inside the stack nor sensor faults are considered, and no detail is given on the reasons leading to the fully isolability of the defined faults (i.e., if it is always the case or it is only related to the chosen model and faults).

To cope with the remarked limitations, the present work proposes an advanced model-based diagnostic algorithm for PEMFC fault detection and isolation. The

4

considered algorithm is based on a lumped PEMFC system model inspired by [29, 30, 31], with further attention on detailed membrane water transport and stack thermal dynamics. The developed model is analysed by means of Structural Analysis theory, as done in [14, 22, 25], to identify the correlations among physical variables, mathematical equations and a set of faults which may occur in the system. The novelty of this work lies on the extended number and type of faults, which are related not only to PEMFC system auxiliaries malfunctions but also to internal stack degradation. A set of residual generators is designed by means of Causal Computation analysis [23, 32], and the maximum theoretical fault isolability achievable with a minimal number of sensors is investigated. The related results are discussed paying also attention to both mathematical and physical motivations. Furthermore, the effective fault isolability is proved through fault simulations at a specific fault magnitude with an advanced residual evaluation technique, in order to account for the quantitative deviation of the monitored variables due to fault occurrence.

2. PEMFC system model overview

In this section the PEMFC system model used for the development of the diagnostic algorithm is introduced. An overview on the system layout and component interaction is given, in addition to the description of the faults considered for the algorithm evaluation (performed in Sections 4 and 6). It is worth observing that the proposed model is only part of the framework the diagnostic algorithm is related to. Therefore, the focus of the work is not on model development, but rather on the understanding of its structure and the links among the involved physical variables. For this reason, a brief description of the main model equations can be found in

5

the Appendix, to keep the focus here on the diagnostic algorithm design.

2.1. *System layout and modelling assumptions*

The proposed model is based on the work of Arsie et al. [30, 31], including a further analysis on the overall stack thermal dynamics. The considered system is an air-fed pressurized cathode PEMFC system oriented towards automotive applications, and it is made of ten components: a hydrogen tank, a volumetric air blower, two humidifiers, a PEMFC stack, and five nozzles. The schematic representation of the proposed system is given in Figure 1, where the aforementioned component are sketched. In the same picture, the location of the sensors (sketched with circles) and the analysed faults (represented with triangles) is also shown.

At the anode side, the hydrogen tank is assumed being an infinite capacity reservoir, capable of instantaneously providing the required amount of hydrogen. At the cathode side, the air blower feeds the PEMFC with oxygen taken from the surroundings at ambient conditions, with given pressure, temperature and relative humidity. The supply manifolds connect the hydrogen tank and the air blower to the PEMFC anode and cathode inlets, respectively. Both gas streams are here humidified to reach the required amount of water and ensure the correct operation of the stack. The humidity regulation follows a feed-forward control, setting in advance a relative humidity set-point and then injecting the required amount of water to reach the desired value. The nozzle are used to control the inlet and outlet mass flows of the electrodes and between the hydrogen tank and the anode supply manifold. The PEMFC stack is modelled through an equivalent cell, whose electrodes are modelled as distinct control volumes, whereas the electrolyte properties are evaluated by means of maps.

The model design process follows a Mean Value Modelling (MVM) approach

6

[5, 29, 30, 31]. Each control volume, enclosed in a dashed rectangle in Figure 1, is characterized by a dynamic mass balance, which takes into account species inlet and outlet flows and generation and consumption terms. Other relevant assumptions are: (i) the stack is represented by an equivalent single cell, thus neglecting performance variation within the stack length; (ii) the water inside the system is assumed being only vapour (no liquid water is considered); (iii) only stack thermal dynamics is taken into account. The reader is addressed to the Appendix for more details.

A PEMFC system oriented towards automotive applications is generally equipped with several sensors, to monitor the stack and auxiliary components behaviour. Nevertheless, in the present paper, only two sensors are considered, related to stack voltage and temperature measurements. These sensors allow increasing model redundancy, so as to generate suitable residuals for fault detection and isolation. This choice is mainly driven by the interest in understanding which is the maximum fault isolability achievable with minimum number of sensors, and how accurate is the diagnostic performance with only use of the stack as "sensor" for the overall system state.

The system configuration shown in Figure 1 and the assumptions illustrated above correspond to a test case study. This choice has been mainly driven by the need to investigate the diagnostic algorithm performance firstly in a basic layout, so as to single out the main issues to be tackled during algorithm development and application, before addressing other complex scenarios. Indeed, this work represents a first step towards the definition of a systematic approach for diagnostic algorithm design, which can be applied to any system layout and model structure. Therefore, other configurations, such as anode gas recirculation or dead-end mode,
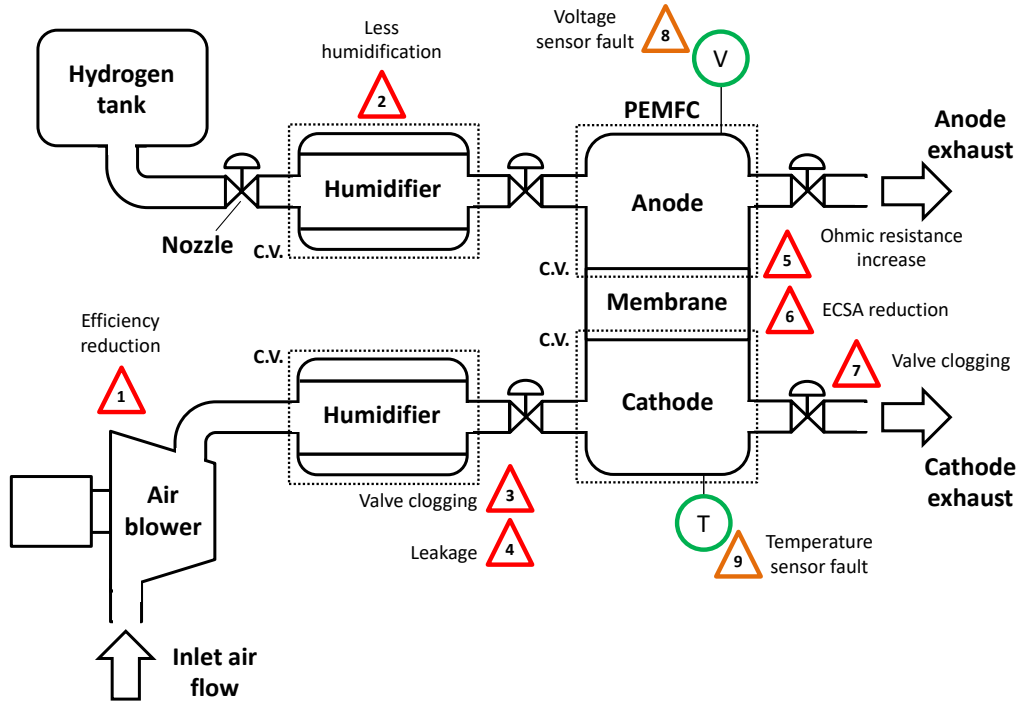
Figure 1: PEMFC system scheme with main components, sensors and analysed faults: faults 5 and 6 occur at stack level, faults 8 and 9 refer to sensors, whereas all the others affect system components.

as well as modelling assumptions, e.g., two-phase water transport (as addressed by the authors in [33, 34, 35]), can be straightforwardly introduced.

2.2. Fault models

The model equations illustrated in the Appendix represent the nominal (i.e., unfaulty) behaviour of the PEMFC system illustrated in Figure 1. To account for faulty states affecting different system components, the following assumptions are considered: (i) no detailed physical model of the fault causes is introduced, though only their effects on the simulated variables is analysed; (ii) only one fault

Table 1: PEMFC system faults, affected components and influenced system variables.

Fault	Affected component	Influenced Variable
f_{cmp}	Air blower	η_{cmp}
f_{inj}	Anode humidifier	$\dot{m}_{H_2O,an,sm,inj}$
f_{leak}	Cathode inlet	$\dot{m}_{ca,sm,out}$
f_{vin}	Cathode inlet valve	$C_{d_{ca,sm,out}}$
f_{ohm}	Stack	σ_{mem}
f_{ecsa}	Stack	$ECSA$
f_{vout}	Cathode outlet valve	$C_{d_{ca,out}}$
f_{vsens}	Voltage sensor	y_V
f_{tsens}	Temperature sensor	y_T

is occurring at a time (i.e., no combination of multiple faults is considered); (iii) each fault affects only one model equation. The faults considered in this paper are listed in Table 1, with the affected component and influenced variable. On the same line of the models assumptions introduced in the previous section, a test case study involving only one fault occurring at a time is here tackled. Nevertheless, coupling the proposed algorithm with the approaches presented by the authors in [32, 36] will allow also addressing multiple faults diagnosis issues.

The first fault concerns the increase in the air blower mechanical losses, due to, e.g., the growth of the electric motor friction due to lack of lubricant [9]. This fault is modelled as done in [5], considering the blower efficiency affected as:

$$\eta_{cmp}^{F} = \eta_{cmp}\left(1 - f_{cmp}\right) \tag{1}$$

The coefficient f_{cmp} represents the fault magnitude and is limited to the range

[0,1]: in normal conditions, the coefficient is 0, and assumes higher values during fault occurrence, reaching a maximum of 1 (i.e., system failure).

The second fault deals with the reduction in the water injected at the anode supply manifold to regulate anode humidification. This fault can be either caused by a malfunction in the humidifier components (e.g., injector clogging) or by a controller/actuator failure. The fault is modelled as follows:

$$\dot{m}_{H_2O,an,sm,inj}^F = \frac{\dot{m}_{H_2O,an,sm,inj}^F}{\tau_{inj}} \left(1 - f_{inj}\right) \tag{2}$$

where the coefficient f_{inj} expresses the fault magnitude, in the same way of f_{cmp}.

The third fault concerns a leakage at cathode inlet, which may be related to the sealing degradation of the cathode inlet nozzle valve or to a pipe crack. Therefore, to model the effects of this fault, the overall cathode inlet flow is reduced by a certain amount proportional to the parameter f_{leak} as:

$$\dot{m}_{ca,in}^F = \dot{m}_{ca,in} \left(1 - f_{leak}\right) \tag{3}$$

The parameter f_{leak} is 0 in normal conditions (i.e., no leakage), whereas 1 when failure occurs (i.e., the leakage amount corresponds to the overall mass flow leaving the cathode supply manifold).

The fourth fault represents a clogging or malfunction of the nozzle valve located at cathode inlet, inducing a reduction in the nozzle section. The effect of this fault can be directly related to the discharge coefficient expressed in equation (A.11):

$$C_{d_{sm,ca,out}}^F = C_{d_{sm,ca,out}} \left(1 - f_{vin}\right) \tag{4}$$

where the coefficient f_{vin} expresses the fault magnitude.

10

The fifth and sixth faults both involve the PEMFC stack: the former one deals with an increase in the fuel cell ohmic resistance, due to membrane degradation:

$$\sigma_m^F = \sigma_m \left(1 - f_{ohm}\right) \tag{5}$$

whereas the latter consists in the ECSA reduction, caused by, e.g., catalyst dissolution or delamination:

$$ECSA^F = ECSA \left(1 - f_{ecsa}\right) \tag{6}$$

The seventh fault represents a clogging or malfunction of the nozzle valve located at cathode outlet, inducing a reduction in the nozzle section. As done for the fault expressed in equation (4), the effect of this fault can be directly related to the discharge coefficient:

$$C_{d_{ca,out}}^F = C_{d_{sm,ca,out}} \left(1 - f_{vout}\right) \tag{7}$$

In addition to system component faults, also the stack voltage and temperature sensors are assumed being affected by a fault. Usually, sensors faults can be related to several causes, such as short or open circuit, wrong calibration, etc. [37], and they are here modelled as done for the previous faults:

$$y_V^F = y_V \left(1 - f_{vsens}\right) \tag{8}$$

$$y_T^F = y_T \left(1 - f_{tsens}\right) \tag{9}$$

The presented faults are only a limited set selected among a much larger group of events which can occur in an automotive PEMFC system. The choice was made driven by the need for investigating the capability of the proposed diagnostic

algorithm to detect and isolate faults involving both stack and auxiliary components, rather than only the stack. Indeed, as shown in Table 1, only two faults affect the stack, whereas five faults involve the auxiliaries and two the installed sensors. Moreover, the faults have been chosen also according to their modelling complexity, so as to respect the assumption of a fault affecting only one model equation.

In the next section, a brief description of the diagnostic methodology accounted in this work is given, with some explanatory examples.

3. Diagnostic algorithm framework

The methodology used in this work entails Structural Analysis theory and Causal Computation to generate suitable residuals based on model equations sensitive to the considered faults. Before illustrating and discussing the achieved results, a brief overview on the main theoretical features of the accounted methodology is given. For further details, the reader is addressed to the works listed in the Bibliography [20, 23, 27, 32].

3.1. Model-based diagnosis

A common approach in model-based diagnosis is to design residuals that are zero in non-faulty case and non-zero in faulty cases. Therefore, fault isolation can be achieved by using a collection of residuals sensitive to different sets of faults. In structural approaches, the decoupling of faults is achieved by computing residuals based on different equation subsets (i.e., different testable sub-models). Thus, faults not affecting a selected sub-model will not affect the corresponding residual [23].

Residual generators can be derived from a sub-model in many ways, but in this work *sequential residual generators* are chosen, see for example [20] or [38] for a

12

more formal description. This approach is here considered due to the capability of handling large models with differential and algebraic equations as well as maps in a straightforward manner.

To give a brief idea on how sequential residual generators are built, an illustrative example is given in the following set of 3 equations:

$$c_1: \quad \dot{x}_1 = -x_1 + u$$
$$c_2: \quad \dot{x}_2 = -x_2 + x_1 \tag{10}$$
$$c_3: \quad y = x_2$$

with 2 unknown variables x_1 and x_2 and two known parameters: an input u and a sensor signal y. This model has one more equation than unknowns and thus implies *analytical redundancy*, which can be used for diagnosis. Any two of these three equations can be used for computing the two unknown variables values, and then the remaining equation, usually called *redundant equation*, can be used to build a residual. It is worth observing that the redundant equation can be any of the aforementioned equations, hence c_1, c_2, or c_3 can be selected as the redundant equation and the corresponding sequentially computed residuals are illustrated in Table 2. One characteristics of the three alternatives resides in which mathematical operations are required to compute the residual: in Case 1, the first two equations should be integrated; in Case 2, the variable x_1 is obtained through integration and the residual by differentiating x_2; in Case 3, both x_1 and x_2 need to be differentiated.

In a computational sequence, the order in which the unknowns are computed is usually expressed as *causality*, which is given once a redundant equation has been selected: if only integration or differentiation is needed, it is referred to the causality type as *integral* or *derivative*, respectively, whereas if both differentiation and integration are required, it is generally recalled as *mixed causality* [27, 38].

Table 2: Example of residual generation based on equation (10).

Case 1	Case 2	Case 3
$\dot{x}_1 = -x_1 + u$	$x_2 = y$	$x_2 = y$
$\dot{x}_2 = -x_2 + x_1$	$\dot{x}_1 = -x_1 + u$	$x_1 = \dot{x}_2 + x_2$
$r = y - x_2$	$r = \dot{x}_2 + x_2 - x_1$	$r = \dot{x}_1 + x_1 - u$

Causality of equations is closely related to the *differential index* of differential-algebraic equations (see for example [39] for formal definitions). With respect to the scope of the paper, only a brief example is given. The following Differential-Algebraic Equation system:

$$\dot{x}_1 = g_1(x_1, x_2)$$
$$0 = g_2(x_1, x_2), \quad \frac{\partial g_2}{\partial x_2} \text{ invertible} \tag{11}$$

has low differential index and can be integrated using standard Ordinary Differential Equation techniques, since the algebraic variable x_2 can be locally solved using the algebraic constraint g_2. It is also straightforward to design observers or Kalman filters using standard state-space techniques. However, if the model is in the form:

$$\dot{x}_1 = h_1(x_1, x_2)$$
$$0 = h_2(x_1) \tag{12}$$

it is said to have *high differential index*. Here, the algebraic variable x_2 cannot be solved for in the algebraic constraint h_2. Therefore, the solution implies differentiation, which is a numerically sensitive operation. It is also not straightforward to design observers and residual generators for such a model.

The type of causality is tightly linked to the issues concerning residual generator design. Indeed, derivative causality implies, for instance, the need for

differentiating noisy signals, whereas integral causality to treat unknown initial conditions (and the related stability issues). Intuitively, mixed causality combines both potential problems.

In the illustrative example proposed in Table 2, it is possible to choose integral, derivative or mixed causality by selecting the specific redundant relation. However, there are some cases (as for the PEMFC system model presented in the Appendix) where mixed causality cannot be avoided given any choice of redundant equation.

In conclusion, crucial steps for proper residual generator design are: i) find a suitable equation subset to test (i.e., sub-model) which allows good fault isolability; ii) given a sub-model, select a redundant equation and thus a computational sequence leading to the desired causality; iii) define which variable shall be computed through each equation involved in the computational sequence. All the remarked steps can be approached with the use of *Structural Analysis*.

3.2. Structural Analysis theory

The Structural Analysis of a mathematical model consists in an analytical method that takes into account only which variables appear in each model equation. Formally, the structure of a model can be represented through a bipartite graph $G = (C \cup X; E)$, with a node set C corresponding to the mathematical relations (i.e., equations, look-up tables, maps, rules, etc.) referred to as *constraints*, a node set X corresponding to the model variables, and an edge set E with $(c,x) \in E$ if the variable $x \in X$ is included in the constraint $c \in C$ [20]. The variable set X can be partitioned with respect to unknown, known, and fault variables.

A model bipartite graph can be represented by a biadjacency matrix, in which the constraints are listed on the rows and the unknown variables on the columns. If the j-th variable is linked to the i-th constraint, the related (i,j)-th cell shows a

15

mark (i.e., one), otherwise is empty (i.e., zero).

3.3. Dulmage-Mendelsohn decomposition

The Dulmage-Mendelsohn decomposition [40] offers an advanced methodology to rearrange a model biadjacency matrix in an upper block triangular form, as shown in Figure 2. In this picture, the gray areas present both ones and zeros, whereas the white areas only zeros. The overall constraint set C and unknown variable set X are divided into three subsets each: C^-, C^0 and C^+ the former, and X^-, X^0 and X^+ the latter.

The subset G^- with sets C^- and X^- represents the underdetermined part of the model, in which more unknown variables than constraints are present. The subset G^0 with sets C^0 and X^0 comprises instead the exactly determined part, where the unknown variables are perfectly balanced by the constraints. This subset is further split in other subsets, represented as G_i ($i = \{1, \cdots, n\}$) and called Hall components. These components are exactly determined parts which need to be simultaneously solved as a whole system of equations. The last subset G^+ with sets C^+ and X^+ represents the overdetermined part of the model, where more constraints than unknown variables are available. Clearly, the overdetermined part of the model comprises redundant constraints which can be used to generate residuals for diagnostic purposes.

3.4. Fault detectability and isolability

Observing which constraints are influenced by the considered faults and in which Dulmage-Mendelsohn decomposition subset they are located, it is possible to check the structural *detectability* and *isolability* of the faults. Given a constraint c_{f_i} included in the overall model constraint set C (i.e., $c_{f_i} \in C$) and influenced by a
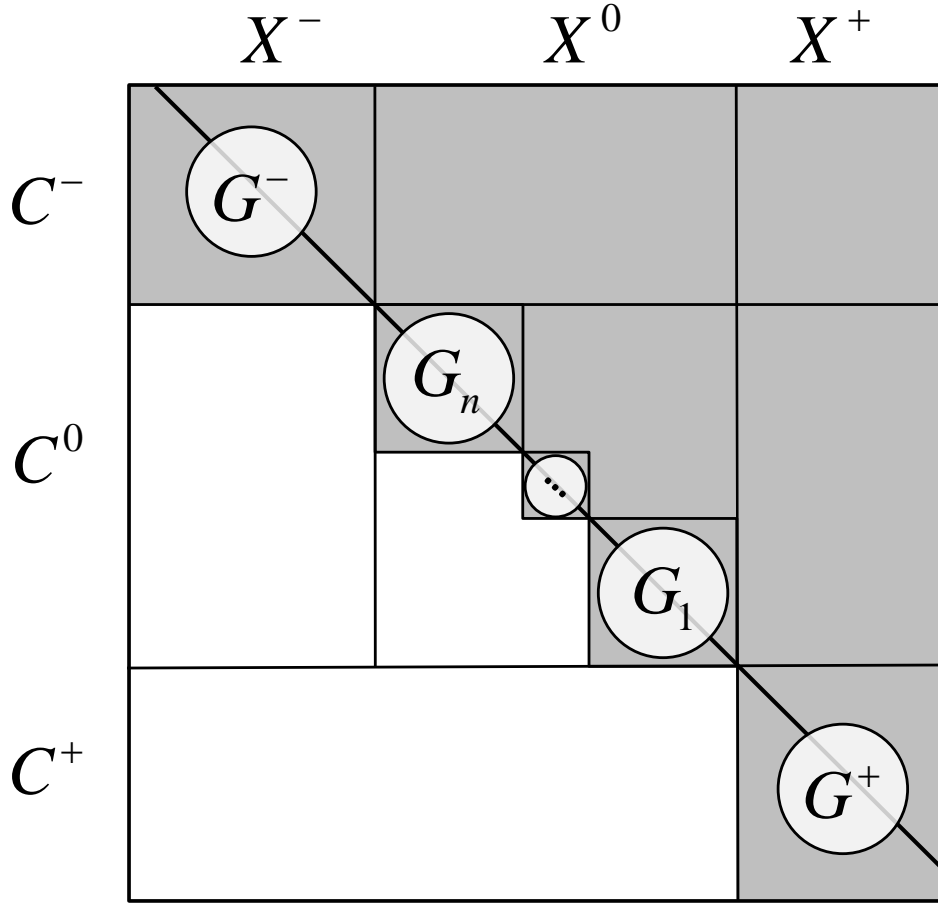
Figure 2: Generalized Dulmage-Mendelsohn decomposition structure, adapted from [23, 27].

fault f_i, it can be stated that the fault f_i is structurally detectable if the constraint c_{f_i} falls in the overdetermined part of the model C^+ (i.e., $c_{f_i} \in C^+$). Moreover, fault f_i is structurally isolable from another fault f_j if, removing from the whole model the constraint c_{f_j} influenced by fault f_j, the constraint c_{f_i} still falls in the overdetermined part of the new model (i.e., $c_{f_i} \in (C \setminus \{c_{f_j}\})^+$) [27].

The structural isolability result obtained in this way relies on the assumption

that residual generators can be constructed for any of the three causalities: derivative, integral or mixed causality. Nevertheless, fault detectability and isolability computed through Dulmage-Mendelsohn decomposition are obtained by analysis of a model abstraction. They may not be realizable considering the real system because only a qualitative analysis of the violation is performed, without considering the quantitative deviation of the related physical variables. However, it should be kept in mind that if fault isolability is not fulfilled through structural analysis, it cannot be achieved as well through quantitative residual computation.

3.5. Testable sub-models

Structural Analysis can be used to find testable sub-models useful for fault detection and isolation. Sub-models suitable for test design shall present the following properties: firstly, they should provide model redundancy to compute residuals (i.e., the model must have more equations than unknowns); secondly, models including few faults are preferred, since they allow more specific fault isolation.

A concept that entails these two structural properties is the so called Minimal Test Equation Support (MTES) (see [32] for a formal and detailed definition). An MTES consists in an overdetermined set of equations (inherently fulfilling the first property) being sensitive to a minimal set of faults (in line with the second property). This implies that it is not possible to define two different MTESs with the same redundancy degree sensitive to the same set of faults. Under certain conditions, MTESs can show model redundancy equal to 1, meaning that there is exactly one more equation than unknown variable. The model proposed in the equation set (10) is an example of such a case, and also the PEMFC model presented in this work applies to this specific situation (as shown in Section 4). An

efficient algorithm described in [32] is here used to find all possible MTESs.

3.6. Residual generation

The characterization of MTESs gives only the sets of constraints which can be used to generate residuals. To design a sequential residual generator from an MTES model C' of redundancy 1, a redundant equation $c \in C'$ needs to be singled out and a computational sequence to be determined. For each $c \in C'$, a computational sequence is computed by applying the Dulmage-Mendelsohn decomposition to the exactly determined system of equations $C' \setminus \{c\}$. Since $C' \setminus \{c\}$ is exactly determined, no underdetermined nor overdetermined part will appear in the structure shown in Figure 2; consequently, all the equations are partitioned in Hall components.

The causal order is given by initially solving the unknowns in the first Hall component G_1 (using the related constraints) and then continuing solving all the other components following an increasing order (i.e., from G_2 to G_n). Moreover, by analysing the Hall components it is possible to determine if the computational sequence is integral, derivative or mixed causality [27].

In the following sections, the results concerning the application of the described approach on the PEMFC system model proposed in the Appendix are presented and discussed. Structural fault detectability and isolability analysis is performed to check the diagnostic capabilities of the proposed model with respect to the considered faults. Then, suitably MTESs are defined through the algorithm presented in [32]. Causal computation analysis is performed and sequential residual generators are designed aiming at maximum structural fault isolability.

4. Diagnosability Analysis of the PEMFC Model

The PEMFC system model presented in the Appendix is a first order differential-algebraic model, with 108 algebraic equations, 11 differential equations, 2 sensor equations, 119 unknown variables, 13 known variables and 9 faults. Therefore, the overall number of equations is 121 with degree of redundancy 2 due to the considered installed sensors.

The structural isolability of the considered faults through the proposed model can be investigated by applying the Dulmage-Mendelsohn decomposition to the model biadjacency matrix, with a canonical decomposition [23] of the overdetermined part, as shown in Figure 3. The exactly determined part consists of only two equations and two unknown variables, whereas the remaining part of the model corresponds to the overdetermined part (enclosed in a blue rectangle).

The figure also reveals fault *equivalence classes* indicated by the grey boxes, first introduced in [23] and used in [27]. The equivalence classes corresponds to parts of the model that cannot be separately monitored, i.e., two faults appearing in the same equivalence class cannot be isolated from each other. The horizontal red lines represent the equations influenced by the faults. Thus, it can be observed that each affected equation falls inside the overdetermined part of the model, i.e., $c_{f_i} \in C^+$, implying that all the faults are structurally detectable (as assessed in Section 3.4). The isolability result can be seen by observing that no two faults affect the same equivalence class except the faults f_{ohm} and f_{ecsa}, which thereby are not isolable from each other.

As a result, the structural isolability of the model can be computed and afterwards represented through the isolability matrix in Figure 4-a. A dot in position (i, j) indicates that fault f_i cannot be isolated from fault f_j. The ideal performance

20
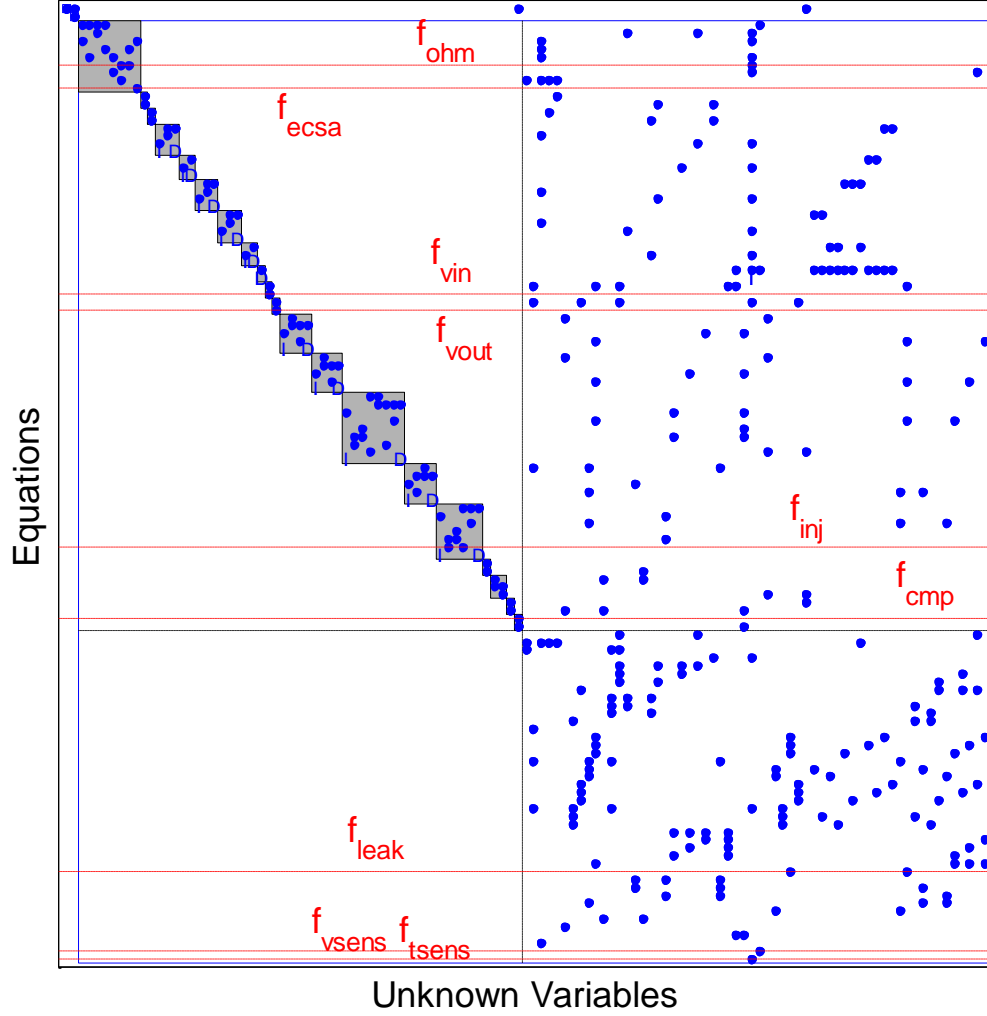
Figure 3: Dulmage-Mendelsohn decomposition structure of the PEMFC system model described in the Appendix.

would be a pure diagonal matrix, meaning that each fault can be univocally isolated from all the others. From the result presented in Figure 4-a, it is clear that all faults in the auxiliary units can be univocally isolated; the same conclusion can be drawn for any fault in the stack with respect to those of the auxiliaries. Nevertheless,

based on the model structure, it is not possible to isolate the two stack faults f_{ohm} and f_{ecsa} from each other.

The former result can be interpreted through a physical analysis of the considered faults and their effects. Membrane degradation can be due to either thermal, mechanic or chemical causes [41], with a consequent reduction in membrane ionic conductivity. This event induces an increase in membrane resistance and thus cell ohmic losses see equation (A.42)). Catalyst degradation directly affects the ECSA and, consequently, the activation losses (see equation (A.40)). In both cases, the fuel cell voltage is affected by both ohmic and activation losses (see equation (A.38)). Also cell temperature is influenced, since cell voltage appears in the thermal balance of equation (A.34). For this reason, both faulty events directly influence cell voltage and temperature in the same way, and with the accounted sensors it is not possible to correctly distinguish between their effects.

Figure 4-b shows the isolability matrix under *integral causality* where it is also indicated which faults correspond to stack, sensor, and auxiliaries. The result is computed using a similar argument as in Figure 4-a, and the details of the procedure can be found in [27]. The reason this result is interesting is that the integral causality case shows what isolability performance can be achieved only using low-index filters, i.e., where standard techniques such as state-observers can be used as residual generators. Thus, Figure 4-b shows what is possible if solutions are restricted to those similar to Case 1 in Table 2. If better performance is desired, solutions including derivatives as in Case 2 and Case 3 have to be used (i.e., high-index models have to be treated).

To investigate what is needed to isolate the stack faults, an extended sensor analysis, using the same technique as outlined above, reveals that the addition of
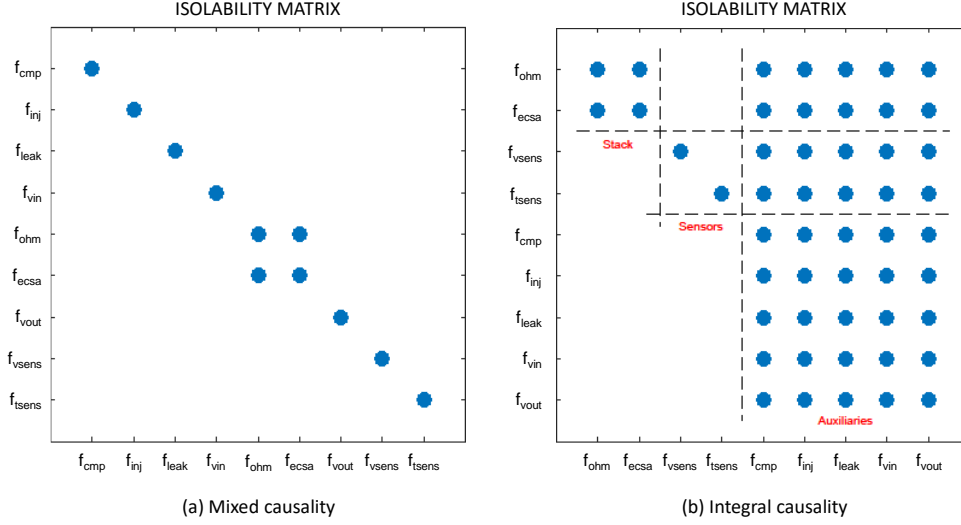
22

Figure 4: Isolability matrices for the PEMFC system model with mixed (a) and integral (b) causalities.

further measurements at balance of plant level (e.g., stack inlet and outlet pressure sensors, air blower mass flow meter and temperature sensors, etc.) does not enable stack fault isolation in any event. This conclusion may imply the need for more advanced measurements, which can give a deeper insight into stack phenomena, such as Electrochemical Impedance Spectroscopy, Current Voltammetry, etc. The use of such techniques, coupled with proper models and diagnostic algorithms, may improve fault detection and isolation at stack level and also reduce the incidence of degradation phenomena on overall system performance.

5. Diagnosis System Design

The analysis in the previous section showed that full structural fault isolation, except between the two stack faults, is possible but also that any diagnosis system

Table 3: Theoretical Fault Signature Matrix based on the identified residuals showing the structural fault sensitivity of the 8 MTES sets.

	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8
f_{cmp}	1	1	1	0	1	1	1	1
f_{inj}	1	1	0	1	1	1	1	1
f_{leak}	1	0	1	1	1	1	1	1
f_{vin}	1	1	1	1	0	1	1	1
f_{ohm}	0	1	1	1	1	1	1	1
f_{ecsa}	0	1	1	1	1	1	1	1
f_{vout}	1	1	1	1	1	0	1	1
f_{vsens}	1	1	1	1	1	1	0	1
f_{tsens}	1	1	1	1	1	1	1	0

will require solving high-index problems (i.e., low-index techniques are not sufficient). Applying the algorithm proposed in [32], 8 MTES sets are singled out. This means that, with two sensor equations there exists 8 testable sub-models that have all different fault signatures (i.e., are each sensitive to a different set of faults). A noteworthy result consists in the design of 8 residuals by using only 2 sensor measurements, being an advancement with respect to other approaches proposing a number of residuals equal to the number of installed sensors (e.g., [5, 8, 9]). The 8 MTES sets consists of 110 to 118 different equations, meaning that almost all model equations are included in each testable sub-model. Table 3 shows the Fault Signature Matrix (FSM) for the residuals linked to the defined MTES sets: each set is sensitive to all but one fault, except residual r_1 which is sensitive to both stack faults.

Using the approach proposed in [42], it is verified that all 8 sub-models are observable but only sub-model 1, 7, and 8 correspond to low-index sub-models. Thus, these three MTES sets can be directly transformed into residual generators using the technique described in Section 3.6, resulting in a solution corresponding to Case 1 in Table 2.

The 5 high index cases cannot be transformed into state-space form and thus be solved using only integration. Simulation of high-index systems is numerically difficult, since it generally corresponds to numerical differentiation. There are several approaches for high-index problems simulation, ranging from direct Baumgartner stabilization and projection methods to specially designed integration routines [43]. In this work, a simple PID-based approach is proposed. This technique is applicable at most to index 2 problems. The main principle is illustrated with a simple example. Consider a high-index MTES set residual generator, based on measurement signals y_1 and y_2 and unknown variables x_1 and x_2, with this latter being a scalar algebraic variable:

$$
\begin{aligned}
\dot{x}_1 &= g_1(x_1, x_2) \\
0 &= g_2(x_1, y_1) \\
r &= y_2 - h(x_1, x_2)
\end{aligned}
\tag{13}
$$

The sub-model proposed in equation (13) is high-index since the algebraic variables x_2 cannot be solved in the algebraic constraint g_2. To avoid differentiation, consider x_2 as a control signal that should stabilize the algebraic constraint, achieving the following simulation sub-model:

$$\dot{x}_1 = g_1(x_1, x_2)$$

$$x_2 = K_p \, g_2(x_1, y_1) + K_I \int g_2(x_1, y_1) \, dt \qquad (14)$$

$$r = y_2 - h(x_1, x_2)$$

where controller gains K_p and K_I are chosen to stabilize the residual generator. This form of the residual generator is by construction low-index and can be simulated using standard integrator routines, thus avoiding dealing with numerical differentiation.

6. Simulation Results

The computational sequence considered for residuals generation through the defined MTES sets refers to a mixed causality assumption, so as to achieve maximum fault isolability (see Figure 4-a). Nevertheless, the equation selected as redundant constraint for each residual generator consists in a sensor equation, being either equation (A.46) or equation (A.47) in accordance with the constraints included in the corresponding MTES sets. Moreover, to evaluate percentage residual deviation from normal conditions, the proposed residual formulations are:

$$r_V = \frac{y_V - V_{fc}}{V_{fc}} * 100 \qquad (15)$$

$$r_T = \frac{y_T - T_{fc}}{T_{fc}} * 100 \qquad (16)$$

The two residual forms in equations (15) and (16) are associated to each residual as described in Table 4. Clearly, the feature changing from one residual to the other is the way V_{fc} and T_{fc} are evaluated by means of the computational sequences

Table 4: Residuals properties.

Residual	Form	Problem type
r_1	r_T	Low index
r_2	r_T	High index
r_3	r_T	High index
r_4	r_V	High index
r_5	r_T	High index
r_6	r_T	High index
r_7	r_T	Low index
r_8	r_V	Low index

based on each MTES set. In Table 4, also information about the problem index is presented, remarking which MTES set requires the use of the PID-based approach defined in equation (14).

To prove fault detectability and isolability, a real case scenario is emulated: the defined PEMFC system model is used to simulate the considered faulty events at a specific magnitude and the quantitative residual deviation is analysed. White noise with Gaussian distribution is superimposed to stack voltage and temperature measurements to reproduce more realistic device outputs, as also proposed in [14]: in nominal operation, the mean stack voltage value is 0.72 V with noise standard deviation of 10^{-3} V, whereas the mean stack temperature value is 325 K with noise standard deviation of 0.07 K.

Only one fault is induced at a time and the residuals deviation is evaluated through a dedicated simulation run. All the considered faults are simulated assuming a fault magnitude of 10%, which means that all fault parameters in equations

(1) through (9) are set equal to 0.1. Residuals are computed by means of equations (15) and (16) (see Table 4), and an example of their raw behaviour in normal and faulty conditions (i.e., f_{ecsa} in the proposed example) is presented in Figure 5 and Figure 6, respectively.

As expected, residuals behaviour in normal condition oscillates around an average zero value with deviations only due to noise. Instead, in faulty conditions, e.g., when the stack fault f_{ecsa} occurs after about 3 h from the simulation start, the residuals deviate from zero with a magnitude depending on the influence the specific fault has on system components (as also observed in [5, 9]). Indeed, some residuals can diverge more than others from their nominal value, as can be seen comparing residuals r_3 and r_4 in Figure 6, which, according to the FSM of Table 3, should both deviate, but the r_4 bias is much more evident than that of r_3.

To enhance residuals deviation and, thus, fulfil detection and isolation of incipient faults (i.e., faults with small magnitude), raw residuals are first normalized with respect to their standard deviation in normal condition and then treated with the Cumulative Sum (CUSUM) sequence [44]:

$$\tau_i(k) = max\left(0, \tau_i(k-1) + |r_i(k)| - v\right) \tag{17}$$

where the drift factor v is chosen as the 33$^{\text{th}}$ percentile of the i-th residual $|r_i|$ distribution during normal condition.

The results concerning treated residuals behaviour with respect to all the proposed faults is summarized in Figure 7. Each frame gathers the CUSUM sequence behaviour of all residuals with respect to each fault occurrence. To allow a better visualization of the results, the y-axes of the frames presents the logarithm of τ. All simulations start with the system behaving normally and then the faults (one
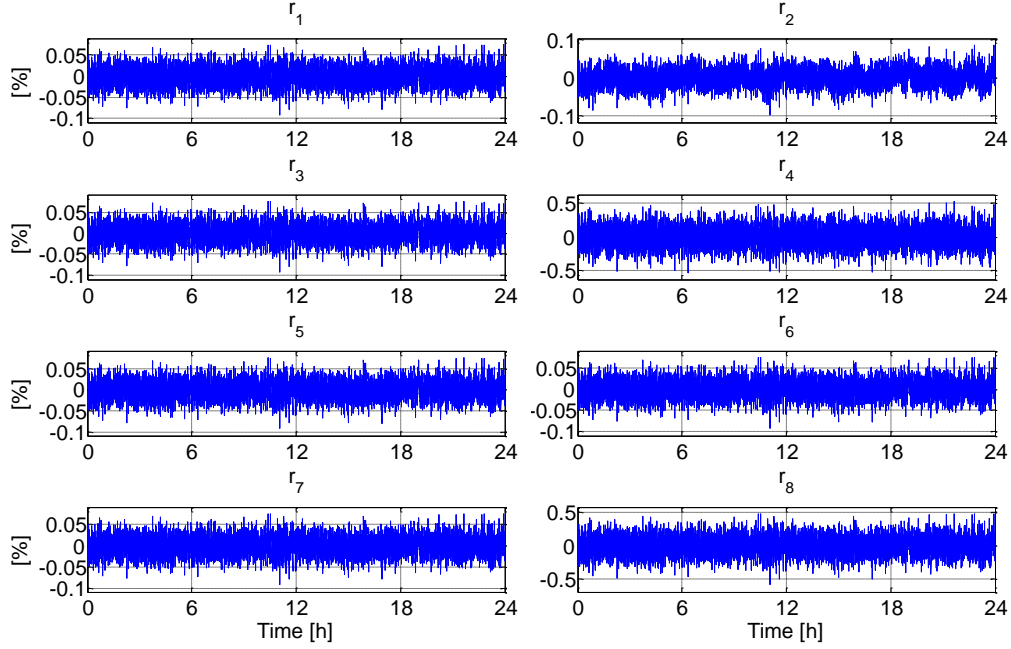
28

Figure 5: Raw residuals behaviour in normal conditions.

for each simulation) are induced after 3 h; at this time, the related fault parameter is gradually brought from 0 to 0.1 in 100 s. Due to the assumption made in Section 2.2 of neglecting detailed models of fault causes, the dynamic effects (i.e., transient behaviours) of fault occurrence are not accounted. Therefore, only the values assumed by the residuals at steady-state after fault induction should be considered for proper fault detection and isolation.

The horizontal dot-dashed black line visible in all frames of Figure 7 represents the threshold level through which fault detection is achieved. This level is set equal to the maximum value of the residual distributions during normal condition increased by a factor of 5.

It is worth observing that both drift factor v in equation (17) and threshold
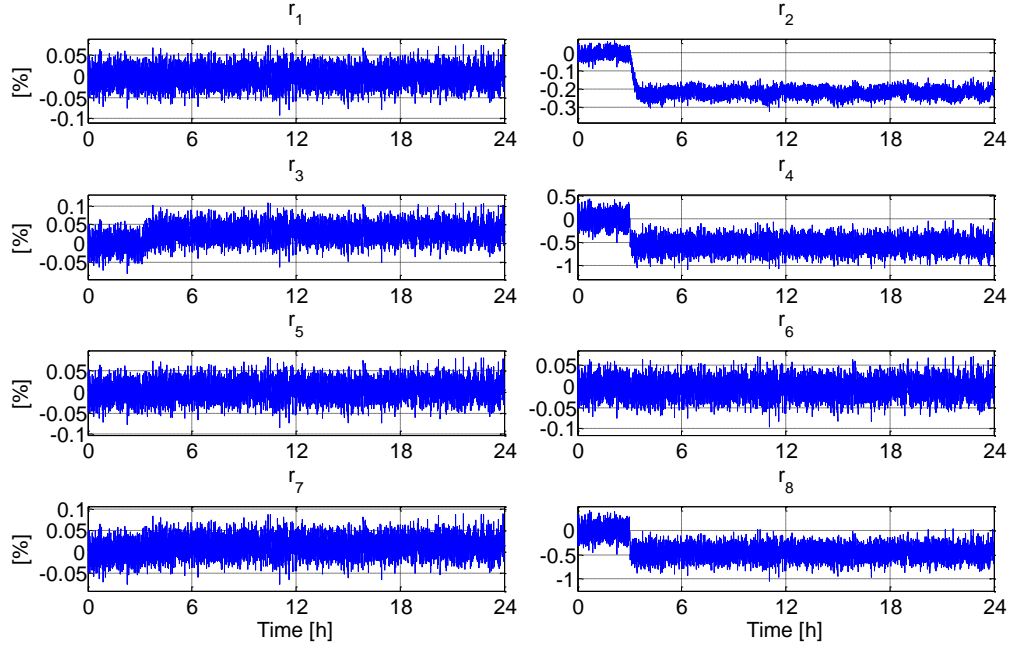
29

Figure 6: Raw residuals behaviour in faulty conditions: example of stack fault f_{ecsa}.

level shall be chosen as a trade-off between incipient fault detection and algorithm robustness [5, 6, 9]: indeed, to achieve the former objective, both drift factor and threshold level should be kept small, but this may lead to an increase in false alarm incidence (i.e., noise may exert a greater influence on CUSUM sequence drift); therefore, to increase algorithm robustness (i.e. reduce false alarm probability) both parameters should be increased, but the probability of missed fault could rise as well.

Once the residuals overcome the threshold level, fault detection is achieved, but correct isolation can be fulfilled only after a settlement period: indeed, as observed in [5, 8, 9], a certain amount of time should pass before performing a correct inference on fault type. For instance, observing the CUSUM sequences
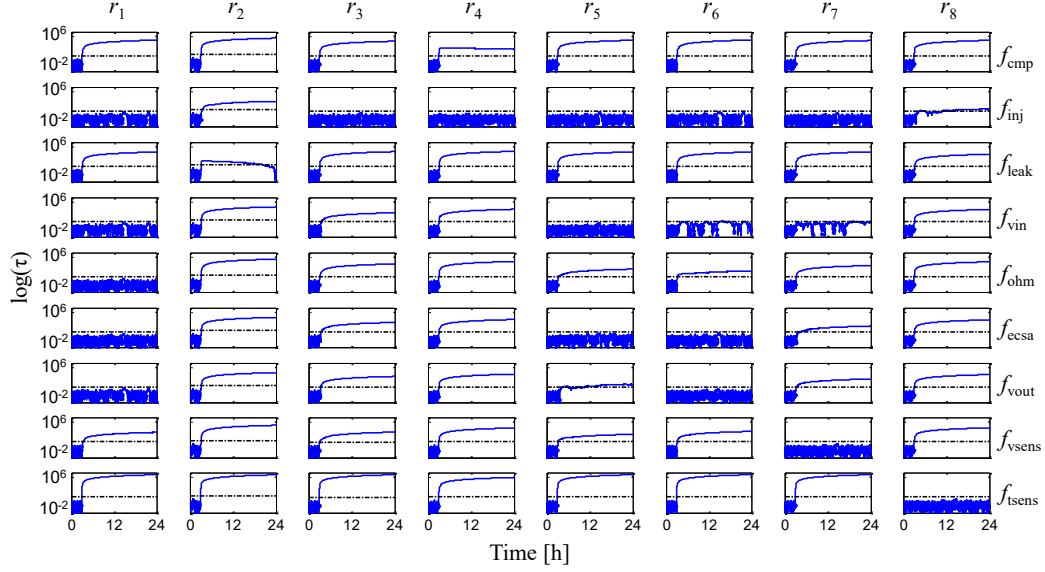
Figure 7: Residuals CUSUM sequence behaviour related to the occurrence of each proposed fault.

during compressor and leakage faults, respectively f_{cmp} and f_{leak}, all residuals drifts from normal behaviour suddenly after 3 h, thus inducing the same residual pattern; however, for fault f_{leak}, residual r_2 goes back below the threshold after about 20 h, changing the pattern and allowing correct isolation with respect to f_{cmp}.

Therefore, fault detection is achieved in all proposed cases nearby the faulty event (with almost no time delay), whereas fault isolation needs a certain time. Notably, the fault requiring the longest settlement time is the leakage, which was proved exerting a small influence especially at small magnitude on Solid Oxide Fuel Cell (SOFC) systems [9]. Through this analysis it is possible to assert that the achieved settlement time is adequate since the analysis has been performed considering incipient faults. Evidently, events with higher magnitude will take shorter time to be both detected and isolated, allowing quicker recovery response.

Table 5: Improved Fault Signature Matrix based on the fault simulation results presented in Figure 7.

	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8
f_{cmp}	1	1	1	1	1	1	1	1
f_{inj}	0	1	0	0	0	0	0	1
f_{leak}	1	0	1	1	1	1	1	1
f_{vin}	0	1	1	1	0	0	0	1
f_{ohm}	0	1	1	1	1	1	1	1
f_{ecsa}	0	1	1	1	0	0	1	1
f_{vout}	0	1	1	1	1	0	1	1
f_{vsens}	1	1	1	1	1	1	0	1
f_{tsens}	1	1	1	1	1	1	1	0

From the results shown in Figure 7, the final diagnostic outcome is summarized in Table 5, where an improved FSM is proposed. This new FSM shows different fault patterns with respect to that of Table 3, since residual sensitivity, fault magnitude and isolation time are now considered (on the same line of what presented in [5, 9]). A significant result consists in the achievement of a FSM with all rows (i.e., fault patterns) different from one another, solving the issue of univocal fault isolability for faults f_{ohm} and f_{ecsa}. However, it should be considered that this result is strictly related to the defined faults magnitude and chosen drift factor and threshold level.

7. Conclusions

This work presented the development of an advanced model-based diagnostic algorithm oriented towards PEMFC systems fault detection and isolation. The

considered diagnostic approach is based on Structural Analysis and Causal Computation theories. One of the most innovative aspects related to this work is the link of such theories to fuel cells applications, which have not yet been thoroughly analysed with the considered approaches.

A lumped parameter model has been conceived so as to simulate an automotive PEMFC system, with attention to reactants transport, electrochemical phenomena, stack thermal dynamics and membrane water transport. Nine different faults have been investigated, involving both auxiliary components, stack and sensors. Moreover, only two sensors have been considered, i.e., stack voltage and temperature. This choice was made to investigate maximum fault isolability with minimal number of sensors, and to test the capability of detecting and isolating faults at system level by only using the stack as sensor. This aspect is of crucial relevance especially for the industrial perspective, since the reduction in on-board sensors availability will lower system manufacturing and deployment costs.

The attained model is a first order differential-algebraic one, with 121 equations and 119 unknown variables, implying a degree of redundancy of 2. This model was first analysed by means of Structural Analysis theory, applying Dulmage-Mendelsohn decomposition, showing only exactly determined and overdetermined parts. Moreover, all the considered faults fell in the overdetermined part, thus fulfilling detectability requirements. Stack faults are belonging to the same equivalence class, so they are not structurally isolable from each other. With respect to all the other faults, Causal Computation analysis has been performed to assess fault isolability upon given causal assumption: the achieved results shown that assuming mixed causality leads to the maximum theoretical isolability, but high-index problems need to be addressed. Therefore, all the considered faults

are structurally detectable with only the stack as system sensor, and only the two stack faults cannot be structurally isolated from each other. This conclusion may imply the need for more advanced measurements (e.g., Electrochemical Impedance Spectroscopy, Current Voltammetry, etc.) to improve fault detection and isolation.

To achieve residuals computation, 8 Minimal Test Equation Support (MTES) sets with redundancy 1 have been computed with the algorithm in [32], 5 out of 8 requiring to face high-index problems, here solved with a PID-based approach. The algorithm has been tested through fault simulation analysis, assuming a fault magnitude of 10%. The acquired residuals have also been treated by means of Cumulative Sum (CUSUM) sequence, to enhance residuals deviation. The achieved results proved univocal fault isolability. Moreover, the need for waiting a certain settlement time was also observed. These results proved the strength of the proposed approach, which can cope with several fault at any fuel cell system components level, by suitably coupling both mathematical and physical information to maximize fault diagnosis.

Finally, the present paper offers a preliminary assessment of the proposed diagnostic algorithm in a test case study, to remark the main issues concerning its development and application. The design procedure adopted for the proposed algorithm showed systematic features, which make it suitable to any system layout, model structure as well as component fault. Therefore, its application to other complex scenarios (e.g., gas recycling or dead-end operation, two-phase flow modelling, membrane crossover fault, etc.) is straightforward.

Acknowledgments

Appendix A: PEMFC model equations

In this section a brief description of the equations characterizing the considered PEMFC system model (whose layout is represented in Figure 1) is given. Model variables definition can be found in the Nomenclature.

A.1. Basic thermodynamics

A dynamic mass balance concerning the i-th species is defined for a j-th control volume as:

$$\frac{dm_{i,j}}{dt} = \dot{m}_{i,j,in} - \dot{m}_{i,j,out} + \dot{m}_{i,j,gen} - \dot{m}_{i,j,cons} \tag{A.1}$$

Then, the species and gas mixture properties are evaluated as follows [29, 30, 31]:

$$p_{i,j} = \frac{m_{i,j}\bar{R}T_j}{M_i V_j} \tag{A.2}$$

$$p_j = \sum_i p_{i,j} \tag{A.3}$$

$$x_{i,j} = \frac{p_{i,j}}{p_j} \tag{A.4}$$

$$M_j = \sum_i x_{i,j}M_i \tag{A.5}$$

$$c_{p_j} = \sum_i x_{i,j}c_{p_i} \tag{A.6}$$

35

$$\phi_j = \frac{p_{H_2O,j}}{p_j^{sat}} \tag{A.7}$$

$$log_{10}\left(p_j^{sat}\right) = -1.69 \cdot 10^{-10}T_j^4 + 3.85 \cdot 10^{-7}T_j^3 - 3.39 \cdot 10^{-4}T_j^2 + 0.143T_j - 20.92 \tag{A.8}$$

$$\omega_j = \frac{p_{H_2O,j}}{p_j - p_{H_2O,j}} \tag{A.9}$$

A.2. Nozzles

According to the ratio between the nozzle downstream and upstream pressures, the flow can be considered un-chocked or chocked if this ratio is respectively higher or lower then the critical pressure ratio δp_c:

$$\delta p_c = \left(\frac{2}{k+1}\right)^{\left(\frac{k}{k-1}\right)} \tag{A.10}$$

Considering the nozzle being located at the j-th component outlet, the flow leaving the component is:

$$\dot{m}_{j,out} = \begin{cases} \frac{C_{d_j}A_j\sqrt{M_j}}{\sqrt{\bar{R}T_j}}p_j\left(\frac{p_{j,down}}{p_j}\right)^{\frac{1}{k}}\sqrt{\frac{2k}{k-1}\left(1-\left(\frac{p_{j,down}}{p_j}\right)^{\frac{k-1}{k}}\right)} & \text{if } \frac{p_{j,down}}{p_j} > \delta p_c \\ \frac{C_{d_j}A_j\sqrt{M_j}}{\sqrt{\bar{R}T_j}}p_j\sqrt{k}\left(\frac{2}{k+1}\right)^{\frac{k+1}{2(k-1)}} & \text{if } \frac{p_{j,down}}{p_j} \leq \delta p_c \end{cases} \tag{A.11}$$

Once the total mass flow $\dot{m}_{j,out}$ is defined, each species flow can be computed as:

$$\dot{m}_{i,j,out} = \dot{m}_{j,out}x_{i,j}\frac{M_i}{M_j} \tag{A.12}$$

A.3. Hydrogen tank and air blower

The storage pressure and temperature represent the upstream conditions for the anode supply manifold. The tank outlet flow is expressed through equation (A.11), and the anode supply manifold species inlet flows are evaluated through equation (A.12), being:

$$\dot{m}_{tank,out} = \dot{m}_{an,sm,in} \tag{A.13}$$

To evaluate the air blower outlet flow and efficiency, two maps are used, shown in Figures A.1-a and A.1-b, respectively, both function of the blower angular speed n_{cmp} and pressure ratio β_{cmp} [30, 31]:

$$\dot{m}_{cmp} = \mathrm{map}(n_{cmp}, \beta_{cmp}) \tag{A.14}$$

$$\eta_{cmp} = \mathrm{map}(n_{cmp}, \beta_{cmp}) \tag{A.15}$$

The blower power and outlet temperature are computed as:

$$P_{cmp} = \dot{m}_{cmp} \frac{c_{p_{air}} T_{amb}}{\eta_{cmp} \eta_{em}} \left(\beta_{cmp}^{\frac{k-1}{k}} - 1 \right) \tag{A.16}$$

$$T_{cmp} = T_{amb} \left(1 + \frac{1}{\eta_{cmp}} \left(\beta_{cmp}^{\frac{k-1}{k}} - 1 \right) \right) \tag{A.17}$$

To ensure the correct evaluation of the species outlet flows upon the knowledge of $\dot{m}_{cmp}$, the following assumption is made: the ratio between the water and dry air masses does not change from blower inlet to outlet, implying the specific humidity conservation [30, 31]:

$$\omega_{cmp} = \omega_{amb} \tag{A.18}$$

From the previous statement, it is possible to write the blower outlet species mass flows as:

$$\dot{m}_{O_2,cmp} = \frac{M_{O_2}}{M_{air}} \left(\frac{\dot{m}_{cmp}}{1 + \omega_{cmp}} \right) x_{O_2,amb} \tag{A.19}$$
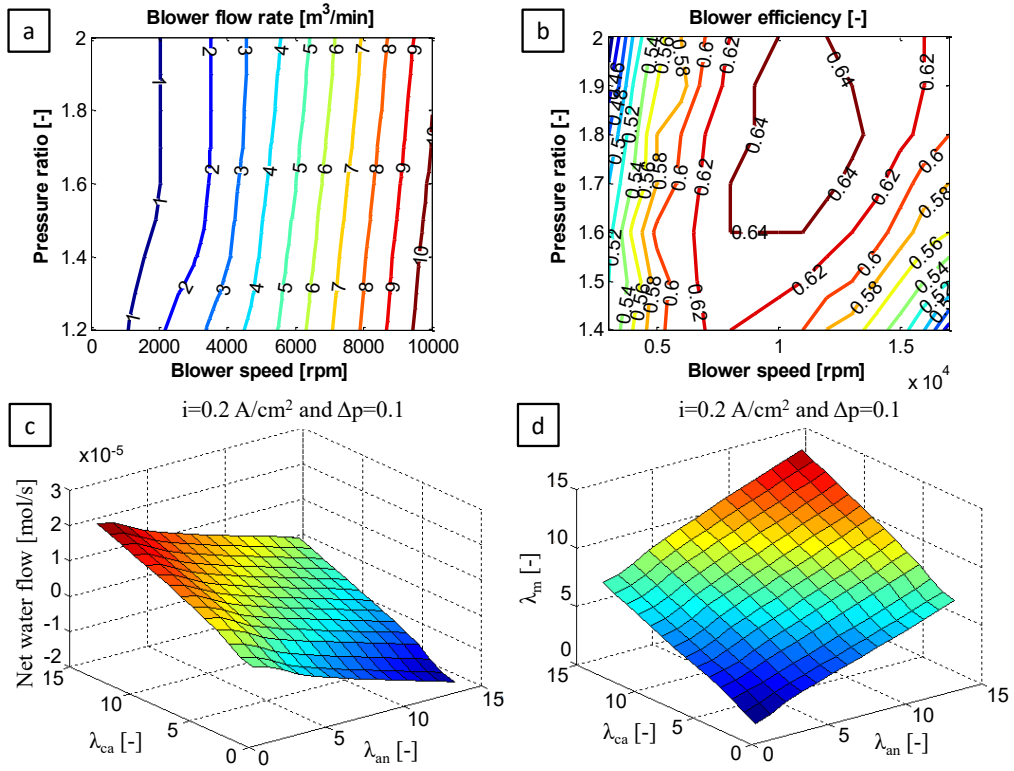
37

Figure A.1: Air flow rate (a) and efficiency (b) maps of the volumetric blower, and electrolyte net water flow (c) and water content (d) maps, adapted from [30, 31].

$$\dot{m}_{N_2,cmp} = \frac{M_{N_2}}{M_{air}} \left(\frac{\dot{m}_{cmp}}{1 + \omega_{cmp}} \right) x_{N_2,amb} \tag{A.20}$$

$$\dot{m}_{H_2O,cmp} = \dot{m}_{cmp} - \dot{m}_{O_2,cmp} - \dot{m}_{N_2,cmp} \tag{A.21}$$

A.4. Supply manifolds

The mass balance proposed in equation (A.1) is here applied neglecting the consumption term. The generation term holds only for water balance, due to the injection required for flow humidification. The anode supply manifold inlet flow is evaluated with equations (A.12) and (A.13), whereas the cathode supply manifold species inlet flow with equations (A.19) through (A.21), being:

$$\dot{m}_{i,ca,sm,in} = \dot{m}_{i,cmp} \tag{A.22}$$

Water humidification is achieved through direct water injection. This term represents the generation term for the water mass balance at supply manifolds:

$$\dot{m}_{H_2O,j,inj} = \frac{m_{H_2O,j,des}}{\tau_{inj}} \tag{A.23}$$

where:

$$m_{H_2O,j,des} = \frac{p_{j,sat} V_j M_{H_2O}}{\bar{R} T_j} \left(\phi_{j,des} - \phi_j \right) \geq 0 \tag{A.24}$$

The outlet flows from both supply manifolds are evaluated through equation (A.11). Consequently, the electrode inlet flow components are computed by means of equation (A.12), being:

$$\dot{m}_{i,an,in} = \dot{m}_{i,an,sm,out} \tag{A.25}$$

$$\dot{m}_{i,ca,in} = \dot{m}_{i,ca,sm,out} \tag{A.26}$$

A.5. PEMFC Electrodes and Electrolyte

The mass balance in equation (A.1) applies to both cathode and anode, assuming that the consumption terms appear only with respect to the hydrogen and oxygen mass balances [2]:

$$\dot{m}_{H_2,an,cons} = \frac{I}{2F} M_{H_2} \tag{A.27}$$

$$\dot{m}_{O_2,ca,cons} = \frac{I}{4F} M_{O_2} \tag{A.28}$$

whereas the generation term appears only in the water balance at cathode side:

$$\dot{m}_{H_2O,ca,gen} = \frac{I}{2F} M_{H_2O} \tag{A.29}$$

The inlet flows coming from the supply manifolds are evaluated by means of equations (A.25) and (A.26) and the outlet flows through equations (A.11) and (A.12). In the electrodes mass balance, the net water flow through the electrolyte from one electrode to the other should be also considered. This term represents the overall water flow achieved by combining the three main water transport mechanisms involving the electrolyte, that are electro-osmotic drag, convection and back-diffusion [2, 45]. This term is computed by means of the numerical map proposed in Figure A.1-c, assuming as positive flux the one going from cathode to anode. The other map shown in Figure A.1-d allows the evaluation of the water content λ_m of the polymeric membrane, here assumed being a Nafion® 117 type, essential for the evaluation of the fuel cell voltage behaviour. Both maps, derived from [45] and already used in [30, 31], are function of the average water content at cathode and anode sides (λ_{ca} and λ_{an}, respectively), pressure difference across the electrolyte $\Delta p = p_{ca} - p_{an}$ and current density i, expressed as the ratio between the

current request I and the fuel cell geometrical area A_{fc}:

$$i = \frac{I}{A_{fc}} \tag{A.30}$$

$$\text{Net water flow} = \text{map}(\lambda_{ca}, \lambda_{an}, \Delta p, i) \tag{A.31}$$

$$\lambda_m = \text{map}(\lambda_{ca}, \lambda_{an}, \Delta p, i) \tag{A.32}$$

The j-th electrode water content is function of the relative humidity level at the corresponding electrode [31]:

$$\lambda_j = \begin{cases} 0.3 + 10.8\phi_j - 16\phi_j^2 + 14.1\phi_j^3 & \text{if } \phi_j \geq 0 \ \& \ \phi_j < 1 \\ -30.41 + 61.98\phi_j - 25.96\phi_j^2 + 3.7\phi_j^3 & \text{if } \phi_j \geq 1 \end{cases} \tag{A.33}$$

A.6. Thermal balance

An overall dynamic thermal balance model is implemented for the fuel cell stack to simulate the dynamic behaviour of the bulk temperature. The balance is based on the total energy amount of the inlet and outlet gases, the produced electric power and the heat exchanged with the surrounding, as done in [5, 46]:

$$K_{fc}\frac{dT_{fc}}{dt} = \sum_{i,j} \left[\dot{E}_{i,j,in}\left(T_{j,in}\right) - \dot{E}_{i,j,out}\left(T_{fc}\right) \right] - V_{fc}I - \dot{Q} \tag{A.34}$$

The energy flows are function of the species mass flow and enthalpy as follows:

$$\dot{E}_{i,j,in} = \dot{m}_{i,j,in}h_i\left(T_{j,in}\right) \tag{A.35}$$

$$\dot{E}_{i,j,out} = \dot{m}_{i,j,out}h_i\left(T_{fc}\right) \tag{A.36}$$

$$h_i(T) = h_i^f\left(T_0\right) + \int_{T_0}^{T} c_{pi}dT \tag{A.37}$$

A.7. Electrochemical model

The performance of the single PEMFC is represented by its voltage behaviour V_{fc}, computed defining the ideal Nernst potential and the different losses as [2, 10, 31, 47, 48] :

$$V_{fc} = E_N - V_{Act} - V_{Ohm} - V_{Dif} \tag{A.38}$$

$$E_N = 1.23 - 0.9 \cdot 10^{-3} \left(T_{fc} - 298 \right) + \frac{T_{fc}\bar{R}}{2F} \ln \left(p_{H_2,an} p_{O_2,ca}^{0.5} \right) \tag{A.39}$$

$$V_{Act} = \frac{T_{fc}\bar{R}}{2\alpha F} \ln \left(\frac{i}{i_0} \right) \tag{A.40}$$

$$i_0 = i_0^{Pt} ECSA \tag{A.41}$$

$$V_{Ohm} = \frac{t_m}{\sigma_m} i \tag{A.42}$$

$$\sigma_m = (0.005139\lambda_m - 0.00326) \exp\left[1268 \left(\frac{1}{303} - \frac{1}{T_{fc}} \right) \right] \tag{A.43}$$

$$V_{Dif} = i\tilde{\omega} T_{fc} \ln \left(\frac{i_{lim}}{i_{lim} - i} \right) \tag{A.44}$$

$$i_{lim} = -\frac{2FD_{O_2N_2}}{V_m t_{ca}} \varepsilon_{ca}^{1.5} \left(\frac{T_{fc}}{273} \right)^{0.823} ln(1 - x_{O_2,ca}) \tag{A.45}$$

A.8. Sensors

The equations representing the stack voltage and temperature measurements are:

$$y_V = V_{fc} \tag{A.46}$$

$$y_T = T_{fc} \tag{A.47}$$

Nomenclature

Acronyms

CUSUM	Cumulative Sum
FSM	Fault Signature Matrix
MTES	Minimal Test Equation Support
MVM	Mean Value Modelling
PEMFC	Polymer Electrolyte Membrane Fuel Cell
SOFC	Solid Oxide Fuel Cell

Roman symbols

A	Area [m^2]
C_d	Nozzle discharge coefficient [-]
c_p	Specific heat at constant pressure [$J\,kg^{-1}\,K^{-1}$]
$D_{O_2N_2}$	Binary diffusion coefficient of O_2 and N_2 [$m^2\,s^{-1}$]
E_N	Nernst potential [V]
$\dot{E}$	Energy flow [W]
$ECSA$	Electrochemical Surface Area [m^{-2}]
F	Faraday constant (96485 [$C\,mol^{-1}$])
f	Fault coefficient [-]
h	Specific enthalpy [$J\,kg^{-1}$]
h^f	Specific enthalpy of formation at T_0=298.15 K [$J\,kg^{-1}$]
I	Current [A]
i	Current density [$A\,m^{-2}$]
i_0	Exchange current density [$A\,m^{-2}$]
i_{lim}	Limiting current density [$A\,m^{-2}$]
K	Heat capacity [$J\,K^{-1}$]
k	Polytropic constant [-]

M	Molar mass [J mol^{-1} K^{-1}]
m	Mass [kg]
$\dot{m}$	Mass flow [kg s^{-1}]
n	Angular speed [rpm]
P	Power [W]
p	Pressure [bar]
$\dot{Q}$	Heat exchanged [W]
$\bar{R}$	Universal gas constant (8.314 [J mol^{-1} K^{-1}])
r	Residual [-]
T	Temperature [K]
t	Time [s]
t_{ca}	Cathode thickness [m]
t_m	Membrane thickness [m]
V	Volume [m^{-3}]
V_{Act}	Activation losses [V]
V_{fc}	Fuel cell voltage [V]
V_{Dif}	Diffusion losses [V]
V_m	Standard molar volume (0.022414 [m^3 mol^{-1}])
V_{Ohm}	Ohmic losses [V]
x	Molar fraction [-]
y_V	Voltage measurement [V]
y_T	Temperature measurement [T]
Greek symbols	
α	Charge transfer coefficient [-]
β	Pressure ratio [-]

δp_c	Critical pressure ratio [-]
ε	Porosity [-]
η	Efficiency [-]
λ	Water content [-]
ν	Drift factor [-]
ϕ	Relative humidity [-]
σ	Conductivity [Ω^{-1} m^{-1}]
τ	CUSUM residual [-]
τ_{inj}	Injection time [s]
ω	Absolute humidity [-]
$\tilde{\omega}$	Empirical constant [Ω m^2 K^{-1}] (from [47])
Subscripts	
amb	ambient
air	air
an	anode
an, sm	anode supply manifold
ca	cathode
ca, sm	cathode supply manifold
cmp	air blower
cons	consumed
des	desired
down	downstream
em	electric motor
f c	fuel cell
gen	generated

i	*i*-th species (O_2, N_2 and H_2O at cathode side, H_2 and H_2O at anode side)
in	inlet
inj	injected
j	*j*-th control volume (*amb*, *an*, *an*,*sm*, *ca*, *ca*,*sm*, *cmp* and *tank*)
leak	leakage
m	membrane
out	outlet
tank	hydrogen tank
tsens	temperature sensor
vin	inlet valve
vout	outlet valve
vsens	voltage sensor
Superscripts	
F	faulty
Pt	catalyst surface
sat	saturation

References

[1] F. Barbir, PEM Fuel Cells, Theory and Practice, 2nd Edition, Academic Press, 2013.

[2] J. Larminie, A. Dicks, Fuel Cell Systems Explained, 2nd Edition, Wiley, 2003.

[3] U. S. Department of Energy, Fuel cell technologies office multi-year re-

search, development, and demonstration plan – 3.4 fuel cells, 2016. (accessed september 2016).

[4] Fuel Cells and Hydrogen Joint Undertaking., Multi-Annual Work Plan 2014-2020. http://www.fch.europa.eu/page/multi-annual-work-plan (accessed april 2017).

[5] P. Polverino, C. Pianese, M. Sorrentino, D. Marra, Model-based development of a fault signature matrix to improve solid oxide fuel cell systems on-site diagnosis, Journal of Power Sources 280 (2015) 320–338.

[6] D. Marra, C. Pianese, P. Polverino, M. Sorrentino, Models for Solid Oxide Fuel Cell Systems, Exploitation of Model Hierarchy for Industrial Design of Control and Diagnosis Strategies, Springer, 2016.

[7] X. Wu, B. Zhou, Fault tolerance control for proton exchange membrane fuel cell systems, Journal of Power Sources 324 (2016) 804–829.

[8] T. Escobet, D. Feroldi, S. de Lira, V. Puig, J. Quevedo, J. Riera, M. Serra, Model-based fault diagnosis in pem fuel cell systems, Journal of Power Sources 192 (2009) 216–223.

[9] P. Polverino, A. Esposito, C. Pianese, B. Ludwig, B. Iwanschitz, A. Mai, Online experimental validation of a model-based diagnostic algorithm dedicated to a solid oxide fuel cell system, Journal of Power Sources 306 (2016) 646–657.

[10] P. Polverino, C. Pianese, Model-based prognostic algorithm for online rul estimation of pemfcs, 2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol), Barcelona, Spain, Sept. 7-9, 2016 (2016) 599–604.

47

[11] R. Isermann, Fault-Diagnosis Systems, An Introduction from Fault Detection to Fault Tolerance, Springer, 2006.

[12] J. Liu, W. Luo, X. Yang, L. Wu, Robust model-based fault diagnosis for pem fuel cell air-feed system, IEEE Transactions on Industrial Electronics 63 (2016) 3261–3270.

[13] D. Rotondo, R. Fernandez-Canti, S. Tornil-Sin, J. Blesa, V. Puig, Robust fault diagnosis of proton exchange membrane fuel cells using a takagi-sugeno interval observer approach, International Journal of Hydrogen Energy 41 (2016) 2875–2886.

[14] A. Rosich, R. Sarrate, F. Nejjari, On-line model-based fault detection and isolation for pem fuel cell stack systems, Applied Mathematical Modelling 38 (2014) 2744–2757.

[15] A. Esposito, L. Russo, C. Kaendler, C. Pianese, B. Ludwig, N. Yousfi-Steiner, High fuel utilization in solid oxide fuel cells: Experimental characterization and data analysis with continuous wavelet transform, Journal of Power Sources 317 (2016) 159–168.

[16] E. Pahon, N. Yousfi-Steiner, S. Jemei, D. Hissel, P. Mocoteguy, A signal-based method for fast pemfc diagnosis, Applied Energy 165 (2016) 748–758.

[17] M. Ibrahim, U. Antoni, N. Yousfi-Steiner, S. Jemei, C. Kokonendji, B. Ludwig, P. Mocoteguy, D. Hissel, Signal-based diagnostics by wavelet transform for proton exchange membrane fuel cell, Energy Procedia 74 (2015) 1508–1516.

[18] C. Damour, M. Benne, B. Grondin-Perez, M. Bessafi, D. Hissel, J. Chabriat, Polymer electrolyte membrane fuel cell fault diagnosis based on empirical mode decomposition, Journal of Power Sources 299 (2015) 596–603.

[19] A. Escobet, A. Nebot, F. Mugica, Pem fuel cell fault diagnosis via a hybrid methodology based on fuzzy and pattern recognition techniques, Engineering Applications of Artificial Intelligence 36 (2014) 40–53.

[20] M. Blanke, M. Kinnaert, J. Lunze, M. Staroswiecki, Diagnosis and Fault-Tolerant Control, Springer, 2006.

[21] R. Isermann, Fault-Diagnosis Applications. Model-based condition monitoring: Actuators, drivers, machinery, plants, sensors, and fault-tolerant systems, Springer, 2011.

[22] C. Svard, M. Nyberg, E. Frisk, M. Krysander, Automotive engine fdi by application of an automated model-based and data-driven design methodology, Control Engineering Practice 21 (2013) 455–472.

[23] M. Krysander, J. Aslung, M. Nyberg, An efficient algorithm for finding minimal overconstrained subsystems for model-based diagnosis, IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans 38 (2008) 197–206.

[24] C. Sundstrom, E. Frisk, L. Nielsen, Selecting and utilizing sequential residual generators in fdi applied to hybrid vehicles, IEEE Transactions on Systems, Man and Cybernetics: Systems 44 (2014) 172–185.

[25] Q. Chen, Q. Ahmed, G. Rizzoni, M. Qiu, Desing and evaluation of model-based health monitoring scheme for automated manual transmission, Journal of Dynamic Systems, Measurement, and Control 138 (2016) 101011–1/10.

[26] D. Vey, J. Lunze, Structural reconfigurability analysis of multirotor uavs after actuator failures, 2015 IEEE 54th Annual Conference on Decision and Control (CDC), December 15-18, 2015, Osaka, Japan.

[27] E. Frisk, A. Bregon, J. Aslund, M. Krysander, B. Pulido, G. Biswas, Diagnos-ability analysis considering causal interpretation for differential constraints, IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans 42 (2012) 1216–1229.

[28] A. Rosich, E. Frisk, J. Aslund, R. Sarrate, F. Nejjari, Fault diagnosis based on causal computations, IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans 42 (2012) 371–381.

[29] J. T. Pukrushpan, A. G. Stefanopoulou, H. Peng, Control of Fuel Cell Power Systems, Principles, Modeling, Analysis and Feedback Design, Springer, 2005.

[30] I. Arsie, A. Di Domenico, C. Pianese, M. Sorrentino, Modeling and ana-lysis of transient behavior of polymer electrolyte membrane fuel cell hybrid vehicles, ASME Journal of Fuel Cell Science and Technology 4 (2007) 261–271.

[31] I. Arsie, A. Di Domenico, C. Pianese, M. Sorrentino, A multilevel approach to the energy management of an automotive polymer electrolyte membrane

fuel cell system, ASME Journal of Fuel Cell Science and Technology 7 (2010) 011004/1–11.

[32] M. Krysander, J. Aslund, E. Frisk, A structural algorithm for finding testable sub-models and multiple fault isolability analysis, 21st International Workshop on Principles of Diagnosis (DX-10), Portland, Oregon, USA (2010).

[33] P. Polverino, A. Esposito, C. Pianese, Experimental validation of a lumped model of single droplet deformation, oscillation and detachment on the gdl surface of a pem fuel cell, International Journal of Hydrogen Energy 38 (2013) 8934–8953.

[34] A. Esposito, P. Polverino, C. Pianese, Y. Guezennec, A lumped model of single droplet deformation, oscillation and detachment on the gdl surface of a pem fuel cell, in: Proceedings of FuelCell2010 Eighth International Fuel Cell Science, Engineering and Technology Conference June 14-16, 2010 Brooklyn, New York, USA, 2010.

[35] A. Esposito, C. Pianese, Y. Guezennec, Coupled modeling of water transport and air–droplet interaction in the electrode of a proton exchange membrane fuel cell, Journal of Power Sources 195 (2010) 4149–4159.

[36] D. Jung, A generalized fault isolability matrix for improved fault diagnosability analysis, 2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol), Barcelona, Spain, Sept. 7-9, 2016 (2016) 504–509.

[37] D. Capriglione, C. Liguori, C. Pianese, A. Pietrosanto, On-line sensor fault detection, isolation, and accomodation in automotive engines, IEEE Instru-

mentation and Measurement Technology Conference, Anchorage, AK, USA 21-23 May 2002.

[38] C. Svärd, M. Nyberg, Residual generators for fault diagnosis using computation sequences with mixed causality applied to automotive systems, IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans 40 (6) (2010) 1310–1328.

[39] K. E. Brenan, S. L. Campbell, L. R. Petzold, Numerical solution of initial-value problems in differential-algebraic equations, Vol. 14, Siam, 1996.

[40] A. Dulmage, N. Mendelsohn, Coverings of bipartite graphs, Canadian Journal of Mathematics 10 (1958) 516–534.

[41] T. Jahnke, G. Futter, A. Latz, T. Malkow, G. Papakonstantinou, G. Tsotridis, P. Schott, M. Gèrard, M. Quinaud, A. Quiroga, M. Franco, K. Malek, F. Callevallejo, T. Ferreira de Morais, R. Kerber, P. Sautet, D. Loffreda, S. Strahl, M. Serra, P. Polverino, C. Pianese, M. Mayur, W. Bessler, C. Kompis, Performance and degradation of proton exchange membrane fuel cells: State of the art in modeling from atomistic to system scale, Journal of Power Sources 304 (2016) 207–233.

[42] K. Murota, Matrices and matroids for systems analysis, Vol. 20, Springer Science & Business Media, 2009.

[43] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II, Vol. 14 of Springer Series in Computational Mathematics, Springer-Verlag Berlin Heidelberg, 1996.

[44] E. Page, Continuous inspection schemes, Biometrika 41 (1954) 100–115.

[45] S. Mazumder, A generalized phenomenological model and database for the transport of water and current in polymer electrolyte membranes, Journal of Electrochemical Society 152 (2005) A1633–A1644.

[46] S. Strahl, A. Husar, P. Puleston, J. Riera, Performance improvement by temperature control of an open-cathode pem fuel cell system, Fuel Cells 14 (2014) 466–478.

[47] G. Maggio, V. Recupero, L. Pino, Modeling polymer electrolyte fuel cells: an innovative approach, Journal of Power Sources 101 (2001) 275–286.

[48] T. E. Springer, T. A. Zawodzinski, S. Gottesfeld, Polymer electrolyte fuel cell model, Journal of Electrochemical Society 138 (1991) 2334–2342.