REAL-TIME MACHINE LEARNING BASED OPEN SWITCH FAULT DETECTION AND ISOLATION FOR MULTILEVEL MULTIPHASE DRIVES

A Dissertation

by

NILOOFAR TORABI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Hamid A. Toliyat
Committee Members,	Prasad Enjeti
	Shankar P. Bhattacharyya
	Behbood B. Zoghi
Head of Department,	Miroslav Begovic

May 2018

Major Subject: Electrical Engineering

ABSTRACT

Due to the rapid proliferation interest of the multiphase machines and their combination with multilevel inverters technology, the demand for high reliability and resilient in the multiphase multilevel drives is increased. High reliability can be achieved by deploying systematic preventive real-time monitoring, robust control, and efficient fault diagnosis strategies. Fault diagnosis, as an indispensable methodology to preserve the seamless post-fault operation, is carried out in consecutive steps; monitoring the observable signals to generate the residuals, evaluating the observations to make a binary decision if any abnormality has occurred, and identifying the characteristics of the abnormalities to locate and isolate the failed components. It is followed by applying an appropriate reconfiguration strategy to ensure that the system can tolerate the failure.

The primary focus of presented dissertation was to address employing computational and machine learning techniques to construct a proficient fault diagnosis scheme in multilevel multiphase drives. First, the data-driven nonlinear model identification/prediction methods are used to form a hybrid fault detection framework, which combines module-level and system-level methods in power converters, to enhance the performance and obtain a rapid real-time detection. Applying suggested nonlinear model predictors along with different systems (conventional two-level inverter and three-level neutral point clamped inverter) result in reducing the detection time to 1% of stator current fundamental period without deploying component-level monitoring equipment. Further, two methods using semi-supervised learning and analytical data mining concepts

are presented to isolate the failed component. The semi-supervised fuzzy algorithm is engaged in building the clustering model because the deficient labeled datasets (prior knowledge of the system) leads to degraded performance in supervised clustering. Also, an analytical data mining procedure is presented based on data interpretability that yields two criteria to isolate the failure. A key part of this work also dealt with the discrimination between the post-fault characteristics, which are supposed to carry the data reflecting the fault influence, and the output responses, which are compensated by controllers under closed-loop control strategy. The performance of all designed schemes is evaluated through experiments.

DEDICATION

I would like to dedicate this dissertation to my loving mother, Fereshteh Faramarz, and my wonderful father, Ghasem Torabi, who always believed in me and supported my intellectual and emotional growth. Their unconditional love and continuous encouragement have been an invaluable source of inspiration at every step of my life.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude and appreciation for my adviser and committee chair, Professor Hamid Toliyat, for giving me an incredible opportunity to work in his group, for his vision, leadership, and continuous support. Not only did I learn everything technical I know from him, but I learned to thrive on solving problems, to enjoy and be proud of being an engineer. Professor Toliyat elevated the quality and accuracy of my work enormously at each step and level of my research including this thesis. I will always be grateful and indebted to him for all I learned from him. I consider him an exceptional mentor, a great teacher, and an excellent researcher and I wish to follow his example in my future career. I am also grateful to the rest of my advisory committee members, Professor Prasad Enjeti, Professor Shankar Bhattacharyya, and Professor Behbood B. Zoghi, for their invaluable feedback, guidance, and insight throughout this research study.

I am very thankful to my colleagues in the Advanced Electric Machines and Power Electronics Laboratory at Texas A&M University. I like to especially thank Farid Naghavi and my former colleague Dr. Vivek Sundaram. It was a delight to work with both of them. Their help during my Ph.D. studies was crucial and is truly appreciated. I am also very grateful to all my friends, classmates, and Electrical Engineering department faculty and staff for making my time at Texas A&M University a pleasant experience.

Finally, I am immensely indebted to my family for their encouragement and sacrifices throughout my life, their love, help, and unwavering support. I owe everything

I am and all I have to my parents, Fereshteh Faramarz and Ghasem Torabi. My mother always was and will be my role model in my life. She is all that I ever dream to achieve both at the moral and personal level as a great human being. I am grateful to my father, for supporting me and believing in me throughout my life, for teaching me how to love, for picking me up every time I fell, for always telling me not to be scared, for giving me my self-confidence. I would like to thank my dear brother Nima, and my adorable sister Nikta, who are the best siblings one can ever wish for. They believed in me and motivated me. None of this would have happened without their love and support.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Hamid Toliyat, Professor Prasad Enjeti, and Professor Shankar P. Bhattacharyya of the Department of Electrical and Computer Engineering and Professor Behbood B. Zoghi of the Department of Engineering Technology and Industrial Distribution.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a research assistantship from Texas A&M University.

There are no outside funding contributions to acknowledge related to the research and compilation of this document.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1 Background

The concept of the multiphase machine, as an envisioned technology, has been investigated, and employed widely in recent years. Due to the significant advantages such as improving efficiency, reducing the torque pulsation, reducing the rotor and dc link current harmonic, reducing the current per phase without increasing the voltage per phase, reducing the stress per phase, improving the torque per ampere, and increasing power for the same volume machine in the same frame, multiphase electric machines have been recognized as an attractive alternative to the conventional three-phase machines. Also, noise characteristics enhancement and loss reduction are the other benefits of multiphase systems. The additional degrees of freedom in multiphase machines enable injecting harmonic currents or supply multi motors from the power converter as well. Besides, improvement of reliability and higher fault tolerance capability are achieved in multiphase electric machines.

Potential benefits of multiphase machines have led to substantial developments. The high-phase order drives are mostly engaged to particular applications with higher demand for reliability such as electric and hybrid vehicles, aerospace applications, electric ship propulsion, electric locomotive traction, electric submarines, drones, and formula-E car racing [1-2].

In multiphase systems, it is requisite that power rating of the converter can meet the required power level of the machine at driven full load. Nevertheless, it is not possible

to increase the converter ratings over a specific range because the power rating of their semiconductor devices is limited. Hence, multilevel multiphase converters have been brought into play to develop the high power level converters using lower rated semiconductor components. The multilevel converters have been employed broadly for high-voltage high-power applications. They offer high power density, low voltage stress on devices, lower total harmonic characteristics, and lower electromagnetic interference noise in power systems [3].

The multilevel inverter concept was introduced in the early 70s with three-level inverters and perceived as an attractive solution for high-level voltage and high-level and medium-level power because of their beneficial specifications compared with the conventional two-level counterparts. Multilevel power converters, developed with arrays of lower rated power semiconductors and capacitors, allow the generating of high-quality load voltage at the high power level. Significant advantages presented by multilevel converters, focusing on improvements in the output quality and increasing their nominal power, are waveform quality, lower common-mode voltage and current with low distortion and low switching frequency [4].

Also, the presence of a higher number of levels is associated with a higher number of redundant states, which offers an essential property for the operation without significant performance deterioration in the equipment under the fault condition. However, the superiority of employing multiphase machines vs. multilevel power converters in high power applications is still open to discussion.

There have been extensive works on upgrading multilevel inverter circuit topologies. The most popular families of multilevel converters are neutral point clamped (NPC), flying capacitor, and cascaded H-bridge converters. Recently, a new three-level T-type converter was proposed for high-efficiency low-voltage applications as well. T-type is the combination of two-level converter and three-level diode clamped inverters. The latest multilevel converter topologies have addressed lower switching loss and improved power quality while they avoid series switching.

Nevertheless, there are some concerns in multilevel inverter applications. The primary concerns in operating the multilevel converters are the capacitor voltage-balancing strategies, suppressing methods of circulating current, and their reliability. Generating high-quality electricity, which has become an important topic in recent years, necessitates guaranteeing optimal performance of multilevel inverters.

1.2 Background on Power Inverters Fault Diagnosis

Reliability, performance efficiency, security, and maintainability are vital requirements in multiphase motor drives. Reliability can be achieved by deploying systematic preventive real-time monitoring, control, and fault diagnosis strategies. Fault diagnosis, as a control methodology preserving the continued reliable operation of a system after fault occurrence, is carried out in three consecutive steps. Firstly, fault detection process, as a binary decision-making function, determines if any fault has occurred. Secondly, the fault isolation scheme identifies the fault characteristics and

locates the failed device. It is followed by the third step which is applying an appropriate reconfiguration strategy to ensure that the system can tolerate the failure [5-6].

The growing complexity of electric drives and increasing demand for their reliability requires improving the fault tolerance of different equipment. Therefore, the earliest possible fault diagnosis including fault detection, fault characteristic identification, and fault isolation is necessary for such systems to ensure their reliable operation.

By and large, the fault diagnosis scheme is an indispensable part of multiphase electric drives, because having a higher number of semiconductors eventuates in a dramatic increase in the probability of failure in the entire system. Besides, due to the high number of components, locating a failed switch is more problematic in principle [7]. Also, in multilevel converters, due to the higher number of switches, each switch is conducting in shorter time interval over every cycle. Therefore, the waveforms are influenced by the component failure in a relatively shorter period. Additionally, the higher number of switches results in higher number of redundant states, which makes locating the failed switch more challenging.

Therefore, detection and isolation of the fault and continuing the operation under failure conditions are of supreme importance. In this regard, fault diagnosis of various electric machines and drives have been investigated [8-9].

A wide range of different electric drives Fault detection and isolation (FDI) techniques have been investigated and developed in recent years. FDI techniques for electric drives can be classified into two primary categories. These include component-

based (component-level) and system-based (system-level) techniques [10]. In component-based FDI techniques, each component must be monitored individually which necessitates deploying extra hardware. Although such techniques are reputed to be fast and reliable, they involve higher cost as well as the additional required space to accommodate the hardware equipment. On the other hand, System-based techniques, which are divided into types (waveform analysis methods and artificial intelligence methods) evaluate the output waveforms to identify the characterizations of occurred faults. In system-based methods, fault indicators (residuals) are generated by comparison of observed waveforms with predicated waveforms. Subsequently, any large deviation from permitted value will be interpreted as fault incident and initiates the fault isolation and reconfiguration schemes [11].

System-based techniques, inspired by the analysis of the output waveforms distortion, has resulted in proposing methods such as slope method, current vector method using Park's transformation, and normalized average current method [12-13]. These methods require the evaluation of at least one fundamental period of measured signals. However, it is desired to detect the fault occurrence as early as possible to prevent consequential damages and to ensure the reliability of the system.

In [14], the slope method is applied in grid-connected converters and back-to-back NPC converters with wind turbine application, whereas, due to the similarity of distortion between two upper switch faults or between two lower switch faults in grid side inverter, the exact faulty switch is not detectable. However, slope method can isolate the failed switch on the generator side converter. It is shown in [15] that injecting current makes it

possible to discriminate between two switches in a pair. In contrast to the grid-connected converters, the output signals distortion is for different switches in motor drive applications. In [16-17], average current Park's vector is used to identify the failed switch. The presented real-time detection method in [18] has decreased the detection time considerably by analyzing the voltage waveforms and recognizing the possible switching states. Nevertheless, this method can detect and isolate the single and the multiple faults in switches and clamped diodes for ac drive applications, the complexity of multiphase drives makes is inexorable. Also, evaluating the pole voltage and instant voltage error is applied for IGBTs and Sic-MOSFETs in [19-21], respectively.

Artificial intelligence methods have also developed for fault diagnosis and prognosis in electric drives. However, extracting the patterns, selecting the features, and labeling the data is expensive due to the use of field knowledge. Therefore, the offline training process is their main drawbacks [22-23]. In [24], a system identification method is proposed which the current is estimated using the Recursive Least Square (RLS) technique. The input of RLS algorithm is constructed by a set of functions varied with angular speed. A nonlinear system identification method originated from the concept of model-based FDI, is suggested in [25-26] as well.

Frequent semiconductor switches faults are classified into two categories: short circuit (SC) switch faults and open circuit (OC) switch faults. Overvoltage, high temperature, and inappropriate gate drive signal lead to short circuit failure. Usually, internal protection circuits of the modules detect the short circuit faults and trip the system. On the other hand, thermal cycling, excessive collector current, and losing of the gate

drive signal result in open circuit failure [27]. Most of the modules do not include any protection circuit to detect the open circuit faults. Despite the short circuit faults, open circuit faults are not harmful to the system. However, they degrade the quality of output waveforms and performance of the system. It is worth to mention that in some of the applications it is essential to provide a seamless operation. Seamless operation and perfect reconfiguration require earliest possible fault detection, either open circuit fault or short circuit fault. Technically speaking, internal protection circuits open the switches after detecting the short circuit fault incidence quickly. Therefore, from system-level diagnosis perspective, it is realistic to consider that the open circuit switch fault diagnosis methods are applicable for short circuit faults.

1.3 Problem Summary

The complexity of fault diagnosis methods in multiphase multilevel drives with a higher number of semiconductor switches is inevitable. As it was reviewed, applying module-level (component-based) fault diagnosis methods is engaged with the higher cost as well as the additional required space to accommodate the extra supervisory hardware. In this regard, proposing a system-based method which can characterize the failure employing system-level observations is aimed. However, the individualities of a multiphase multilevel voltage source inverters (VSI) have prompted some concerns.

Due to the defined objectives, drive systems under closed-loop controlled operation are usually concerned with controlling the voltage, current, flux, torque, or speed. The main challenge is that there is a conflict of interest between designing a model-

based fault diagnosis scheme and a closed-loop control strategy. The state of the art for control is to minimize and compensate the deviations of targeted signals from the desired behavior, and the effect of failures regard to objectives. While the aim of supervisory function which includes monitoring, automatic protection, and fault diagnosis is to detect these deviations with the purpose of producing the proper fault indicators to make the decision and take appropriate action as the reconfiguration procedure. As a result, it is essential to apply the techniques that can recognize post-fault behaviors even under controlled operation.

In current-controlled voltage source inverters, the discrimination between fault characteristics and responses which are compensated by controllers is enormous. Because, in controlled systems such as motor drive applications, output waveforms are dramatically influenced by the effect of controllers, e.g., output currents, which are supposed to carry the data reflecting the fault influence, are under control loop impact.

Additionally, real-time fault diagnosis methods based on waveform analysis in low-speed applications is more challenging, because they require a more extended period to observe significant abnormality in the output, plot the trajectories, and recognize the related fault patterns. Therefore, the window of data which carry adequate information is relatively larger. Furthermore, investigating the transient response of the system to various fault conditions offers a solution to make the decision based on the data over a shorter time interval. However, technically speaking, in multiphase drives the faulty leg should be opened before reconfiguration step. Accordingly, the informative data which results in efficient and reliable real-time fault isolation schemes are limited to the time interval

between fault occurrence and the instant of opening the leg. Noticeably, in data-driven methods, there is a tradeoff between the length of this time interval and having a seamless reconfiguration. It is desired to minimize the width of data window and maximize the extracted information from transient responses to the different failure scenarios simultaneously. On the other hand, the higher number of fault states adds extensive complexity to obtain an efficient model for classifying the fault state. In recent years, machine learning techniques have raised an intense interest in both industry and research. However, it is often difficult, expensive, and time-consuming to obtain labeled data in industrial systems for training the classifier under faulty conditions through deployed supervisory, control, and data acquisition systems. On the contrary, unlabeled data can be available easier while they contain significant information. Combining these two sets of data results in semi-supervised methods.

Motivated by those mentioned above, this paper is concerned with the intelligent and real-time system-level identification and characterizations of open circuit faults for multi-level multi-phase inverters under speed and current closed loop control operation.

1.4 Dissertation Organization

The primary concentration of this dissertation is a two-fold. First, to develop a real-time fault detection scheme in multiphase multilevel inverters. Second, identifying and isolating the failed components. In this regard, in chapter 2 the principles of the multiphase machine models and operation of multilevel neutral point clamped inverters are reviewed.

In chapter 3, general fault diagnosis framework is discussed. Further, considering the physical characteristics of system and measurement availabilities, an efficient monitoring structure and model-based fault detection scheme are proposed, which are combining system-level signals (stator currents) with component-level (pulse width modulation switching signals) in a nonlinear model identification procedure. Two nonlinear adaptive model identifiers, self-recurrent wavelet neural network, and self-recurrent wavelet neuro-fuzzy network, are reviewed and used to predict the stator current based on pulse width modulation switching signals. Taking pulse width modulation switching signals into account result in enhancement of the identifier's performance. Efficient model identifier predicts the stator currents and generates the robust residuals, which are insensitive to noise but sensitive to failure.

In chapter 4, the proposed fault detection scheme is applied on a multiphase conventional two-level inverter, where a real-time fault indicator is introduced to isolate the failed switch. Moreover, to evaluate the performance of real-time indicator, the results are compared with the results of two supervised classifiers, support vector machines and linear discriminant analysis classifiers. The evaluations illustrate that the proposed real-time indicator is faster and more reliable. In chapter 5, the semi-supervised learning concept for fault isolation in multiphase multi-level inverters is introduced and the results are presented. Another fault isolation structure for the multiphase multi-level inverters, based on analytical data mining methods is discussed in chapter 6. Finally, in chapter 7, the conclusion and suggestion to extend this study are provided.

2. REVIEW OF MULTIPHASE MACHINES AND MULTILEVEL INVERTERS

2.1 Multiphase Machines

The most remarkable benefit of the multiphase machines is the redundancy capability by the presence of the additional phases. The n-phase machine can continue to operate with (n-3) healthy phases, which offers higher reliability in post-fault operation. Therefore, a multiphase machine is attractive alternative in the sensitive applications where the seamless and uninterrupted operation is required. However, earliest possible detection and an efficient fault tolerant control strategy to operate after detecting the fault is vital.

Different strategies for normal operating and post-fault operating of the multiphase machines are available. It has been shown that hysteresis controllers as one of the earliest studies on multiphase machine control schemes, closed loop voltage by frequency (V/f) control, constant and variable switching frequency direct torque control (DTC), and vector control (field oriented control) schemes can be applied to multiphase induction machines [28-30]. Modeling and control of multiphase machines are usually performed in the synchronous frame using the synchronous frame transformation, to convert the phase variables to d-q-x-y-0, same as conventional three-phase machines. The torque and flux control in the balanced machine is taken to place with independent of the d-q current components, coupled with the rotor. The x-y current components create current harmonics and have effect on loss, but they have no influence on the torque performance. In the

machines that are not entirely balanced, supplementary controllers for x-y current leads to the elimination of the existing x-y current components.

The basic equations of a multiphase induction machine under balanced operating conditions are given in [31-33]. Assuming that the machine windings are distributed sinusoidal and linear, the stator and rotor voltage equations of the multiphase induction machine can be written as follows:

$$\begin{aligned}[V_s] &= [R_s][I_s] + \frac{d\lambda_s}{dt} \\ [V_r] &= [R_r][I_r] + \frac{d\lambda_r}{dt}\end{aligned} \tag{1}$$

and the stator and rotor air gap flux linkage equations are given by:

$$\begin{aligned}\lambda_s &= [L_{ss}][I_s] + [L_{sr}][I_r] \\ \lambda_r &= [L_{rr}][I_r] + [L_{rs}][I_s]\end{aligned} \tag{2}$$

The current and voltage vectors are:

$$\begin{aligned}[I_s] &= [i_{s1} \quad i_{s2} \quad \ldots \quad i_{sn}]^T \\ [V_s] &= [v_{s1} \quad v_{s2} \quad \ldots \quad v_{sn}]^T \\ [I_r] &= [i_{r1} \quad i_{r2} \quad \ldots \quad i_{rn}]^T \\ [V_r] &= [0 \quad 0 \quad \ldots \quad 0]^T\end{aligned} \tag{3}$$

In these equations, $[R_s]$, $[R_r]$, the stator and rotor resistance matrices, are $n \times n$ matrix given by $R_s = r_s \times I$ and $R_r = r_r \times I$ where I is the identity matrix. The $[L_{ss}]$, $[L_{rr}]$ are the stator and rotor self-inductance matrices, respectively. Also, $[L_{sr}]$ and $[L_{rs}]$ are the mutual inductance matrices. This model lies in the coupled nonlinear expressions of the

inductance matrices. To obtain a decoupled model, two transformation matrices, as *Tc* and *Tz*, are required to split the original model into two decoupled subspaces [34-35].

The electromechanical energy conversion subspace (*dq*-subspace) and the subspace which gives the losses (zero-subspace).

For n-phase induction machines, two orthogonal vectors might be defined as:

$$\begin{aligned}\left[d_0\right]&=\left[\cos(\varphi_0+\varphi_1)\quad \cos(\varphi_0+\varphi_2)\quad \cos(\varphi_0+\varphi_3)\quad ...\quad \cos(\varphi_0+\varphi_n)\right]^T\\ \left[q_0\right]&=\left[\sin(\varphi_0+\varphi_1)\quad \sin(\varphi_0+\varphi_2)\quad \sin(\varphi_0+\varphi_3)\quad ...\quad \sin(\varphi_0+\varphi_n)\right]^T\end{aligned} \tag{4}$$

d_0 and q_0, which refer to the *n*-phase normal operation. In symmetric *n*-phase winding, φ_i for i=1,2,…,n are the stator current phase angles. In five-phase motor, there are defined as $\varphi_1 = 0$, $\varphi_2=\frac{4\pi}{5}$, $\varphi_3=\frac{6\pi}{n}$, $\varphi_4=\frac{8\pi}{n}$, $\varphi_5=\frac{8\pi}{n}$. Moreover, as an example for the asymmetric multiphase machine, in six phase motor stator current phase angles are defined as $\varphi_1 = 0$, $\varphi_2=\frac{\pi}{6}$, $\varphi_3=\frac{6\pi}{n}$, $\varphi_4=\frac{8\pi}{n}$, $\varphi_5=\frac{8\pi}{n}$.

φ_0 is fixed to obtain two orthogonal vectors [*d*] and [*q*]. These two vectors are orthogonal if $d\times q^T=0$. Therefore:

$$\begin{aligned}&\cos(\varphi_0+\varphi_1)\sin(\varphi_0+\varphi_1)+\cos(\varphi_0+\varphi_2)\sin(\varphi_0+\varphi_2)+\cos(\varphi_0+\varphi_3)\sin(\varphi_0+\varphi_3)\\ &+...+\quad \cos(\varphi_0+\varphi_n)\sin(\varphi_0+\varphi_n)=0\end{aligned} \tag{5}$$

In normal operation, φ_0 is zero. If fault happens at the phase *x* and *j* denotes the active phases, then the term corresponding to *x* will be removed from equation above. It will be re-written as:

$$\begin{aligned}&\cos(\varphi_0+\varphi_1)\sin(\varphi_0+\varphi_1)+\cos(\varphi_0+\varphi_2)\sin(\varphi_0+\varphi_2)+\cos(\varphi_0+\varphi_3)\sin(\varphi_0+\varphi_3)\\ &+...+\quad \cos(\varphi_0+\varphi_n)\sin(\varphi_0+\varphi_n)=0\end{aligned} \tag{6}$$

If j denotes to the set of active phases, under failure operation at phase *x*, fulfilling the equation above leads to the φ_0 calculation as:

$$\varphi_0 = \frac{-1}{2}\tan^{-1}\left(\frac{\sum_j \sin(2\varphi_j)}{\sum_j \cos(2\varphi_j)}\right) \tag{7}$$

The transformation matrix T is the as follows:

$$T = \begin{bmatrix} [T_c] \\ [T_z] \end{bmatrix} \tag{8}$$

where

$$T_c = \begin{bmatrix} [d]/\|[d]\| \\ [q]/\|q\| \end{bmatrix} \quad , \quad T_z = null\left([T_c]\right) \tag{9}$$

$[d]$ and $[q]$ are obtained in the same manner for active phases. By applying the transformation *T* to the machine model equations results in:

$$\begin{bmatrix} v_{sd} \\ v_{sq} \end{bmatrix} = \begin{bmatrix} R_s + L_{sd}\,p & 0 \\ 0 & R_s + L_{sq}\,p \end{bmatrix}\begin{bmatrix} i_{sd} \\ i_{sq} \end{bmatrix} + \begin{bmatrix} M_d\,p & 0 \\ 0 & M_q\,p \end{bmatrix}\begin{bmatrix} i_{rd} \\ i_{rq} \end{bmatrix}$$
$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} M_d\,p & 0 \\ 0 & M_q\,p \end{bmatrix}\begin{bmatrix} i_{sd} \\ i_{sq} \end{bmatrix} + \begin{bmatrix} R_r + L_r\,p & 0 \\ 0 & R_r + L_r\,p \end{bmatrix}\begin{bmatrix} i_{rd} \\ i_{rq} \end{bmatrix} \tag{10}$$

where

$$\begin{aligned} L_{sd} &= L_{ss} + \|d\|^2 L_{ms} \quad & M_d &= \|d_0\|.\|d\| L_{ms} \\ L_{qs} &= L_{ss} + \|q\|^2 L_{ms} \quad & M_q &= \|q_0\|.\|q\| L_{ms} \\ L_r &= L_{rr} + \|d_0\|^2 L_{ms} = L_{rr} + \|q_0\|^2 L_{ms} \end{aligned} \tag{11}$$

In normal operation, $[d] = [d_0]$ and $[q] = [q_0]$.

The stator voltage equations in zero subspace, which are related to losses, are:

$$
\begin{aligned}
v_{sz1} &= R_s i_{sz1} + L_{ss} \frac{di_{sz1}}{dt} \\
v_{sz2} &= R_s i_{sz2} + L_{ss} \frac{di_{sz2}}{dt} \\
v_{sz3} &= R_s i_{sz3} + L_{ss} \frac{di_{sz3}}{dt}
\end{aligned} \tag{12}
$$

Eventually, the expression for the torque of a multiphase motor in the stationary *d-q* reference frame is given by:

$$
T_m = PL_r(i_{dr} i_{qs} - i_{qr} i_{ds}) \tag{13}
$$

2.2 Multilevel Inverters

The conventional two-level inverter can create only two different output voltage levels for the load: $V_{dc}/2$ or $-V_{dc}/2$.To produce an AC output voltage these two voltages are usually switched with PWM. It creates harmonic distortions in the output voltage, EMI, and high *dv/dt*. On the other hand, the diode-clamped multilevel inverter employs clamping diodes and cascaded dc capacitors to build up ac voltage waveforms with multiple levels. The inverter can be configured as different-level topologies, but only the three-level inverter, often known as neutral-point clamped (NPC) inverter, has applied in wide range of applications in high-power medium-voltage drives [36-39]. Compared to the two-level inverter, the main feature of the NPC inverter is reducing *dv/dt* and improving the THD in the ac output voltages. Additionally, NPC can reach higher voltage without switching devices in series. In this study, the operation of the NPC inverter under the various open-circuit fault scenario will be discussed. The configuration of a five-phase three-level NPC connected to a five-phase machine is indicated in Figure 1. Prior to

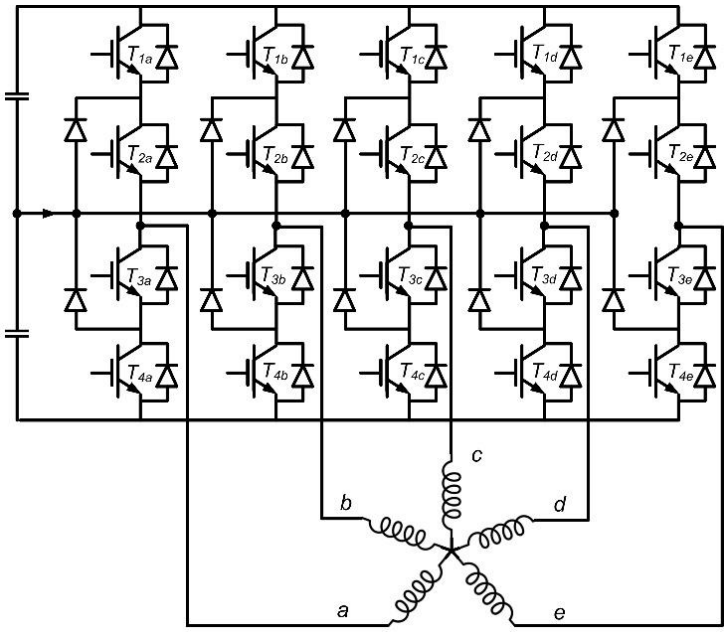

Figure 1 Five-phase three-level NPC inverter connected to a five-phase machine.

analysis the fault scenarios, the switching operation states for each leg of the NPC inverter in general modulation are shown in Table 1 where state [P] denotes the upper two switches of the inverter leg are on. Therefore, the terminal voltage is $+V_{dc}$. State [N] represents the state with two lower switches on, when the terminal voltage is $-V_{dc}$. Further, state [O] signifies two inner switches on, that results in clamping the terminal voltage to zero.

Table 1 Definition of switching states.

Switching State	Device switching status (per leg)				Inverter terminal voltage v_{az}
	T_1	T_2	T_3	T_4	
[P]	on	on	Off	off	$+V_{dc}$
[O]	off	on	On	off	0
[N]	off	off	On	on	$-V_{dc}$

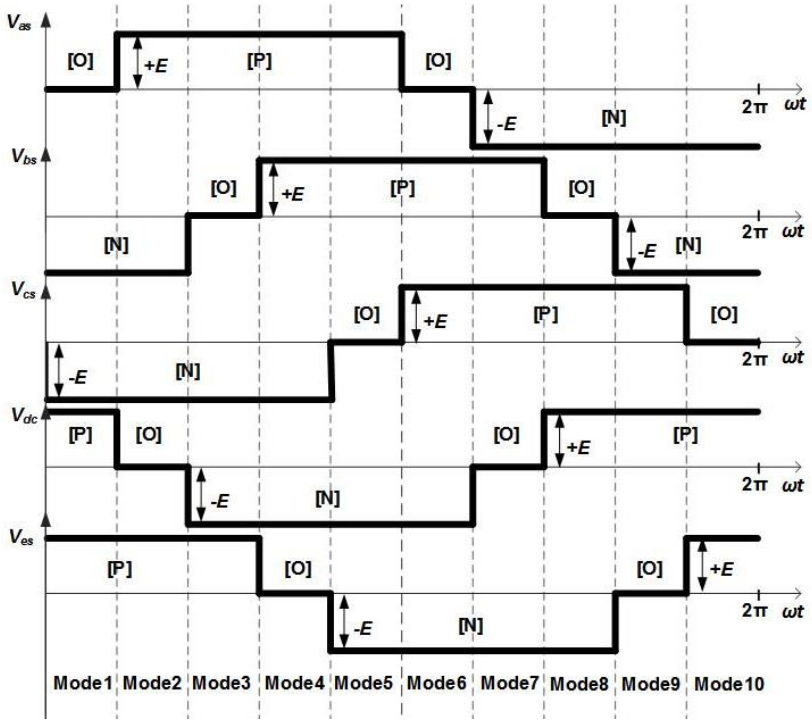

Figure 2 Switching state arrangement and inverter terminal waveforms for five-phase NPC inverter.

There are two possible cases of conduction in state [O] whereas positive load current yields in turning on the D_{z1} and inner switch of the upper part. Also, negative load current forces D_{z2} to turn on while the inner switch of lower part conducts [33]. A simple example of switching state arrangement and five inverter terminal waveforms is shown in Figure 2. To investigate the operation of the inverter under normal and fault condition. As the figure indicates, this example includes ten different modes. In each mode, two phases are in [P] switching state, two phases are in [N] switching state, and one phase is in [O] switching state. Also, there are two commutation cases in each switching state: commutation with positive phase current or negative phase current.

Due to symmetrical property, all possible fault cases in one mode have been investigated, and the result is generalized. In the mode 1, nine switches and one clamping diode are on in normal operation. Lower pair of switches in phase *b* (T_{3b}, T_{4b}) in [N] switching state, lower pair of switches in phase *c* (T_{3c}, T_{4c}) in [N] switching state, upper pair of switches in phase *d* (T_{1d}, T_{2d}) in [P] switching state, upper pair of switches in phase *e* (T_{1e}, T_{2e}) in [P] switching state are active. Also, depending on the current of phase a, either T_{2a} and Dz_1 are commutating or T_{3a} and Dz_2. If the current of phase *a* is positive T_{2a} and D_{z1} conduct, but if the current is negative T_{3a} and D_{21} conduct. Therefore, there are ten open switch failure cases in each mode, which are categorized as follows:

i. *Open-switch fault in the top switches of upper pairs:* This situation might happen in the phase with [P] switching state: In the mode 1, T_{1d} and T_{1e} have similar commutation. If T_{1d} or T_{1e} fail, the upper clamping diode of the corresponding phase start to conduct, which lead to zero voltage in the phase terminal. Due to the inductive load, the load current is constant during the transient failure operation. Figure 3 indicates the commutation under the condition of open-switch failure in T_{1d}.

ii. *Open-switch fault in the inner switches of upper pairs:* This case might happen in either the phase with [P] switching state or the phase with [O] switching state when current is positive. In [P] switching state, if one of the switches T_{2d} or T_{2e} fails, there will be no current path in the upper half leg of d. Therefore, parallel diodes in the lower pair (D_{T3d}, D_{T4d}) are forced to turn on are forced to turn on Figure 4. Likewise, parallel diodes (D_{T3e}, D_{T4e}) are one after open switch failure in T_{2e}.

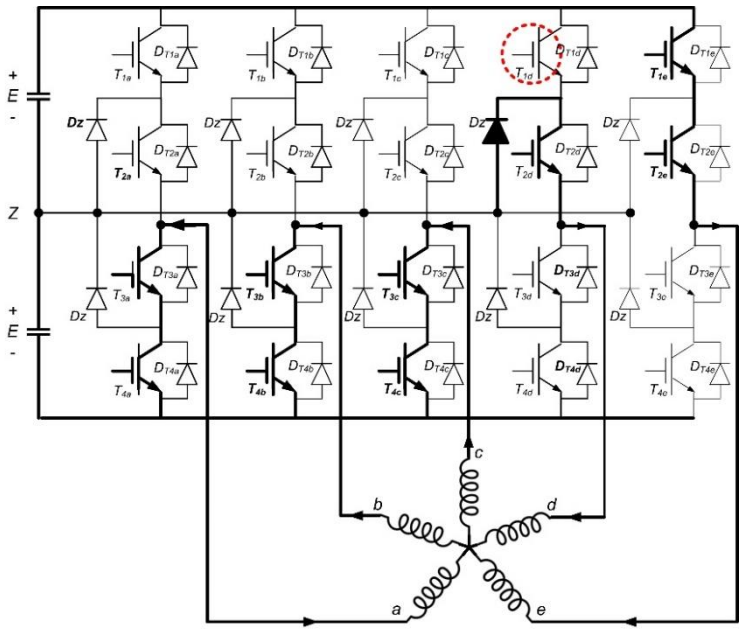

Figure 3 Commutation under T_{1d} open-switch failure in [P] switching state.

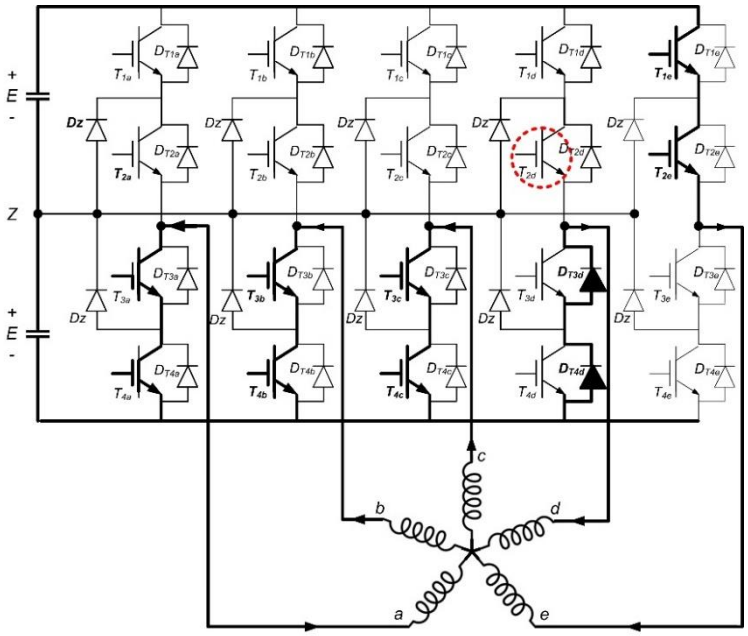

Figure 4 Commutation under T_{2d} open-switch failure in [P] switching state.

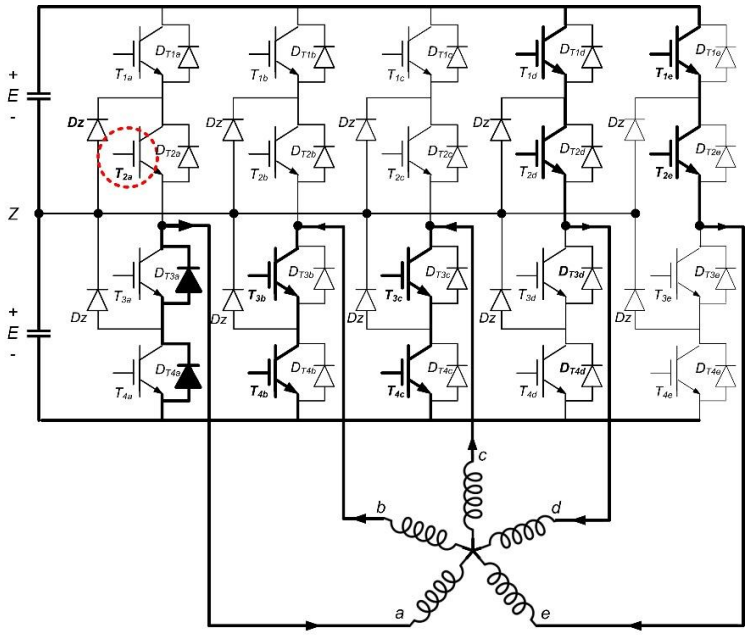

Figure 5 Commutation under T_{2a} open-switch failure in [O] switching state.

Similarly, if the failure occurs in phase a, the leg with [O] switching state when the load current is positive, parallel diodes in the lower pair (D_{T3a}, D_{T4a}) are forced to commutate, shown in Figure 5.

iii. *Open-switch fault in the inner switches of lower pairs:* This case might happen in either the phase with [N] switching state or the phase with [O] switching state when current is negative. If one of the inner switches of lower pairs in the legs with [N] switching state fails, phase *b* or *c* in mode 1, it is not possible to commutate lower pair switches. Therefore, parallel diodes in upper pair of the failed leg (D_{T1b}, D_{T2b}) are forced to turn on (Figure 6). Likewise, parallel diodes (D_{T1c}, D_{T2c}) are one after open switch

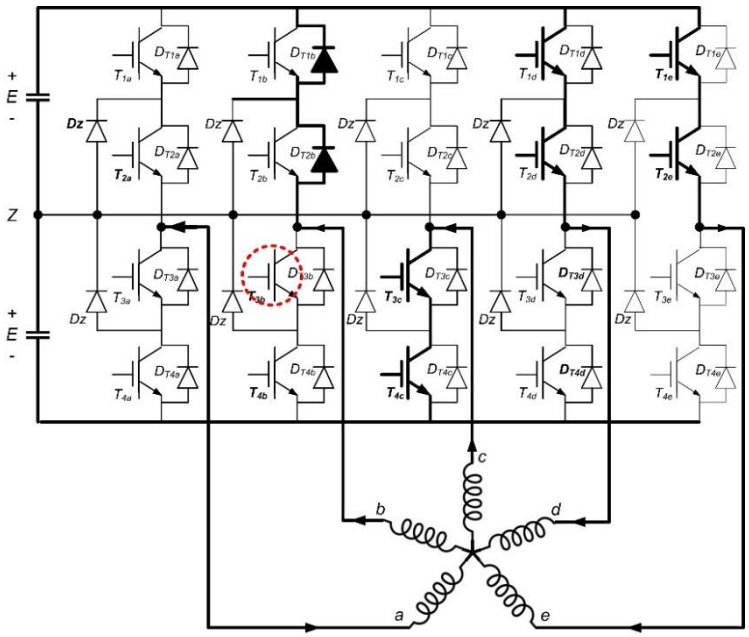

Figure 6 Commutation under T_{3b} open-switch failure in [N] switching state.

failure in T_{3c}. In the same way, if the failure occurs in phase a, the leg with [O] switching state when the load current is negative, parallel diodes in the upper pair (D_{T1a}, D_{T2a}) are forced to commutate, shown in Figure 7.

iv. *Open-switch fault in the bottom switches of lower pairs:* This case might be happened in the phase with [N] switching state: In the mode 1, If T_{4b} or T_{4c} fail, the lower clamping diode of the corresponding phase have to turn on, which leads to zero voltage in the phase terminal. Figure 8 indicates the commutation under the condition of open-switch failure in T_{4b}. The failure T_{4c} is the same.

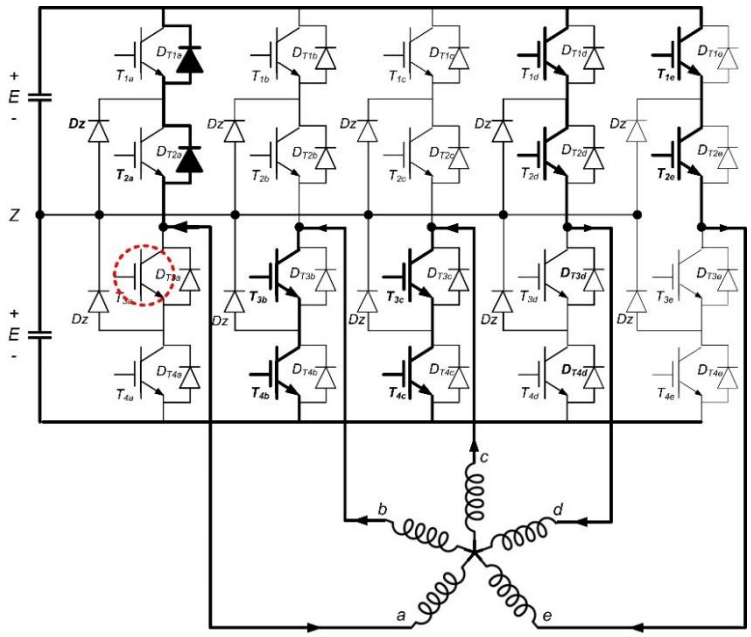

Figure 7 Commutation under T_{3a} open-switch failure in [O] switching state.

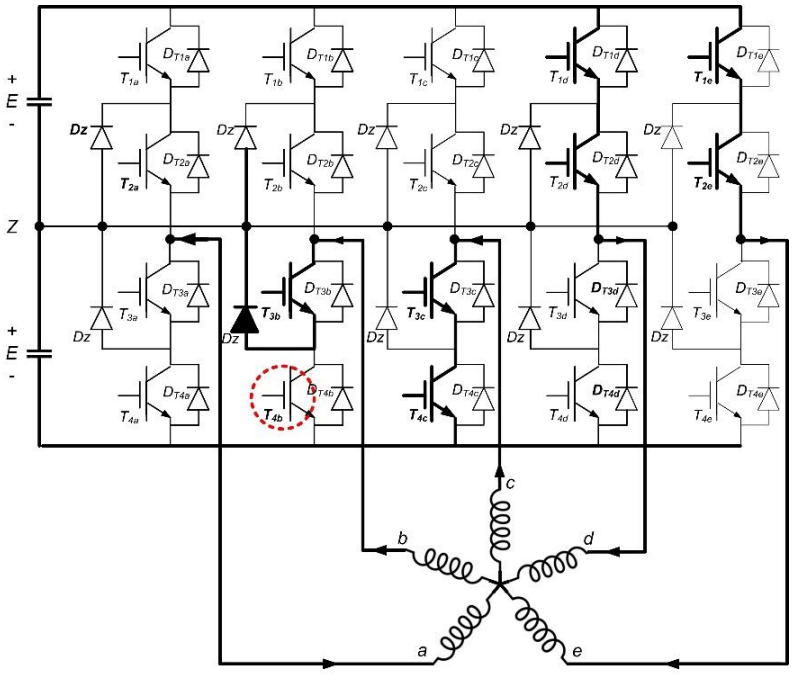

Figure 8 Commutation under T_{4b} open-switch failure in [N] switching state.

It worth to note that in practice, inverters might come with a Complex Programmable Logic Device (CPLD) that checks for the correct switching pattern, correct switch-on, and switch-off sequences, and the correct electrical limits[40]. In fact, CPLDs supervise the on and off sequences of the switches, dead times, and PWM gate signals sequences to make sure that inner switches are switched on before outer switches, and the outer switched are off before inner switches. CPLDs check generated PWM signals by control strategy. There are three switching areas: allowed, potentially destructive, and destructive switching areas which are presented in Table 2 from normal or open switch failure operation perspective. If generated PWM be in the allowed area, it will be delivered to switches directly. If the pattern is one of the potentially destructive or destructive cases, the CPLD will shut down the phase leg until receiving the allowed pattern. Then, it will resume the operation. As it was discussed on the difference between open circuited switches fault and short-circuited switches fault, the short circuit is detected by internal protection systems by detecting the maximum allowed values of current, temperature, or voltage, and thus CPLD shuts down the phase leg. But in open circuit failures, these values remain in allowed or potentially destructive areas and do not lead to shut down an extra fault diagnosis scheme is required to detect and isolate the failure.

Irrespective of controlling employed strategy, the switching patterns might be generated using different modulation techniques, e.g., conventional sinusoidal PWM (SPWM) or the space vector PWM (SVPWM). Similar to the three-phase machine, applying SPWM or SVPWM techniques need different levels of DC bus utilization [41]. In this research study, the sinusoidal pulse width modulation method is employed. The

SPWM uses a triangular carrier wave modulated by a sinusoidal waveform, where the intersection points determine the switching times. In current controlled inverters, the sinusoidal waveforms come from the corresponding control strategy. Also, the indirect field oriented control strategy with PI current controllers is implemented.

Table 2 Allowed, potentially destructive, and destructive switching patterns for 3L-NPC.

T_1	T_2	T_3	T_4	Operation	Area
off	off	off	off	Open switch failure in T_2 or T_3 [O] switching state	Allowed
off	on	off	off	Open switch failure in T_1 [P] switching state	
off	off	on	off	Open switch failure in T_4 [N] switching state	
on	on	off	off	Normal [P] switching state	
off	off	on	on	Normal [N] switching state	
on	off	off	off	Open switch failure in T_2 [P] switching state	Potentially destructive
off	off	off	on	Open switch failure in T_3 [N] switching state	
on	off	off	on	Open switch failure in T_2 and short circuit failure in T_4	
on	off	off	on	Open switch failure in T_3 and short circuit failure in T_1	
on	off	on	off	Open switch failure in T_2 and short circuit failure in T_3	
off	on	off	on	Open switch failure in T_3 and short circuit failure in T_2	
on	on	on	off	Short circuit failure in T_3	Destructive (CPLD shuts down the phase)
on	on	off	on	Short circuit failure in T_4	
on	off	on	on	Short circuit failure in T_1	
off	on	on	on	Short circuit failure in T_2	
on	on	on	on	Short circuit failure in (T_1, T_2) or (T_3, T_4)	

3. GENERAL FAULT DIAGNOSIS FRAMEWORK

Per International Federation of Automatic Control (IFAC) safe process technical committee, a fault is defined as "an unpermitted deviation of at least one characteristic property or parameter of the system from the acceptable/usual/standard condition"[42-43]. Undesired or not permitted states of a dynamic process, which may take place in dynamic plants, sensors, actuators or the switching logic components, are identified by a supervisory function within an automatic control system. It is followed by a taking a proper action for reconfiguration to maintain the optimal seamless operation [44]. Supervisory functions include the following tasks:

(1) The monitoring system that checks the measurable variables concerning uncertainties and tolerances to create a real-time residual (fault signature).

(2) Automatic protection system that evaluates the residuals to initiates an appropriate counteraction under the unpermitted state of dynamic plant (Fault Detection).

(3) Supervision and fault diagnosis system that extracts the fault features from the residuals and identifies the fault characteristics (Fault Isolation).

(4) Supervision and fault diagnosis system that makes decision for counteractions (Fault Reconfiguration).

Generally speaking, fault detection and isolation schemes are inspired by redundancy theory and implemented by hardware or analytical methods. Hardware redundancy principally duplicates a signal using extra sensors or signal processing

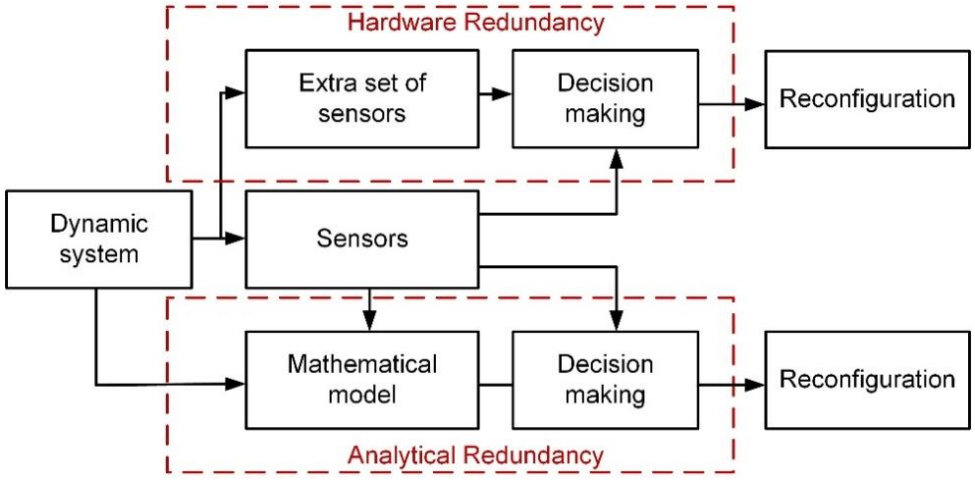

Figure 9 Concept of redundancy.

methods, while analytical redundancy utilizes a mathematical (quantitative or qualitative) model of the system. Quantitative and qualitative methods employ explicit mathematical models and artificial intelligence modeling techniques, respectively.

Figure 9 illustrates the concept of redundancy in dynamic plants.

Although capturing the behavior of dynamic systems through a mathematical model is one of the most critical methods in science and engineering, most of the given industrial systems in control, signal processing, and information science are complex and nonlinear. Generally speaking, nonlinear complexities of inverters, including switching characteristic and dead bands, pose obstacles for some aspects of modeling. Thus, the nonlinear models are applied to deal with nonlinearities. Nonlinear models are obtained from theoretical modeling by a priori knowledge of the nature and the intrinsic of the systems. A practical alternative to theoretical modeling is system identification (empirical modeling).

It is preferred to model of the system online and adaptively to deal with the dynamic of the system. There is a trade-off between the accuracy of the model and the simplicity. In fact, the model is an approximation of the real phenomena; therefore it is crucial to fulfilling the Occam's razor hypothesis. Occam's razor hypothesis states that plurality should not be posited without necessity [45]. In other words, the more straightforward solution is more reasonable. This concept is known as the parsimonious principle.

Fault residual generation methods are categorized as observer-based methods (full-state or unknown input), parity relations approach, optimization-based approach, Kalman filter-based approach, stochastic approach, discrete event systems/hybrid systems approaches, system identification approaches(linear and nonlinear), and AI techniques [46-47].

The first task of constructing a fault detection procedure might be framed as a residual generation. Residual, which is the difference between the measured output $y(t)$and the estimated $\hat{y}(t)$, should meet the invariant relation and fault detectability conditions. Invariance relation requires that the mean of residual in normal operation be zero, while fault detectability condition necessitates that when the fault occurs the residual should deviate from zero [48].

The second step of this procedure is to evaluate the generated residuals. Residual evaluation is a decision-making process that mostly is expressed as a threshold logic of a decision function. Defining a robust residual is the only way to avoid the false alarm with an adequate and acceptable sensitivity to noise and uncertainties. Robust residual

evaluation can be accomplished in many ways, for example by statistical data processing, data reconciliation, correlation, pattern recognition, fuzzy logic or adaptive thresholds. In recent years, there has also been a clear trend towards an increased involvement of knowledge-based and artificial intelligence methods, including qualitative modeling for residual generation and residual evaluation [49]. Efficient

3.1 Efficient Monitoring Systems Based on Measurable Signals

From power electronic devices perspective, module-level decision making in component-based techniques is reputed to be fast and reliable, but they require developing additional supplementary hardware monitoring equipment. On the other hand, module-level decision making procedure in existing system-based techniques is not satisfactory for real-time systems or closed loop controlled systems such as motor drive applications, because output waveforms are influenced by the effect of controllers.

In normal operation, actual (measured) current value is relatively equal to the predicted (approximated) value. However, under fault operation, there might be a significant difference between these two values, which is employed to define a fault indicator (residual). The procedure of generating a robust residual is followed by the residual evaluation stage, which is a logical process to make a decision and transform quantitative knowledge into qualitative, e.g., yes/no statements. The residual evaluation also can be seen as a classification problem.

Considering all the discussed challenges, this study proposes a novel hybrid fault diagnosis framework, which borrows the concept of module-level (component-based) methods to employ the statuses of the component in a system level (system-based) framework. A nonlinear identification model method is applied to develop a model-based fault diagnosis methods. Nonlinear model identification can cover all the nonlinearities and uncertainties of switching devices and handle the noise and disturbances.

Moreover, the measurable signals which are utilized to identify the model are a combination of module-level type and system-level variables. Technically speaking, the pulse width modulated signals, which are generated by the control strategy, indicate valuable information about switch statuses in both normal operations and under fault operations. The PWM signals are inherently component-level signals but it is not required to install additional hardware to measure them. Engaging the output currents as system-level signals and the PWM signals as indirect component-level signals make it possible to construct a hybrid framework for fault diagnosis purposes, which has the advantages of accuracy, a quick reflection of failure on the switching states, and providing precious information about controllers performance without positioning additional monitoring hardware. Employing PWM signals results in a significant enhancement in the performance of fault diagnosis scheme, which will be discussed in this section in details. The general configuration of the system including motor, inverter (conventional or multilevel), and control system (filed oriented control strategy) is shown in Figure 10. As it is indicated, the measurable signals, stator current signals and PWM signals generated by the control system, are utilized in the hybrid fault diagnosis scheme.

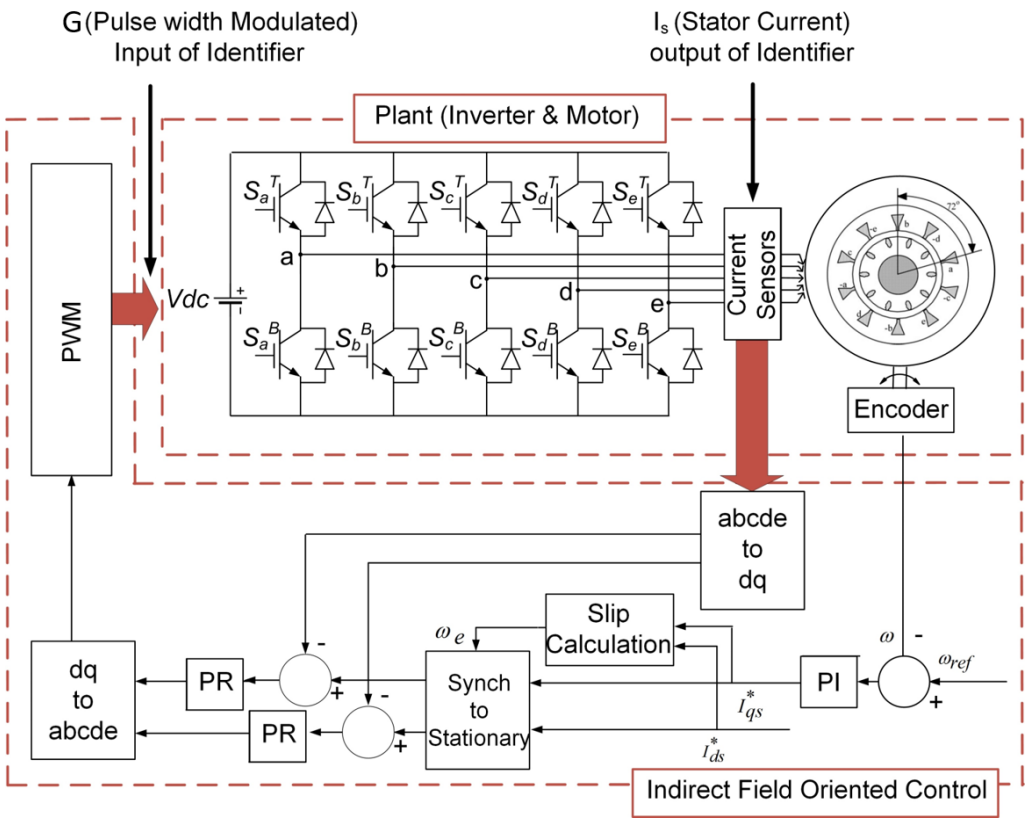

Figure 10 Block diagram of an indirect FOC for five-phase IM driven with a five-phase two-level inverter.

In this study, an analytical redundancy method is preferred. Motor and inverter and control system are considered as an integrated dynamic plant. The monitoring system measures stator currents simultaneously, and it is assumed that any abnormality in switches is reflected in the motor currents responses. To perform automatic protection function, inspiring the model-based fault diagnosis strategies, a model is estimated under normal operating conditions either offline or online. Subsequently, normal operation values of the currents are being approximated adaptively. Then, any significant deviation of measured signals from the predicted value could be interpreted as a fault symptom.

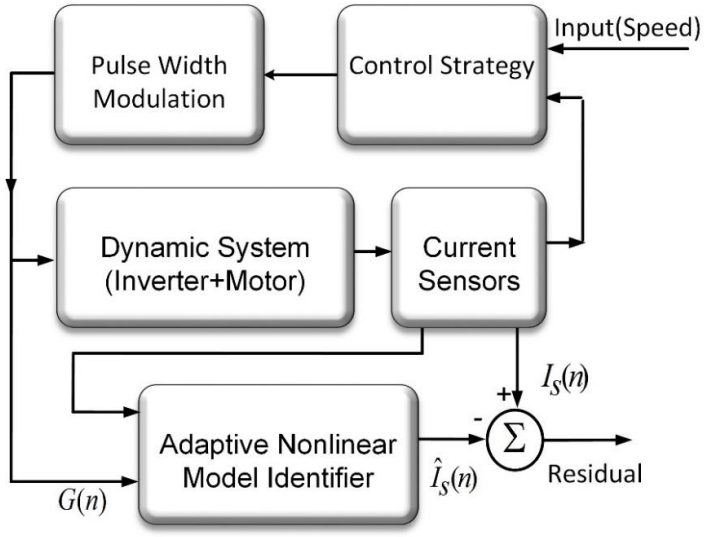

Figure 11 The structure of the model-based nonlinear identifier.

Nonlinear complexities of inverters, including switching characteristic and dead bands, pose obstacles for some aspects of modeling. Therefore, the nonlinear model is used to deal with large nonlinearity. Frequently, nonlinear models are obtained from theoretical modeling by a priori knowledge of the nature and the intrinsic of the systems. A practical alternative to theoretical modeling is system identification or empirical modeling. To deal with the dynamic of the system, it is preferred to model the system online and adaptively.

To deal with all mentioned challenges, a hybrid method is presented which takes advantage of both component-based and system-based concepts. General diagram of process model based fault diagnosis for motor drive application under the current controlled condition is shown in Figure 11.

An adaptive nonlinear identifier estimates the output waveforms (stator currents) and compares the estimated and measured waveforms to create the fault indicators. Pulse

width modulated (PWM) signals, are calculated by the control strategy and are defined as inputs of identifier to increase the speed of response, accuracy, and efficiency of the nonlinear identifier. In this manner, switching pattern has been brought to attention and provides a more accurate image of the system. PWMs are inherently component-level signals, but applying them does not impose additional sensor. Therefore, the stator currents (system-level signals) are approximated based on the component-level signal in a hybrid framework. Accordingly, the fault diagnosis performance is enhanced without additional sensor or hardware installation.

3.2 Residual Generation

From the promptness of fault identification perspective, it is desirable to detect the fault as soon as possible. Also, the employed technology should be able to meet the identifiability condition. The fault diagnosis schemes must be capable of identifying the fault characteristics which are compensated by controllers and also discriminating between the effect of component failure on the output and the impact of controllers. When no fault is present, the residual will be nonzero due to residual disturbances, i.e. measurement disturbances, process disturbances and model uncertainties. Therefore, the residual evaluation must be robust against these disturbances to avoid false alarms.

In practice, there is usually such a significant number of unknown inputs that, in the face of the limitations of available measurement information, a complete decoupling from all unknown inputs is hardly achievable even if the model uncertainties are structured. The situation becomes even worse if the uncertainties are unstructured, in

which case a perfect decoupling in the residual generation stage is impossible. Hence the residuals or any decision functions built from them always deviate from zero even if no fault is present.

In this section, a model-based fault detection approach has been developed to generate appropriate residual. Two nonlinear identifiers, self-recurrent wavelet neural network and self-recurrent wavelet neuro-fuzzy network, are reviewed. The data-driven nonlinear system identifier estimates the model of the entire system, including converter and motor, which predicts the phase currents based on pulse width modulated gate signals. In fact, a nonlinear dynamic for stator current described by the following nonlinear autoregressive moving average (NARMA) model is defined as:

$$\hat{I}_s(t) = h(I_s(t-1),...,I_s(t-n_{dI}),S(t),...,S(t-n_{ds})) + \xi(t) \tag{14}$$

where stator current, $I_s(t)$, and gate signals, $S(t)$, are the output and input of nonlinear identified system, respectively. The function $h(.)$ denoted a smooth nonlinear function and $\xi(t)$ is Gaussian white noise. n_{dI} and n_{ds} are the dynamic orders of the output and input.

3.3 Structure of Nonlinear Identifier for Residual Generation

In the n-phase machine, there are n generated fault residuals, which are defined as differences between measured and estimated stator currents in each phase. These n residuals, calculated in parallel, are as follows:

$$r_x(t) = I_{xs}(t) - \hat{I}_{xs}(t); x \in \{a,b,c,d,e,...\} \tag{15}$$

where $r_x(t)$ is a nonlinear function of the previous state of the observed stator current, $I_s(t-1)$, and recent calculated PWM switching signals, $S(t)$, as:

$$\hat{I}_{xs}(t) = h(I_s(t-1), S(t)); x \in \{a, b, c, d, e, ...\} \quad (16)$$

where $I_s(t)$ is stator current ($n \times 1$) as:

$$\left[I_s(t)\right]_{n\times 1} = [I_{as}(t) \quad I_{bs}(t) \quad I_{cs}(t) \quad I_{ds}(t) \quad I_{es}(t) \quad ...]^T \quad (17)$$

$S(t)$ is the augmented vector of the entire set of generated PWM switching signals. In a conventional two-level inverter, $S_x^T(t)$ and $S_x^B(t)$ denote the top switch and bottom switch of the phase (leg) *x*, respectively. Therefore, $S(t)$ is constructed as:

$$\left[S(t)\right]_{n\times 2} = \left[S_a^T \quad S_b^T \quad S_c^T \quad S_d^T \quad S_e^T \quad ... \quad S_a^B \quad S_b^B \quad S_c^B \quad S_d^B \quad S_e^B \quad ...\right]^T \quad (18)$$

In three-level NPC inverter, $S(t)$ is formed as:

$$\left[S(t)\right]_{n\times 2} = \left[S_a^1 \quad S_b^1 \quad S_c^1 \quad S_d^1 \quad S_e^1 \quad ... \quad S_a^2 \quad S_b^2 \quad S_c^2 \quad S_d^2 \quad S_e^2 \quad ...\right]^T \quad (19)$$

where $S_x^1(t)$ and $S_x^2(t)$ denote the outer top switch and inner top switch of the phase (leg) x, respectively. In three-level NPC inverter, complementary switching mechanism allows to re-write $S(t)$ as:

$$\left[S(t)\right]_{n\times 2} = \left[S_a^3 \quad S_b^3 \quad S_c^3 \quad S_d^3 \quad S_e^4 \quad ... \quad S_a^4 \quad S_b^4 \quad S_c^4 \quad S_d^4 \quad S_e^4 \quad ...\right]^T \quad (20)$$

As a result, the residuals are:

$$r_x(t) = h(I_s(t-1), S(t)) - I_s(t); x \in \{a, b, c, d, e, ...\} \quad (21)$$

where $h(.)$ is a smooth nonlinear function. To identify the nonlinear function with more accuracy, the input of the nonlinear identifier,$x(t)$, is considered with a dynamic as:

$$x(t) = \left[S(t) \quad ... \quad S(n-n_{dS}) \quad I_S(n-1) \quad ... \quad I_S(n-n_{dI})\right] \quad (22)$$

The n_{dS} denotes the dynamic of the PWM switching signal observations, and n_{dI} represents the dynamic of the stator current observations. In this manner, nonlinear system identifier estimates the next state of stator currents precisely.

In this research study, the proposed hybrid fault detection structure employs self-recurrent wavelet neural networks (SRWNN) and self-recurrent wavelet neuro-fuzzy networks (SRWNFN). In the following sections, wavelet network concept, the structure of SRWNN, and SRWNFN will be reviewed in detail.

3.3.1 Self-Recurrent Wavelet Neural Networks

The wavelet theory, as a powerful concept to present the transient behavior of the signals, has been widely studied recently and has found many applications in various areas, such as numerical analysis, engineering, and signal processing. Compared to other basic functions such as Fourier analysis, time-frequency representation of the signal is the most substantial property of wavelet concept. Wavelet theory is extensively employed to localize the non-stationary behavior of the signals both in time and frequency domains [50-56]. Applying wavelet on many functions and operators to transform them into the wavelet domain makes them sparse. This sparseness is an important feature that results in different applications is signal processing and data analysis areas. Wavelet frames are constructed by the operations of translation and dilation of a single fixed function called mother wavelet. Wavelet transformation can be accomplished in three different ways known as the wavelet series, continuous, and discrete wavelet transformation. The wavelet series maps a function of the continuous variable into a sequence of coefficients. In

continuous wavelet transformation, a wavelet $\phi_j(x)$ is derived from its mother wavelet $\phi(z)$ by the relation.

$$\phi_j(x) = \phi(\frac{x - m_j}{d_j}) = \phi(z_j) \tag{23}$$

where m and d are real numbers in R and R^+, respectively. Defined multi-dimensional wavelet for the modeling of multivariable processes is as below:

$$\phi_j\left(\begin{bmatrix} x_i \\ \vdots \\ x_l \end{bmatrix}\right) = \prod_{k=1}^{l} \phi\left(\frac{x_k - m_{jk}}{d_j}\right) \tag{24}$$

where m_j=[m_{j1}, m_{j2}, …, m_{jn}] ^{T} represents the translation vector and d_j=[d_1, d_2, …, d_n] ^{T} indicates the dilation vector.

As a consequence of wavelet network's ability to capture essential features of signals entirely, they are favored in identification, classification and regression problems. The multi-dimensional wavelet networks are written as follows:

$$f_p(x) = \sum_{j=1}^{p} \omega_j \Phi_j(x) \tag{25}$$

where ω_j is the weight vector for the network. Wavelet networks optimize both the position and the dilation of the network.

Despite the fact that wavelet decomposition offers significant advantages, it should be mention that constructing and storing the wavelet parameters in large dimension is most challenging and prohibitive. Sensitivity to dimensions is one of the significant challenges in signal processing. Hence, it is essential to develop algorithms which are less dimension-sensitive and can manage high volume large dimension problems. Neural networks are

well-known as one of the most influential tools to deal with problems with high dimension. Considering the many advantages offered by wavelet theory and conventional neural networks, wavelet networks has come into view as both wavelet theory and neural networks offspring. Wavelet neural networks are able to handle enormous dimension problems well. Hence, they have proliferated in recent years [57].

The structure of self-recurrent wavelet neural network (SRWNN), which is a modified form of wavelet neural networks, mainly consists of four layers [58]:

1) The input layer accepts N_i nodes corresponding to N_i input variables.

2) The mother wavelet layer or wavelon layer consists of N_w mother wavelet nodes (wavelons). Each node has a wavelons group with a self-feedback loop. The effect of the choice of different types of mother wavelet on the performance of the proposed model will be discussed. In this study, mother wavelets are in the form of the first derivative of the Gaussian function as:

$$\phi(x) = -x \exp\left(-\frac{1}{2}x^2\right) \tag{26}$$

Therefore, the wavelet ϕ_{jk} is derived as

$$\phi_{jk}(z_{jk}) = \phi\left(\frac{u_{jk} - m_{jk}}{d_{jk}}\right); z_{jk} = \frac{u_{jk} - m_{jk}}{d_{jk}} \tag{27}$$

where m_{jk} and d_{jk} are the translation and the dilation factors of the wavelets, respectively. If the weight of self-feedback neuron is shown by θ_{jk} and the previous state of the system be stored using $\phi_{jk}(n\text{-}1)$ in the memory, the j^{th} input of the k^{th} wavelet, denoted by u_{jk},

might be written as:

$$u_{jk}(n) = x_k(n) + \phi_{jk}(n-1)\theta_{jk} \tag{28}$$

The storing data in the self-feedback loop is the advantage of self-recurrent wavelet neural network to conventional wavelet neural networks. Even if the SRWNN has fewer wavelons than the WNN, it contains the memory term $\phi_{jk}(n-1)$ which stores the information of previous states of the system. Therefore, SRWNN models the complex dynamics correctly.

3) The product layer is the product of N_w wavelons as follows:

$$\Phi_j(x) = \prod_{k-1}^{N_i} \phi(z_{jk}) = \prod_{k-1}^{N_i} \left(-\frac{u_{jk} - m_{jk}}{d_{jk}} \exp\left(-\frac{1}{2}\left(\frac{u_{jk} - m_{jk}}{d_{jk}} \right)^2 \right) \right) \tag{29}$$

4) The output layer is a linear combination of the consequences obtained from the outputs of layer three as:

$$y(n) = \sum_{j=1}^{N_w} w_j \Phi_j(x) + \sum_{k=1}^{Ni} a_k x_k \tag{30}$$

where w_j, in the first term, is the connection weight between the product node and the output node. Also, the second term shows that output layer accepts N_i input values from input layers which are weighted by a_k. Along these lines, the SRWNN has a multi-input and single-output (N_i x N_w) structure. Additionally, the a_k direct weights offer the features of a direct linear feedthrough network, e.g. the initialization of network parameters based on prior knowledge and enhanced extrapolation outside of examples of the learning data

sets. The parameters of an SRWNN, which are trained in an adaptive process, construct the weighting vector W, augmented as:

$$W = \begin{bmatrix} a_k & m_{jk} & d_{jk} & \theta_{jk} & w_{jk} \end{bmatrix} \tag{31}$$

The goal of SRWNN is minimizing quadratic cost function which might be obtained by the Gradient Descent (GD) method. The quadratic cost function is defined as

$$J(n) = \frac{1}{2}\left(y_d(n) - y(n)\right)^2 = \frac{1}{2}e^2 \tag{32}$$

where $y_d(n)$ is the desired output and $\hat{y}_d(n)$ is the estimated value. The weights of the SRWNN algorithm are tuned during online training process using the adaptive learning rates (ALRs). Thus, the weighted vector is adjusted to minimize the cost function. The Gradient Descent method may be defined as:

$$W(n+1) = W(n) + \Delta W(n) = W(n) + \bar{\eta}\left(-\frac{\partial J(n)}{\partial W(n)}\right) = W(n) + \bar{\eta}\left(e(n).\frac{\partial y(n)}{\partial W(n)}\right) \tag{33}$$

Where $\bar{\eta} = diag[\eta^a \quad \eta^m \quad \eta^d \quad \eta^\theta \quad \eta^w]$ denotes the diagonal form of SRWNN learning rate matrix. Adaptive learning rates (ALRs) are used to train the weights of the SRWNN structure. Thus, the weighted vector is adjusted to minimize the cost function. It is followed by adjusting the parameters [16]. By applying the chain rule recursively, the error term is calculated as:

$$\frac{\partial y(n)}{\partial W(n)} = \begin{bmatrix} \frac{\partial y(n)}{\partial a_k(n)} \\ \frac{\partial y(n)}{\partial m_{jk}(n)} \\ \frac{\partial y(n)}{\partial d_{jk}(n)} \\ \frac{\partial y(n)}{\partial \theta_{jk}(n)} \\ \frac{\partial y(n)}{\partial w_j(n)} \end{bmatrix}^T = \begin{bmatrix} x_k \\ -\frac{w_j}{d_{jk}} \frac{\partial \Phi_j(n)}{\partial z_{jk}(n)} \\ -\frac{w_j}{d_{jk}} z_{jk} \frac{\partial \Phi_j(n)}{\partial z_{jk}(n)} \\ \frac{w_j}{d_{jk}} \phi_{jk}(n-1) \\ \frac{\partial \Phi_j(n)}{\partial z_{jk}(n)} \Phi_j(n) \end{bmatrix}^T \tag{34}$$

Eventually, the stator currents are estimated by the nonlinear function defined through SRWNN as follows:

$$\hat{I}_{xs}(t) = \sum_{j=1}^{N_w} w_j \Phi_j(x) + \sum_{j-1}^{N_i} a_k x_k = h(I_s(t-1), S(t)) \tag{35}$$
$$x \in \{a, b, c, d, e, \ldots\}$$

3.3.2 Self-Recurrent Wavelet Neuro-Fuzzy Network with Adaptive Learning Rate

Similarly, a self-recurrent wavelet neuro-fuzzy network is developed for nonlinear system identification with fuzzy inference as follows [59-60]:

$$R_l : \text{IF } x_1 \text{ is } F_{1l} \text{ ,...,} x_n \text{ is } F_{nl} \text{ , then } \hat{y} \text{ is } \omega_l \tag{36}$$

where $x_i(k)$ is the i^{th} input linguistic variable, F_{il} is the fuzzy set corresponding to x_i in the l^{th} fuzzy implication, $\hat{y}$ is estimated output, and ω_l is the l^{th} consequence link weight. The structure of wavelet neuro-fuzzy network consists of six layers:

1) The input layer accepts N_i nodes corresponding to N_i input variables.

2) Fuzzy membership layer performs membership functions F_{il} in each node. A Gaussian function is used here as the membership function.

3) The product layer is the product of all membership functions from second layer.

4) Normalized firing strengths layer obtains the ratio of the firing strength of every rule to the sum of those rules as:

$$v_l = \frac{\Phi_j(x)}{\sum_{j=1}^{L} \Phi_j(x)} \tag{37}$$

where $0 < v_l < 1$ and $\sum_{l=1}^{L} v_l$=1.

5) The mother wavelet layer with self-feedback neurons layer consists of N_w mother wavelet nodes (wavelons), which are the first derivative of a mother wavelet Gaussian function. O_l , the output of layer 5, is represented by:

$$O_l = \left(\sum_{i=1}^{n} \phi_{il}\right).w_l.v_l = \left(\sum_{i=1}^{n} -z_{il} \exp\left(-\frac{z_{il}^2}{2}\right)\right).w_l.v_l \tag{38}$$

6) The output of layer is the summation of all incoming signals from previous layer as follows:

$$\hat{y} = \sum_{l=1}^{L}\left(\sum_{i=1}^{n} -z_{il} \exp\left(-\frac{z_{il}^2}{2}\right)\right).w_l.v_l \tag{39}$$

The weighting vector of network is represented by:

$$W = \begin{bmatrix} \mu_{il} & \sigma_{il} & b_{il} & a_{il} & w_l \end{bmatrix} \tag{40}$$

By applying the gradient descent method to minimize the error, the training algorithm for the parameters are obtained by

$$W(k+1)=W(k)+\Delta W(k)=W(k)+\bar{\eta}\left(-\frac{\partial J(k)}{\partial W(k)}\right)=W(k)+\bar{\eta}\left(e(k).\frac{\partial J(k)}{\partial W(k)}\right) \tag{41}$$

$\bar{\eta}=diag[\eta^{\mu}\ \eta^{\sigma}\ \eta^{b}\ \eta^{a}\ \eta^{w}]$ denotes the diagonal form of learning rate matrix and the error is shown by $e(k)=y-\hat{y}$. Also:

$$\frac{\partial y(k)}{\partial W(k)}=\begin{bmatrix}\frac{\partial y(k)}{\partial \mu_{il}(k)}\\ \frac{\partial y(k)}{\partial \sigma_{il}(k)}\\ \frac{\partial y(k)}{\partial m_{il}(k)}\\ \frac{\partial y(k)}{\partial d_{il}(k)}\\ \frac{\partial y(k)}{\partial w_{l}(k)}\end{bmatrix}^{T}=\begin{bmatrix}\left(O_l(k)-\hat{y}(k)\right)v_l\frac{x_i(k)-\mu_{il}(k)}{\left(\sigma_{il}(k)\right)^2}\\ \left(O_l(k)-\hat{y}(k)\right)v_l\frac{\left(x_i(k)-\mu_{il}(k)\right)^2}{\left(\sigma_{il}(k)\right)^3}\\ v_l w_l\phi_{il}\frac{\left(z_{il}(k)\right)^2-1}{z_{il}(k)d_{il}(k)}\\ v_l w_l\phi_{il}\frac{\left(z_{il}(k)\right)^2-1}{d_{il}(k)}\\ v_l\sum_{i=1}^{n}-z_{il}(k)\exp\left(-\frac{\left(z_{il}(k)\right)^2}{2}\right)\end{bmatrix}^{T} \tag{42}$$

Eventually, the stator currents are estimated by the nonlinear function defined through SRWNFN as follows:

$$\hat{I}_{xs}(t)=\sum_{l=1}^{L}\left(w_l.v_l.\left(\sum_{i=1}^{n}-z_{il}\exp\left(-\frac{\left(u_{il}-m_{il}\right)^2}{2d_{il}}\right)\right)\right)=h(I_s(t-1),S(t)) \tag{43}$$
$$x\in\{a,b,c,d,e,...\}$$

The SRWNN and SRWNFN identifiers have the capability of storing information

temporarily. Thus, they can approximate the dynamics of the system accurately. Also, they have an on-line adapting ability for modeling the dynamics of the system. Unlike the conventional Gradient Descent algorithm which uses constant learning rate, Adaptive Learning Rates (ALR) are applied for the online training of the identifier, which offers the ability of discrimination between fault occurrence and any changes in the systems which causes abnormality. The proposed methods can adjust model parameters on-line. Consequently, generated residuals are not affected by system operation variations [61].

In spite of the AI methods which are dependent on variables considered in training, the proposed approach is robust to uncertainties, noise, and load changes. Residuals are insensitive to uncertainties while sensitive to faults. Parameters variation, load change, and noise create transients that could be interpreted as a fault. As nonlinear model predictor has adaptive learning rates, it offers the ability of self-tuning to parameters. Such rates are slightly lower than switching frequency. Therefore, the model parameters are not able to be regulated quickly in switch failure cases. However, when the load varies, the temperature fluctuates, or motor operation mode changes, learning rates are self-tuned relatively faster. In this manner, the system can adapt itself to variation to discriminate between fault and operation mode changes and avoid possible false alarms.

3.4 Residual Evaluation for Fault Detection

The primary goals of residual evaluation procedure in the FDI schemes are, first, to make a binary decision on whether a fault has occurred in the system and, second, to identify the fault characteristics and location. The decision on the occurrence of faults in

FDI systems is based upon residuals that reflect the faults. The original idea behind the residual evaluation is to create such a (physical) feature of the residual signal that allows a reliable detection of the fault.

The critical difficulty with residual evaluation is that generated residuals are uncertain, corrupted by noise, disturbances in general. Moreover, uncertainties of the model-based methods are affecting the accuracy of the residuals. Therefore, it is essential to select the most efficient method to evaluate the residuals depend on the type and characteristics of the considered system.

Different strategies to evaluate the residuals have been studied in literature such as threshold logic, statistical decision making, pattern recognition and machine learning methods, and decision making based on intelligent systems. There exist two major residual evaluation strategies depending on the type of the system, stochastic or deterministic systems, two procedures for residual evaluation are considered: statistic testing (for stochastic systems) and norm-based evaluation (for deterministic systems). The second procedure reduces the computational complexity of the systems containing deterministic disturbances or system uncertainties [62].

Besides of less on-line calculation, the norm-based residual evaluation allows a systematic threshold computation using well-established robust control theory. In the pioneering works, root-mean-square (RMS) norms and adaptive threshold, H_∞ framework, and L_2 norm are adopted as the residual evaluation functions. In practice, also other kinds of features are used for the same purpose, for instance, the absolute value in the limit monitoring scheme. The mathematical tools for the statistic testing and norm-based

evaluation are different. In this study, three different residual evaluation approaches based on the application are applied. Firstly, a strategy based on conventional threshold criteria is proposed which evaluates the residuals for n-phase stator currents and report the fault once at least one of the residuals go beyond the pre-defined limit. Secondly, the support vector machine method is applied as a supervised statistical learning decision-making boundary surface. Thirdly, the linear discriminant analysis is used.

3.4.1 Pre-Defined Threshold

The most simple and straightforward method for fault decision consists in a threshold test of the residual or a measure formed from the residual. If constant thresholds are used one has to cope with the problem of the effects of unknown inputs including modelling errors. This means, if the threshold is chosen too small, false alarms occur, if the threshold is chosen too large, small faults cannot be detected. The effect of the modelling errors depends on the operating conditions of the process.

If there are no uncompensated unknown input effects on the residuals due to a perfect decoupling, then the thresholds diminish to zero. Otherwise, thresholds different from zero have to be assigned. In practice, there is usually such a high number of unknown inputs that, in the face of the limitations of available measurement information, a complete decoupling from all unknown inputs is hardly achievable even if the model uncertainties are structured. The situation becomes even worse if the uncertainties are unstructured, in which case a perfect decoupling in the residual generation stage is impossible. Hence the residuals or any decision functions built from them always deviate from zero even if no

fault is present[63].

Due to the model uncertainties, the residual is affected by the known input signals, which are, in general, time varying. To achieve a threshold that is as tight to the residual as possible, the threshold should also depend on the known input signals. Unknown operational conditions, noise, uncertainties, and failures are different inputs to define the thresholds.

An extensive study on the computation of constant thresholds in linear systems can be found in [64], where different kinds of thresholds both under deterministic settings using signal norms of the unknown inputs and stochastic settings using statistical properties of unknown inputs are proposed. Different from constant thresholds, the variable thresholds vary with the instantaneous values of the process input and some system parameters. Different techniques have been reported in the literature for designing thresholds in FD systems. These thresholds can broadly be classified as (1) constant thresholds and (2) adaptive thresholds. The most straightforward approach is to decide that a fault has occurred when the instantaneous value of a residual vector exceeds a constant threshold. The constant thresholds are designed by considering the upper bound of the unknown inputs and admissible uncertainties.

The selection of the threshold dramatically influences the efficiency of the monitoring system. A threshold is a tolerant limit for disturbances and model uncertainties under fault-free operation conditions. A lower threshold setting usually leads the monitoring system to be subject to higher false alarms, and a higher threshold setting causes a typically higher rate of missed detection. Therefore, based on the chosen

evaluation function, the threshold can be defined by:

$$J = \sup_{f = 0,ds,\Delta} J_e \tag{44}$$

where *ds* and Δ represent the disturbance and model uncertainties, respectively. J_e represents the feature of the evaluated residual signal, which could be RMS or peak.

After the residual evaluation and threshold setting, a decision logic has to be carried out. The simplest decision logic is to compare the feature of the evaluated residual J_e with the pre-determined threshold J_{th}.

In this way, decision is made as follows:

$$\begin{cases} J_e \le J_{th} & \text{Normal operation} \\ J_e > J_{th} & \text{Fault operation and alarm} \end{cases} \tag{45}$$

3.4.2 Support Vector Machine Classifier (SVM)

Pattern recognition and data mining tools are well-stablished as decision units to evaluate the residuals. In the data-driven design framework, the multivariate statistical process monitoring techniques utilize input and output information of the process. The basic idea of multivariate statistical process monitoring techniques is to extract the statistical features to describe the desired process behavior from huge amount of process data. The extracted statistical features are used for later process monitoring purpose. The ability to tackle large number of highly correlated variables shows the multivariate statistical methods significant advantage. To successfully apply the standard multivariate statistical schemes, it is a primary assumption that a linear process runs under stationary operating conditions. However, the process dynamics and nonlinearity are very important

aspects in industrial processes.

Due to their simple concluding remarks forms and low design efforts, the multivariate statistical methods are widely applied to process monitoring in numerous large-scale industrial applications. Within multivariate statistical framework, approaches like support vector machines, linear discriminant analysis, principal component analysis, and their variants are intensively studied employed to deal with process dynamics and nonlinearity. Support Vector Machine is one of these methods.

Support Vector Machine, as a machine learning method based on statistical learning theory and structural risk minimization principle, is well-recognized in classification and regression problems [65-67]. Let x_i and y_i be the input and output vector of a given sample set of data as:

$$S = \{(x_i, y_i) \mid x_i \in H, y_i \in \{\pm 1\}, i = 1,...,1\} \tag{46}$$

where H is the features set and l is the number of training samples. In order to construct the support vector machine, an optimal hyper-plane (decision surface) is determined by the following optimization problem:

$$\begin{aligned} &\textit{Minimize} \quad T(w,\xi) = \langle w.w \rangle + c\sum\nolimits_{i=1} \xi_i \\ &\textit{Subject to} \begin{cases} y_i.\left(\langle w.x_i \rangle + b\right) \geq 1 - \xi_i \quad i = 1,...,l \\ \xi_i \geq 0 \\ C > 0 \end{cases} \end{aligned} \tag{47}$$

where ξ_i is a slack variable and c is a the penalty constant. By incorporating kernels and rewriting it in Lagrange multipliers, the dual quadratic optimization problem is given by

$$\begin{aligned} &Maximize \;\; W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ &Subject\ to \begin{cases} \sum_{i=1}^{l} \alpha_i y_i = 0 \quad i = 1, \ldots, l \\ 0 \le \alpha_i \le C \quad i = 1, \ldots, l \end{cases} \end{aligned} \tag{48}$$

α_i is Lagrange multiplier. Thus, by solving the dual optimization problem, the decision surface of SVM is obtained by:

$$f(x) = sign\left(\sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + b \right) \tag{49}$$

$K(x, x_i)$ is the kernel function. There are three commonly used kernel functions, namely, Gaussian radial basis function (GRBF) kernel, Polynomial kernel, and Sigmoid kernel, that in this study radial basis function (RBF) kernel is used which is shown as follows:

$$K(x, x_i) = \exp\left(-\left\| x_i - x_j \right\|^2 / 2\sigma^2 \right) \tag{50}$$

The decision surface of SVM defines an n-dimensional boundary in the space to limit the mapped data set of normal operation. An example is illustrated in Figure 12. The r_1, r_2, r_3 are corresponding to generated residuals.

3.4.3 Linear Discriminant Analysis Classifier

Linear Discriminant Analysis, as the generalized form of Fisher's linear discriminant, is a multivariate statistical analysis method concerned with the relationship between a categorical variable and a set of interrelated variables [68-69].

The data set of *n* samples in an m-dimensional space is denoted by $X = [x_1 \ldots x_n]$

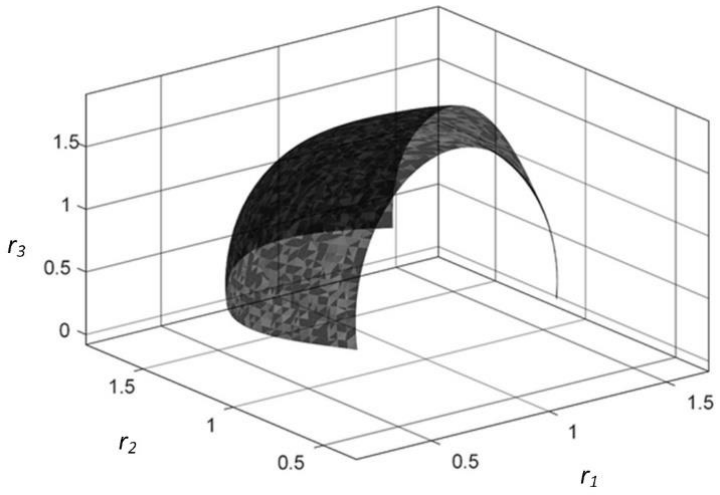

Figure 12 Example for SVM decision surface in 3D space.

and the data set is clustered into C classes. The between-class scatter matrix S_b and the within-class scatter matrix S_w with overall mean $\bar{x}$ are defined as:

$$S_b = \frac{1}{N}\sum_c n_i(\mu_c - \bar{x})(\mu_c - \bar{x})^T$$
$$S_w = \frac{1}{N}\sum_c \sum_{i \in C} n_i(\mu_c - \bar{x})(\mu_c - \bar{x})^T \quad (51)$$

The goal of the linear discriminant analysis is finding a transformation to maximize the between-class scatter S_b and minimizes the within-class scatter S_w, which is fulfilled by maximizing the following objective function:

$$J(w) = \frac{w^T S_b w}{w^T S_w w} \quad (52)$$

when S_b and S_w are Hermitian..In order to classify the data into two clusters, the distance concept is used for multidimensional data discrimination. Suppose there are two different populations G_1 and G_2, where the mean vectors are μ_1 and μ_2, respectively. Also, the

covariance matrices of G_1 and G_2 are $\Sigma_1 = \Sigma_2 = \Sigma$. It is shown that by using the concept of Mahalanobis square distance between a sample and a population, distance discriminant rules might be represented by the following rules [55]:

$$\begin{cases} x \in G_1 \ ; \ if \ \ W_1(x) \le W_2(x) \\ x \in G_2 \ ; \ if \ \ W_1(x) > W_2(x) \end{cases} \quad \text{where} \quad \begin{cases} W_1(x) = a_1^T x + b_1 = \left(\Sigma^{-1}\mu_1\right)^T x - \frac{1}{2}\mu_1{}^T\Sigma^{-1}\mu \\ W_2(x) = a_2^T x + b_2 = \left(\Sigma^{-1}\mu_2\right)^T x - \frac{1}{2}\mu_2{}^T\Sigma^{-1}\mu \end{cases} \tag{53}$$

where given sample x is referring to a point in n-dimensional space.

3.5 Residual Evaluation for Fault Isolation

In multiphase drives, the isolation step includes two main procedures, i.e. detecting the faulted phase leg and locating the failed switch. The number of residuals is same as the number of phases. It is worth mentioning that experimental studies illustrated the possibility of having more than one residual exceeding the threshold. Because not only the failed phase, but also the other phases are affected by fault influence. Hence, it is challenging to make decision on the residual which is corresponding to failed phase.

Residual evaluation for fault isolation purpose is discussed in next two chapters for multiphase two-level conventional converters and multiphase multilevel NPC converters, respectively.

4. REAL-TIME FAULT DETECTION AND ISOLATION STRUCTURE FOR MULTIPHASE CONVENTIONAL INVERTER

The fault isolation task in a fault diagnosis scheme is a signal processing (post-processing) task to determine the characteristics and location of the fault in the process under consideration. Efficiency and complexity of developed fault isolation schemes rely on: (1) the total number of possible fault scenarios, (2) the potential distribution of the faults, (3) the characteristic features of each fault, and (4) prior knowledge of the faults.

Due to the number of the possible fault scenarios and the distribution of faults in the system under consideration, the fault isolation scheme might be decomposed to different levels and be solved in a hierarchical framework.

Considering the decoupling of unknown inputs in FDI process, it was recognized that identifying and isolating the faults can also be formulated as a multi- unknown input single-known output (or multi-known output) decoupling problems. For this purpose, faults are then handled as unknown inputs. If it is possible to design a bank of residual generators that solves unknown input decoupling fault detection and isolation for each possible case, a fault isolation is then achieved. Fault isolation strategy delivers generated residuals to the corresponding fault classes/clusters. In fact, fault isolation problem is a decision unit to perform a classification/clustering model. However, depending on the application, the decision unit, which splits the space to the number of possible fault cases, might be considered as a hierarchical structure.

In multiphase multilevel inverters, the location of the fault has two different location indicators, first, the phase of the faulted leg, second, the specific location of the faulted switch in the leg. Therefore, the fault isolation problem is considered as a two steps task. In this chapter, a hierarchical fault isolation scheme for the multiphase conventional two-level inverter is presented.

4.1 Failed Phase Identification Using Principal Component Analysis

The supervisory system in the n-phase drives creates n real-time residuals, and thus transfer the generated residuals to the decision unit. The fault detection scheme initiates the diagnosis procedure if at least one of the residuals exceed the threshold, or the binary decision-making unit delivers the residuals to fault operation.

It is worth mentioning that in multiphase systems, having more than one residual exceeding the threshold is possible, because not only the failed phase, but also the other phases are affected by fault influence. Due to star connection, all of the stator currents show unpredicted post-fault behavior. The effect of the fault on the output waveform depends on the time of fault occurrence and amplitude of observed output waveform as well. Hence, it is challenging to interpret on the residual which is corresponding to failed phase. It is essential to apply a method which can detect the phase related to the fault even in the case of having a lower deviation in the corresponding phase current.

Multivariate statistical analysis methods, referring to advanced statistical methods for exploring the relationships among more than two variables simultaneously, are a practical tool to recognize the failed phase. This study aims to apply one of the

multivariate techniques, principal component analysis (PCA), to disclose the relationship between the post-fault current behaviors and the phase of the failed leg. If more than one residual exceeds the threshold, principal component analysis (PCA) might be able to recognize the phase-correlated to the failed switch. PCA emphasizes variation and brings out the strongest patterns in a dataset. It makes easy to explore and visualize the significant dissimilarities among dataset data [71-72]. PCA search for a linear combination of variables such that the maximum variance is extracted from the variables. Then, after removing this variance, it seeks the second linear combination explaining the maximum proportion of the remaining variance, and so on. This method is called the principal axis method and results in orthogonal and uncorrelated factors. PCA analyzes common and unique variance.

In fact, Principal components, reflecting both common and unique variance of the variables, are variance-focused approach aiming to reproduce both the total variable variance with all components and to reproduce the correlations. The principal components are set of linear combinations of the original variables weighted by their contribution to explaining the variance in a particular orthogonal dimension Eigenvalues, also called characteristic roots. The eigenvalue for a given factor measures the variance in all the variables which are accounted for by that given factor. Also, the ratio of corresponding eigenvalues is the ratio of explanatory importance of the given factors for the variables. If an element has a low eigenvalue, then it means a small contribution to the explanation of variances in the variables, which may be ignored as redundant with more critical factors. Eigenvalues measure the amount of variation in the total sample accounted for by each

element. An element's eigenvalue may be computed as the sum of its squared factor loadings for all the variables.

Taking these into consideration, a small sliding window is considered over the time axis. If t denotes the time instant corresponding to the earliest moment that one of the residuals exceeded the threshold, P indicates half of the arbitrary sliding window length over time and I_s represents the measured stator currents of n-phases over this time interval; then the matrix Z might be constructed as:

$$Z = \begin{bmatrix} I_{as}(t-P) & \dots & I_{as}(t-1) & I_{as}(t) & I_{as}(t+1) & \dots & I_{as}(t+P) \\ I_{bs}(t-P) & \dots & I_{bs}(t-1) & I_{bs}(t) & I_{bs}(t+1) & \dots & I_{bs}(t+P) \\ I_{cs}(t-P) & \dots & I_{cs}(t-1) & I_{cs}(t) & I_{cs}(t+1) & \dots & I_{cs}(t+P) \\ I_{ds}(t-P) & \dots & I_{ds}(t-1) & I_{ds}(t) & I_{ds}(t+1) & \dots & I_{ds}(t+P) \\ I_{es}(t-P) & \dots & I_{es}(t-1) & I_{es}(t) & I_{es}(t+1) & \dots & I_{es}(t+P) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (54)$$

The highest eigenvalue in the covariance of matrix Z indicates the faulted leg. This matrix includes the information of pre-fault instant, fault instant, and post-fault instant. Thus, it can compare the stator currents behavior over the time interval with the length of 2P. For the n-phase machine and assuming one faulted phase, the n-dimension input is mapped by PCA on a one-dimension eigenspace of one eigenvector, which is calculated by the covariance of the data set Z. The covariance enables PCs to divide the data set into sub-clusters and picture the most different component, assuming that only one phase is faulty. Thus the first PC is observed. The highest eigenvalue and eigenvector correlated to PC_1 indicates that which eigenvector among the dataset has the significant variation. Accordingly, it represents the faulted phase. In this manner, the phase of the failed switch is identified by:

$$\text{Phase } x \text{ is Faulted}: \{x \mid D_x^Z = \max(D^Z)\ ;\ D^Z = diag(\text{cov}(Z))\} \tag{55}$$

Therefore, first, the fault is detected by the binary decision-making unit. Second, the isolation scheme is initiated. Matrix Z is formed based on a window of data. Maximum calculated covariance of matrix Z represents associated phase to fault with the most dissimilarity, which is interpreted as fault operation. Third, detecting the faulted phase is followed by isolating the failed switch of the identified faulted phase.

4.2 Failed Switch Isolation Using Real-Time Isolator

Isolation of the failed switch might be seen as a classic classification or clustering problem, which delivers the transient post-fault response to the associated pre-defined class. In multiphase systems, the switch isolation is the lower level in the hierarchy.

In the conventional inverters with two switches per leg, the classification problem is simplified to a basic classification/clustering problem to choose one of two possible cases. Conventional classifications fall into the domain of supervised learning (grouping set of objects based on training datasets of labeled data), while classification falls into the unsupervised learning domain (grouping set of objects based on similarities).

This study will first propose a real-time criterion to isolate the failed switch. Further, it will compare its performance, speed, and robustness with two different basic classifiers, namely, support vector machine and linear discriminant analysis.

In two-level converters, it is simple to locate the failed switch, because the sign of stator currents are representing commutating switches. Top switches conduct in half positive of the period, while bottom switches are on in half positive of the period. If the

top switch fails, the actual current of failed phase decreases, but the identifier is still predicting a non-zero value for current, and the estimated current is still tracking the normal value. Hence, the predicted current value is higher than the measured value. Quite the opposite, in the case of bottom switch failure, the predicted negative current has a lower value relative to measured negative current value.

As a result, the sign of the residual depends on the fault location. Needless to say, that input of this decision-making function is the residual of the faulted phase, which was recognized in failed phase identification step using PCA. The online decision can be made using the following function:

$$\begin{cases} sign(r_x) = sign(\hat{I}_x(t_d) - I_x(t_d)) > 0 & \text{Top switch is open circuited} \\ sign(r_x) = sign(\hat{I}_x(t_d) - I_x(t_d)) < 0 & \text{Bottom switch is open circuited} \end{cases} \quad (56)$$

Considering the implementation complexity, latency, and preparatory computations criteria, the proposed fault isolation methodology, including principal component analysis and the real-time fault isolator, is uncomplicated and straightforward.

4.3 Overall Fault Diagnosis Scheme for Multiphase Conventional Inverters

The complete configuration of the supervisory, automatic control using indirect field oriented control strategy, and fault diagnosis scheme for a five-phase two-level inverter is illustrated in Figure 13.

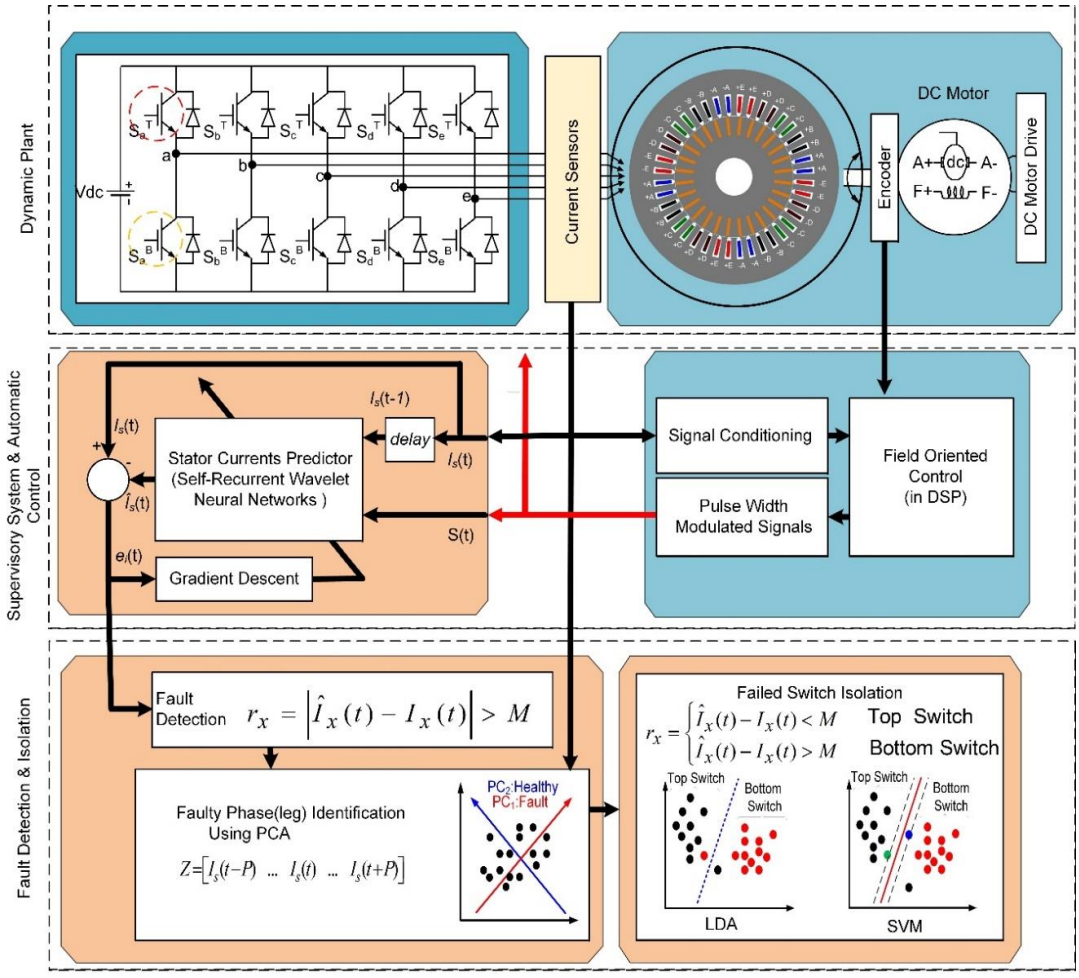

Figure 13 Overal fault diagnosis scheme for multiphase conventional two-level converter.

The structure of the proposed fault diagnosis scheme, which includes several signal processing tasks; (1) control strategy, (2) the adaptive nonlinear identifier for supervisory purpose to approximates the currents and generate fault residual, (3) threshold-based decision unit to detect the abnormalities, (4) faulted phase leg identification using principal component analysis, and (5) failed switch isolation using either real-time fault isolator, the support vector machine classifier, or the linear discriminant analysis classifier.

It worth mentioning that the fault identification and isolation tasks are taken to action only once an abnormality is detected. The supervisory system predicts the nonlinear model adaptively using the concept of model-based techniques. The nonlinear identifier, self-recurrent wavelet neural network, estimate the stator currents based on the previous stator current observations and generated pulse width modulated switching signals. Then, the difference between predicted currents and measured currents are interpreted as the fault residuals. In the n-phase machine, n residuals are generated. Decision-making unit observer the generated residuals and initiate the fault diagnosis scheme once one of the residuals exceeds the pre-defined threshold.

It worth mentioning that input of SRWNN model predictor is defined as equation (22). Number of wavelons and dynamic order of input are design parameters, which can be optimized by minimizing the error (residual) values in normal operation. Generated residuals are given to isolator procedure; either real-time isolator, LDA, or SVM.

4.4 Experimental Results

To validate the proposed diagnosis approach, an experimental evaluation set up is performed by using a five-phase induction motor coupled to a dc-motor which is programmed to operate as a dynamometer for load profile change purpose. The motor is driven by a five-phase conventional two-level converter which is built using two three-phase inverters. The vector control (Indirect Field Oriented Control) strategy is implemented by the Texas Instruments TMS320F28335 Delfino™ digital signal

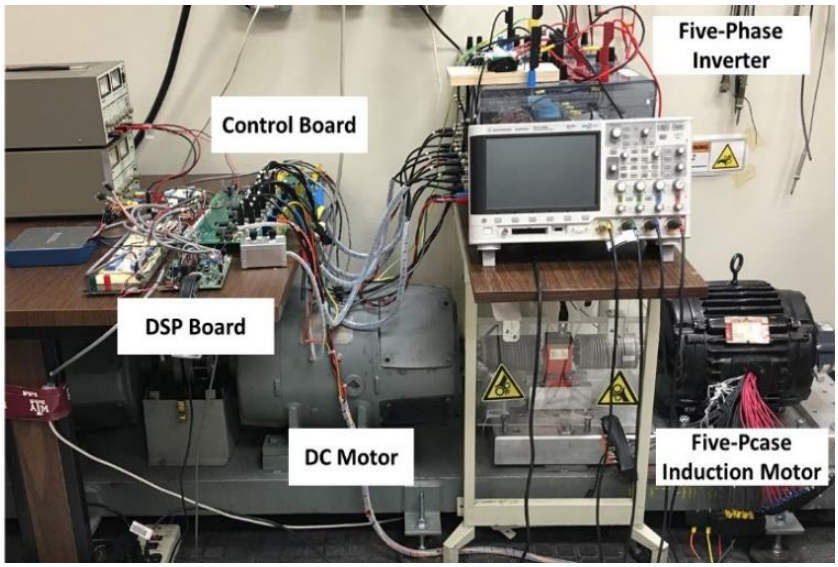

Figure 14 Experimental setup.

Table 3 Induction motor characteristics.

Rated power	5.6 kW
Rated speed	1800 rpm
Pole number	4
L_m	0.117 H
L_{ls}	0.0127 H
R_s	0.6424 Ω
L_{lr}	0.0063 H
R_r	0.373 Ω

controller. The configuration of the experimental setup is shown in Figure 14. Motor specifications are listed in Table 3. The switching frequency of the inverter is set as 6 kHz, and the motor is running at speed 0.3PU.

As it is mentioned, existing fault diagnosis methods, which are making the decision based on waveform patterns, require at least one cycle of fundamental waveforms. In low-speed applications, the detection time is more challenging, because the more extended

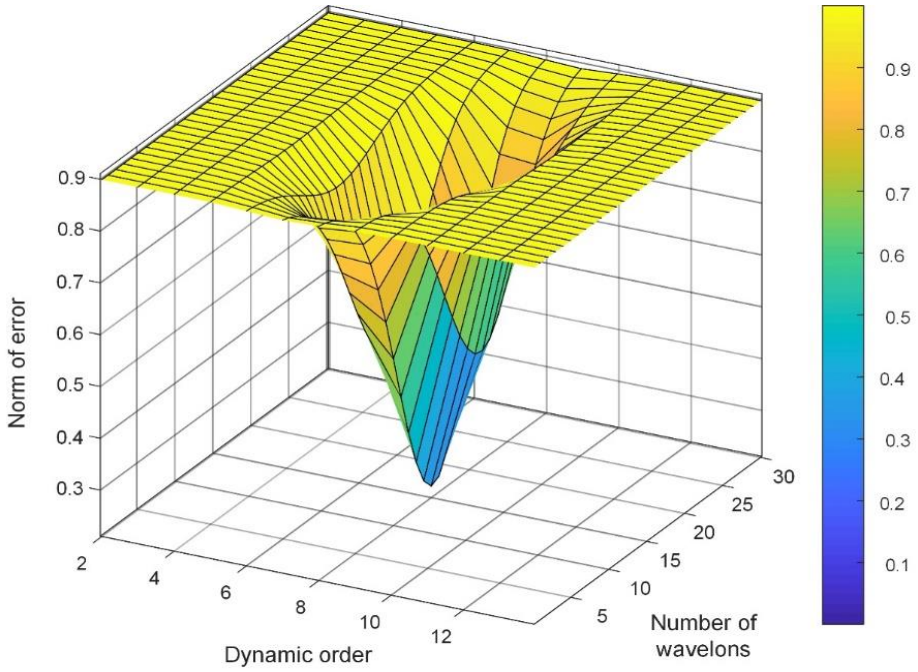

Figure 15 Optimized number of wavelons and the identifier dynamic order.

period is needed to form a trajectory and recognize the associated fault pattern. Thus, the low-speed operation is targeted to show the performance of proposed FDI scheme. Stator currents are estimated based on real-time switching patterns obtained from DSP and previous measured current values. Afterwards, the predicted and measured currents are compared to generate the residuals. The currents are predicted by SRWNN model with 15 wavelons, *Nw*. The dynamic order, N_d, is considered to be 8 as well. To optimize the values, the norm of error is calculated, which is plotted in Figure 15. In the multiphase two-level inverters, the potential fault scenarios are realized as upper or lower part faults. Due to the symmetrical feature of the motor, it is sufficient to investigate the failure condition for one phase. Therefore two open-circuited faults for ${T_a}^T$ and ${T_a}^B$ are explored.

Figure 16 indicates the measured current, estimated current and generated residual under the fault at $T_a{}^T$. The timing of generated residual and measured current is compared in Figure 17 as well. As it is illustrated, the error between estimated and measured values which is plotted as the fault residual exceeds the pre-defined threshold M=±1. The fault happens at t_0=0.5569 second and residual crosses the threshold in at t=0.5572 seconds. Similarly, the measured stator current, estimated current, and generated residual under the fault in T_a^B are indicated in Figure 18 and Figure 19. Generated residual under fault operation exceeds the threshold in less than 0.3 milliseconds after fault occurrence. The fault occurs at t_0=0.9115 seconds, and it is detected at t=0.9119 seconds.

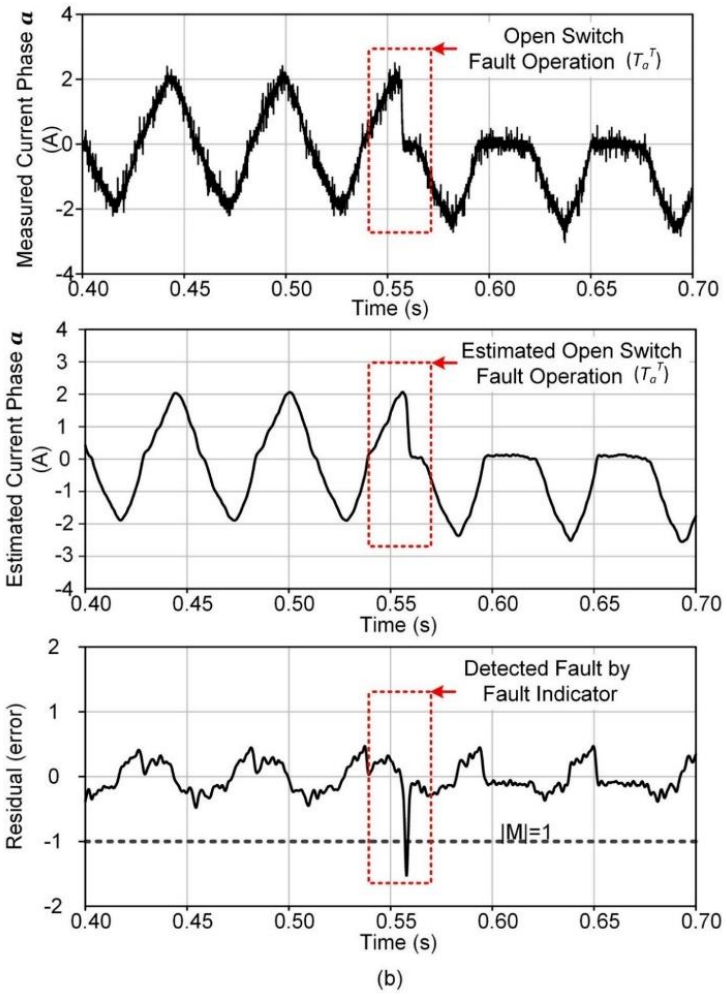

Figure 16 Measured and estimated current, and generated residual under the fault at T_a^T.

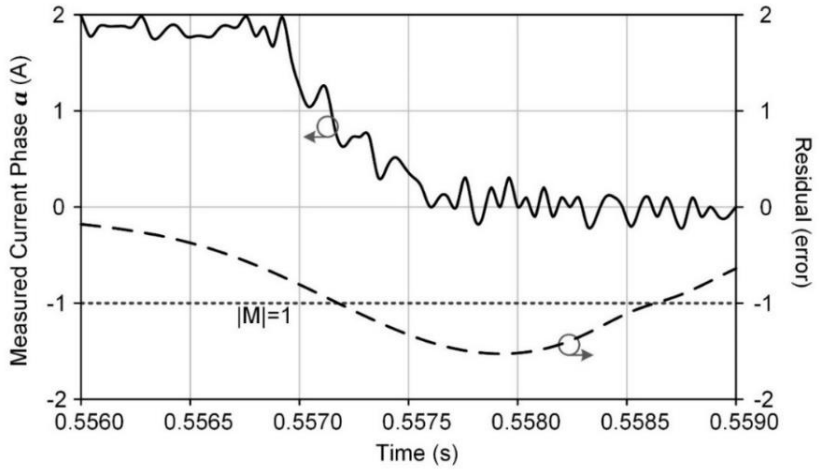

Figure 17 Measured current and generated residual under the fault at T_a^T.

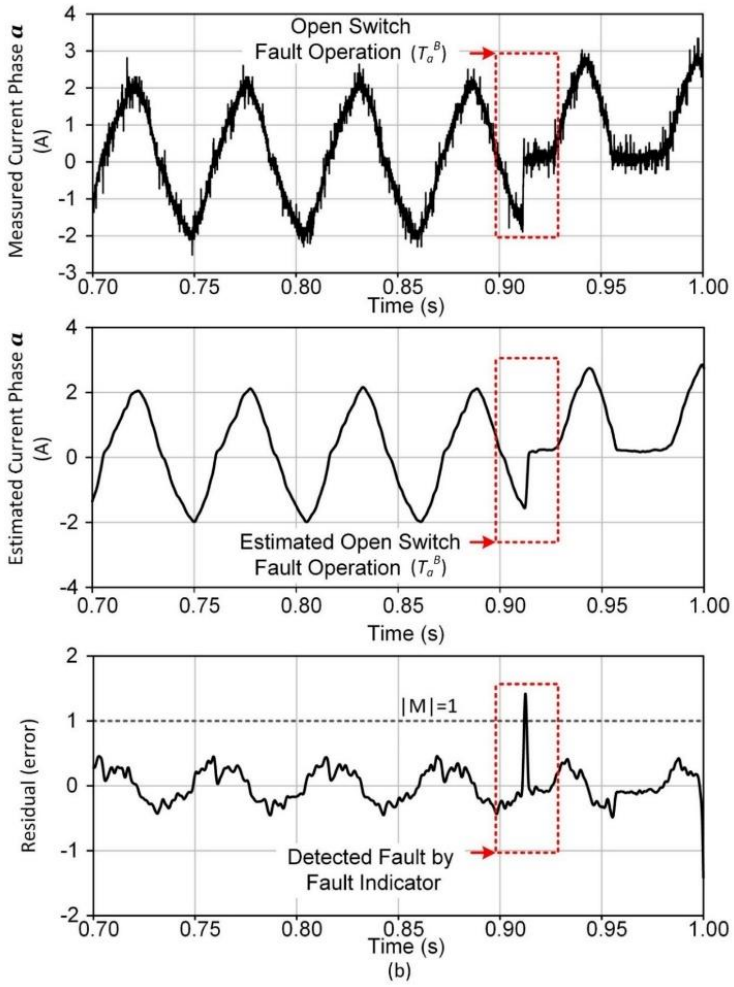

Figure 18 Measured and estimated current, and generated residual under the fault at T_a^B.

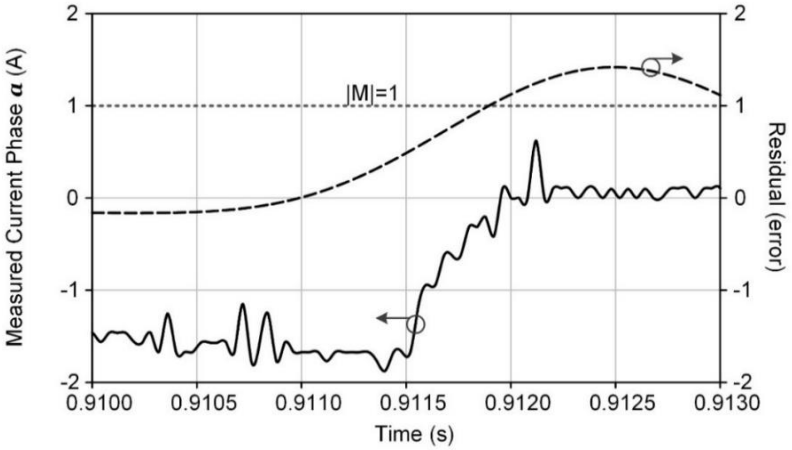

Figure 19 Measured current and generated residual under the fault at T_a^B.

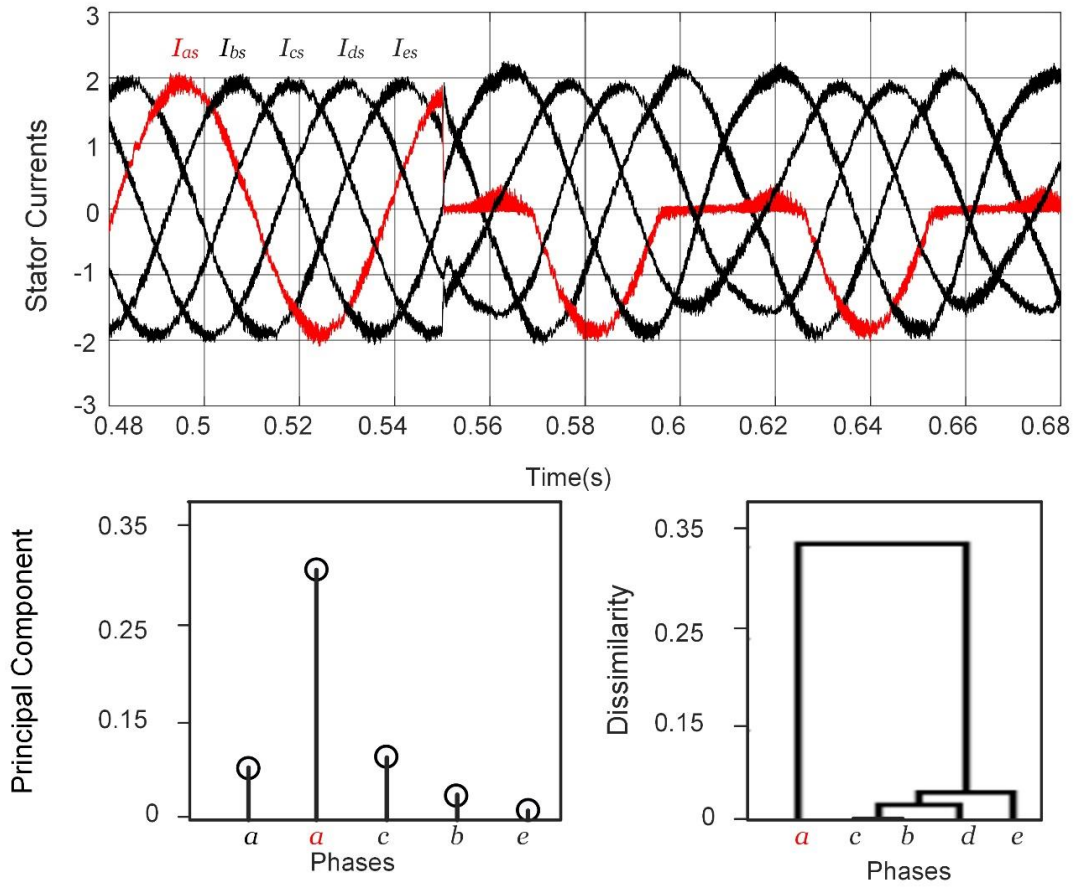

Figure 20 Stator currents, first principal component, and the clustering diagram for the fault in phase a (Second phase).

To identify the faulted phase using the principal component analysis method, the first principal component, PC_1, is calculated. Figure 20 shows, if the fault occurs in phase a, PC_1 corresponding to phase a will have the highest value. Also, clustering diagram confirms that phase a, first phase, has the most different behavior which is interpreted as the failed phase.

The location of the failed switch can be identified by either the sign of fault

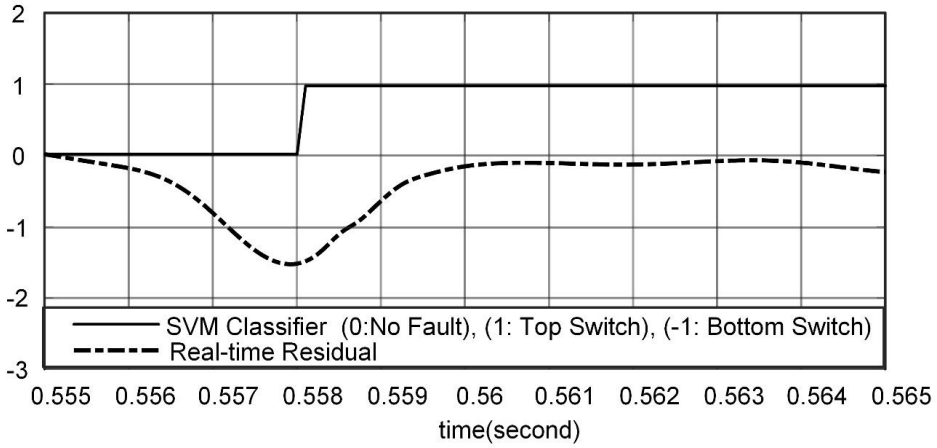

Figure 21 Comparison of the proposed real-time fault isolator with SVM classifier for fault in T_a^T.

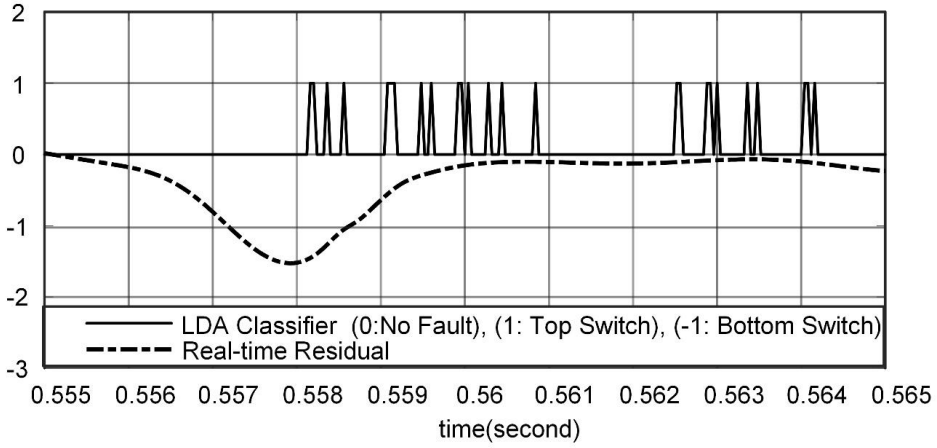

Figure 22 Comparison of the proposed real-time fault isolator with LDA classifier for fault in T_a^T.

indicator or applying a conventional supervised classifier. For comparison purposes, two different types of classifiers, support vector machine and linear discriminant analysis are employed, and their results are compared to the proposed real-time fault isolator. One of the most significant challenges in the art of machine learning is training while applying proposed method isolates the failed switch accurately without necessitating the training process. Proposed fault indicators criteria benefit from real-time detection capability

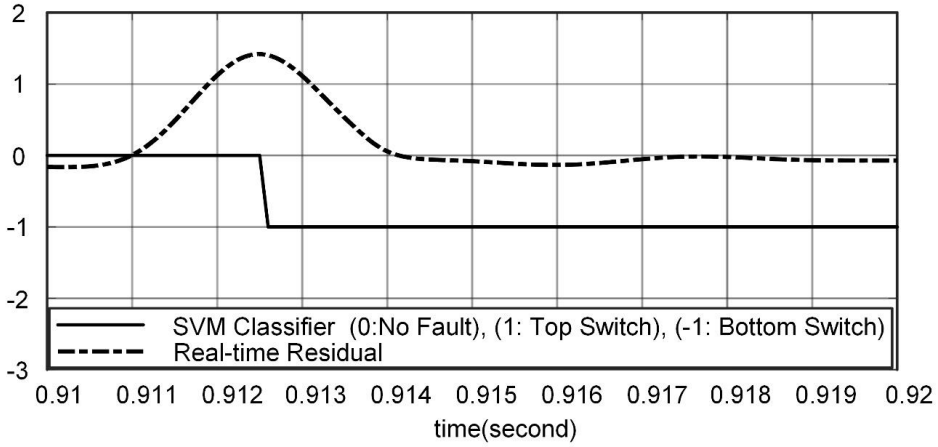

Figure 23 Comparison of the proposed real-time fault isolator with SVM classifier for fault in T_a^B.

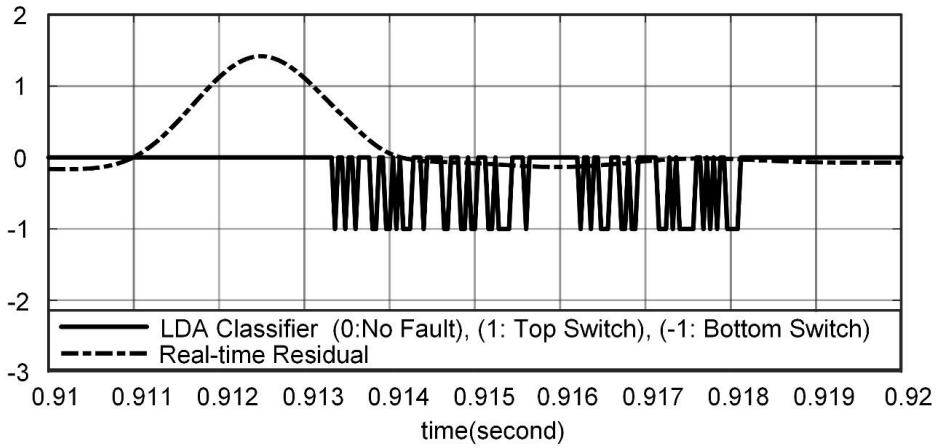

Figure 24 Comparison of the proposed real-time fault isolator with LDA classifier for fault in T_a^B.

without any prior knowledge of potential fault characteristics, while modeling an efficient classifier require training process. The results of SVM and LDA classifiers for under fault operation at T_a^T, comparing with proposed real-time fault identifier are presented in Figure 21 and Figure 22, respectively.

A similar comparison for the fault in T_a^B is presented in Figure 23 and Figure 24. The output of classifiers belongs to {1,-1, 0}, which represent the top switch, the bottom

switch and the no-fault states correspondingly. Both of these classifiers are supervised, which means they need training process. On the other hand, The sign of generated residual for the failure in the top switch of phase a, T_a^T, is negative, while it is positive for the operation under the failure in bottom switch of phase a, T_a^B. Applying proposed method isolates the failed switch accurately without necessitating the training process.

The residual fault indicator does not require any prior knowledge of the system. It does not suffer from chattering problem and is more robust. Comprehensively, real-time fault residual can detect and isolate fault faster than classifiers as well. The robustness capability of the proposed fault detection method is evaluated in Figure 25. The identifier tunes the model parameters and follows the variations. As it is indicated, the motor is loaded at t=0.851; electromagnetic torque increases from 0 to about 6, which is about 0.2PU of rated value for this motor. Subsequently, stator current increases. Although the current variation could lead to a significant deviation between measured and estimated currents, the adaptive learning rates offer to adapt the model parameters faster than the rate of mechanical variations. Thus, residuals remain in the permitted region without exceeding the thresholds. In this manner, the proposed methodology is robust and reliable, which prevents the false alarms and degradations.

4.5 Chapter Summary

In this chapter, a hybrid structure is proposed for monitoring, fault detection, and isolation in multiphase two-level conventional inverters in drive applications. In fact, an adaptive model-based fault detection approach, self-recurrent wavelet neural network, is

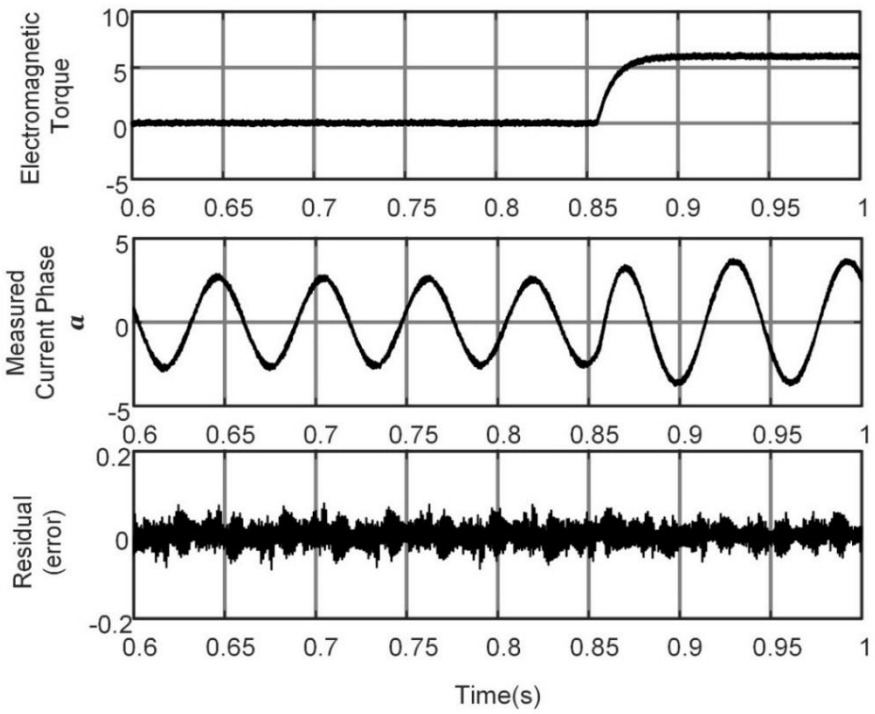

Figure 25 Robustness.

employed, which is inspired by machine learning concept.

This nonlinear identifier, first, approximates the nonlinear system including the inverter, motor, and control system, second, generate n residuals for n phases of the system. Measured stator currents values are compared with their predicted values, and any significant deviation shows abnormality. The fault detection scheme, as a decision unit using the threshold concept, initiates the fault isolation procedure. The fault isolation procedure is accomplished in a hierarchical framework, which identifies the failed phase leg and locates the failed switch respectively.

A five-phase conventional inverter validates the FDI scheme. It is shown that the SRWNN method with adaptive learning rates for tuning the model parameters offers the ability to handle the uncertainties, noise, motor parameters variation, and load variation

challenges in modeling. Due to self-tuned rates, the adaptive model can distinguish between fault, degradation, or operation point variation and the real-time residual exceeds the threshold only once a failure occurs.

The main advantage of the proposed method is the hybrid structure of identifier, which predicts the stator currents (system-level signals) using PWM switching signals (component-level signals). Thus, the performance, accuracy, and speed of identifier are enhanced considerably without deploying extra sensors or supervisory system. This approach applies to current-controlled, or control strategies compensate speed-controlled motors which post-fault transient is affected by closed-loop controllers and symptoms of abnormalities. The suggested isolation criteria are robust, fast, and does not require prior knowledge of system or training process.

5. REAL-TIME FAULT DETECTION AND ISOLATION STRUCTURE FOR MULTIPHASE MULILEVEL NPC INVERTERS USING SEMI-SUPERVISED FUZZY CLUSTERING ALGORITHM WITH PAIRWISE CONSTRAINTS

A hybrid framework to detect the switch failure in multiphase conventional inverter was proposed in previous chapters. An adaptive identifier predicts the stator currents of the motor based on previous observations from the stator currents combined with pulse width modulated switching signals. PWM signals are generated by the vector control scheme reflecting the effect of fault under controlled condition directly. Subsequently, any significant deviation from actual behavior and predicted behavior of stator currents is reported as a failure.

In this study, the same methodology is used for detecting the abnormality in multilevel NPC inverters. However, in multilevel converters switches are conducting in shorter intervals of time and the effect of switch failures on the output waveform is less significant. Also, there are redundant states, and in some of the time intervals, a pair of switches are conducting. Therefore, the observed deviation might be caused by either of conducting switches. Hence, an efficient fault isolation is required to locate the failed switch. The clustering algorithm evaluates the transient behavior and assigns the observed waveforms to the appropriate pattern representing the location of the failure.

5.1 Introduction to Semi-Supervised Learning

A variety of statistical learning and computational intelligence methods have been developed to collect data from the supervisory system for intelligent condition monitoring, fault prognosis, and fault diagnosis routines. From the perspective of machine learning, intelligent methods are mainly classified into two categories: supervised learning and unsupervised learning. Supervised learning, such as artificial neural networks, fuzzy neural networks, and support vector machine, is utilized for fault diagnosis on the labeled monitored data. On the other hand, unsupervised learning, such as self-organizing maps, adaptive resonance theory, and cluster approaches, is applied to make decisions exploiting unlabeled data.

However, labeling the data requires accurate prior knowledge of the system for the training set (detailed event types and their characteristics), which can be expensive and inadequate. The occurrence of the various faults might be infrequent in many real-world applications, as well. Therefore, deficient labeled datasets and complete expert knowledge of the system may lead to degraded performance in supervised learning and cause insufficient supervised learning eventually. Additionally, the underlying condition characteristics of the systems are likely to change over time as a function of the operating conditions or aging of the equipment. Thus, a reliable and efficient diagnostics system should be insensitive to the influence of the operating condition variations [73].

The traditional settings of the supervised learning only utilize a large amount of training labeled data to construct the model which cannot be automatically acquired in the real applications. Unsupervised learning is often inaccurate without labeled data

instruction. Also, in fault analysis scenarios, labeled and unlabeled data are captured from different sources. Thus, the combination of the labeled and unlabeled data, captured from the condition monitoring system, enhances the performance of fault analysis system. Therefore, exploiting unlabeled data to enhance the supervised learning performance has gained attraction during the past few years.

Presently, there are three machine learning paradigms for exploiting unlabeled examples, namely, semi-supervised learning, transductive learning, and active learning. Although unlabeled examples are usually different from the test examples, semi-supervised learning deals with techniques which attempt to exploit unlabeled examples automatically. Transductive learning deals with techniques that also attempt to automatically exploit unlabeled examples but assumes that the unlabeled examples are precisely the test examples. Active learning deals with methods that assume that the learner has some control over the input space, and an oracle can be queried for labels of specific instances, with the goal of minimizing the number of queries required [74].Based on the underlying assumptions, semi-supervised learning frameworks can be roughly categorized into seven types: semi-supervised learning with generative models, semi-supervised learning with low-density separation, graph-based methods, co-training methods, self-training methods, semi-supervised learning clustering, and semi-supervised learning regression[75-78].

A semi-supervised clustering framework is addressed in this study. Clustering algorithms exploit the feature-based similarities between inputs. In SSL clustering, the goal still is clustering based on similarities, but adding any information about the

relationship between samples helps to enhance the performance of the clustering process. Defining constraints based on partial prior knowledge of the system results in more efficient paradigms. There are some labeled data in the format of must-links (two data points must be in the same cluster) and cannot-links (two data points cannot be in the same cluster). These pairwise constraints, even when there is an incomplete prior knowledge of the system, result in the performance improvement. There is a trade-off between satisfying these constraints and optimizing the original clustering criterion (e.g., minimizing the sum of squared distances within clusters). Procedurally one can modify the distance metric to try to accommodate the constraints, or bias the search [79].

In this study, a semi-supervised fuzzy clustering algorithm is used which integrates the obtained data from the simulator, as unlabeled data, with the labeled experimental data. Thus, the uncertainties of the real system are brought into play, and the performance of the classifier is enhanced. A semi-supervised fuzzy clustering algorithm with pairwise constraints by competitive agglomeration is applied to locate the failed switch [79-80]. This method utilizes pairwise constraints between labeled samples and introduces two types of constraints: must-link and cannot-link.

This method is presented by a fuzzy cost function that takes pairwise constraints into account. Pairwise constraints are defined between the samples related to different failed switches. Evaluations indicate that even with a rather low number of constraints, this algorithm can significantly improve the fault switch location process. Aggregate objective function reflects a penalty corresponding to the allocation of a pair of points

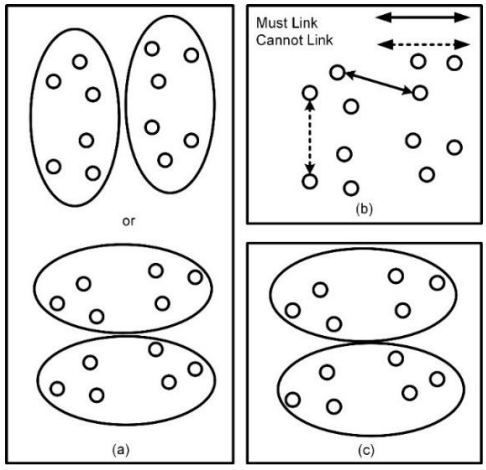

Figure 26 Semi-supervised clustering. (a) Possible clusters. (b) Pairwise constraints. (c) Final clusters.

from must-link in the different cluster and a penalty corresponding to allocating a pair of points from cannot-link in the same cluster [80-81].

Regarding the defined constraints, the cost of violating the pairwise constraints is considered additional to the main C-mean objective function. As it is shown in Figure 26, there is possibility to have more than one clustering state. According to labeled samples, must-link or cannot-link pairs are assigned to one cluster or not; then valid clustering is selected as the final option.

In multiphase multilevel NPC converter, a small data set is acquired from condition monitoring system of the experiment, and the samples are labeled corresponding to related switches. Then, extracted data points belonging to the same switch are categorized as must-link pairs, and data points corresponding to different switches are labeled as cannot-link pairs. These labeled data are combined with unlabeled data obtained from simulators.

The structure of semi-supervised fuzzy clustering with pairwise constrained competitive agglomeration to isolate the failed switch is discussed in this chapter. Also, the feature extraction process using wavelet concept is presented.

5.2 Semi-Supervised Fuzzy Clustering with Pairwise Constrained Competitive Agglomeration

For dataset $X = \{x_1, x_2, \dots, x_N\}$, N is the number of data points. $V = \{\mu_1, \mu_2, \dots, \mu_C\}$ is the prototypes of clusters and C denotes the number of defined clusters which here is the number of switches. *U* represents the matrix of the membership degrees, with u_{ik} being the membership of x_i to the cluster *k*. Let $d(x_i, m_k)$ be the Euclidian distance between the vector x_i and the cluster prototype μ_k (the center of cluster *k*).

Let *M* be the set of must-link constraints; $(x_i, x_j) \in M$ implying that x_i and x_j should be assigned to the same cluster (switch), and C be the set of cannot-link constraints; $(x_i, x_j) \in C$ implying that x_i and x_j should be assigned to different clusters.

The objective function, which consists of the summation of the fuzzy C-means objective function, cost of violating the pairwise must-link, and cost of violating the pairwise must-link cannot-link constraints, is defined as follows:

$$J = \sum_{k=1}^{C}\sum_{i=1}^{N} u_{ik}{}^2 d^2(x_i, \mu_k) - \beta \sum_{k=1}^{C}\left(\sum_{i=1}^{N} u_{ik}\right)^2 + \alpha \sum_{(x_i,x_j)\in M}\sum_{k=1}^{C}\sum_{l\neq k}^{C} u_{ik}u_{jl} + \alpha \sum_{(x_i,x_j)\in C}\sum_{k=1}^{C} u_{ik}u_{jl} \quad (57)$$

$$\text{subject to } \sum_{k=1}^{C} u_{ik} = 1 \text{, for } i \in \{1, 2, \dots, N\}$$

α is the weight given to the supervision. Choosing α is vital for pairwise constraints clustering.

As the number of constraints is remarkably lower than unlabeled data, it is essential to make sure that the constraints have an impact on objective function.

For cluster k ($1 \le k \le C$), μ_k which is the prototype of the cluster, and its cardinality, N_s, are represented by

$$\mu_k = \frac{\sum_{i=1}^{N} (u_{ik})^2 x_i}{\sum_{i=1}^{N} (u_{ik})^2} \tag{58}$$

and

$$N_s = \sum_{i=1}^{N} u_{is} \tag{59}$$

The updated memberships are followed by:

$$u_{rs} = u_{rs}^{FCM} + u_{rs}^{Constraint} + u_{rs}^{Bias} = \frac{\frac{1}{d^2(x_r, \mu_s)}}{\sum_{k=1}^{C} \frac{1}{d^2(x_r, \mu_k)}} + \frac{\alpha}{2d^2(x_r, \mu_s)} \left(\bar{C}_{vr} - C_{vs} \right) + \frac{\beta}{d^2(x_r, \mu_s)} \left(N_s - \bar{N}_r \right) \tag{60}$$

where

$$C_{vs} = \sum_{(x_i, x_j) \in M} \sum_{l=1, l \neq k}^{C} u_{jl} + \sum_{(x_i, x_j) \in C} u_{js} \tag{61}$$

and

$$\bar{C}_{vr} = \frac{\sum_{k=1}^{C} \frac{\sum_{(x_i, x_j) \in M} \sum_{l=1, l \neq k}^{C} u_{jl} + \sum_{(x_i, x_j) \in C} u_{jk}}{d^2(x_r, \mu_k)}}{\sum_{k=1}^{C} \frac{1}{d^2(x_r, \mu_k)}} \tag{62}$$

$$\bar{N}_r = \frac{\sum_{k=1}^{C} \frac{N_k}{d^2(x_r, \mu_k)}}{\sum_{k=1}^{C} \frac{1}{d^2(x_r, \mu_k)}} \tag{63}$$

In [80], the value of α is suggested as:

$$\alpha = \frac{N}{M} \frac{\sum_{k=1}^{C} \sum_{i=1}^{N} \left(u_{ik}\right)^2 d^2\left(x_i, \mu_k\right)^2 x_i}{\sum_{k=1}^{C} \sum_{i=1}^{N} \left(u_{ik}\right)^2} \tag{64}$$

Lastly, β which reflects the importance of competition term is given by:

$$\begin{aligned} \beta(t) = \eta_0 \exp(\frac{-t}{\tau}) \sum_{j=1}^{C} \sum_{i=1}^{N} u_{ij}{}^2 d^2(x_i, \mu_j) + \sum_{j=1}^{C} \sum_{i=1}^{N} u_{ij}{}^2 d^2(x_i, \mu_j) \\ + \alpha \left(\sum_{(x_i, x_j) \in M} \sum_{k=1}^{C} \sum_{l \neq k}^{C} u_{ik} u_{jl} + \sum_{(x_i, x_j) \in C} \sum_{k=1}^{C} u_{ik} u_{jl} \right) \end{aligned} \tag{65}$$

The semi-supervised fuzzy clustering algorithm with pairwise constraints is summarized in below:

ALGORITHM 1 SEMI-SUPERVISED FUZZY CLUSTERING ALGORITHM WITH PAIRWISE CONSTRAINTS

Choose the maximum number of clusters C (number of switches).

Initialize cluster prototypes randomly.

Initialize memberships to u_{ik} = 1/C for all samples (including the extracted feature from the observation under fault operation) and clusters (extracted patterns).

Compute initial cardinality for every cluster using equation (59).

repeat

Compute α and β using equation (64) and (45).

Compute membership's u_{rs} using equation (50).

Compute cardinality for every cluster using equation (59).

until prototypes be stabilized

5.3 Extraction Feature Using Wavelet Multiresolution Analysis

The performance of fault diagnosis method is highly dependent on the extracted features. Different feature extraction techniques have been used in literature. Wavelet analysis, as a prominent tool with time-frequency representation of the signal, provides the ability to localize the non-stationary behavior of signals both in time and frequency domains efficiently. Fundamentally, the wavelet maps a function of the continuous/discrete variable into a sequence of coefficients. Applying the wavelet on functions and operators to transform them into wavelet domain makes them sparse, and this sparseness is a particular feature that results in various applications. Wavelet multiresolution analysis decomposes the original signal x into other signals with the level of resolution i using sequences of high-pass filters and low-pass filters as follows where *i* denotes the level of decomposition. The calculated coefficient from these filters is called approximation and detail coefficients and construct the vector *f*. To deal with high dimensional data issues, calculated approximation and detail coefficients are not used in classification/clustering process directly.

The statistical techniques usually are used to summarize the entire set of wavelet coefficients. In this study, the most famous features in the multiresolution analysis are applied. Mean, standard deviation, root mean square, skewness, kurtosis, form factor, crest factor, energy, Shannon-entropy, and log-energy entropy of this coefficient arrays are brought into play.

Table 4 represents the extracted features in this study, where f_{ij} indicates the i^{th} coefficient from j^{th} level. Variables μ_d, σ_d, and rms_d denote mean, standard deviation, and

root mean square of the extracted wavelet coefficients of the signal. Usually, the coefficients of 3rd level carry enough information for power systems applications.

Table 4 Feature extraction techniques using discrete wavelet transform coefficients.

Feature	**Form**
Mean μ_d,	$x_1 = \frac{1}{N}\sum_{j=1}^{N} f_{ij}$
Standard deviation σ_d^2	$x_2 = \frac{1}{N}\sum_{j=1}^{N}(f_{ij} - \mu_d)^2$
Root mean square rms_d	$x_3 = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(f_{ij})^2}$
Form factor	$x_4 = \frac{\mu_d}{rms_d}$
Crest factor	$x_5 = \frac{peak}{rms_d}$
Energy	$x_6 = \sum_{j=1}^{N} \lvert f_{ij} \rvert^2$
Log Shannon entropy	$x_7 = \sum_{j=1}^{N} \log(f_{ij}^{\,2})$
Shannon entropy	$x_8 = -\sum_{j=1}^{N} f_{ij}^{\,2} \log(f_{ij}^{\,2})$
Skewness	$x_9 = \frac{1}{\sqrt{6N}}\sum_{j=1}^{N}\left(\frac{(f_{ij} - \mu_d)}{\sigma_d}\right)^3$
Kurtosis	$x_{10} = \frac{1}{\sqrt{N}}.\sum_{j=1}^{N}\left(\left(\frac{f_{ij} - \mu_d}{\sigma_d}\right)^4 - 3\right)$

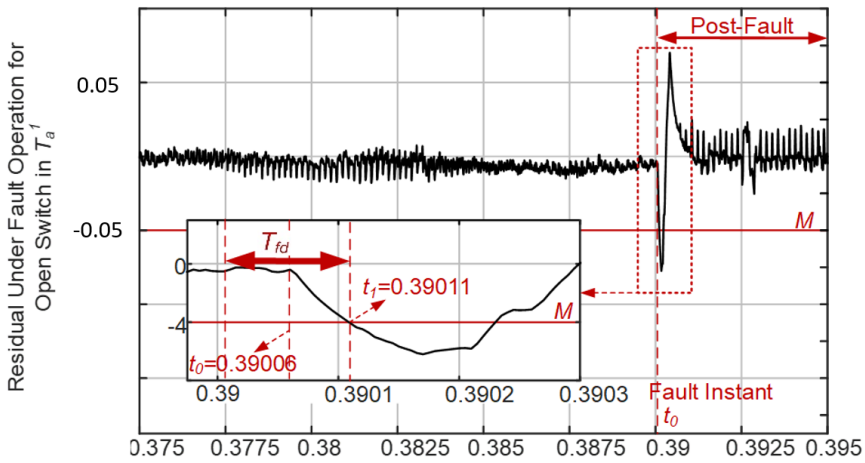

Figure 27 Data window for feature extraction.

If T_{fd} denotes the time interval between fault occurrence and fault detection, features are extracted from a window of data with the length of $2\times T_{fd}$. Let T_d be the moment that fault occurrence is reported, which is the moment that residual exceeds the threshold. There are two option for obtaining $T_{fd;}$ first, assuming a constant $T_{fd,}$ independent from the performance of fault detector, second, defining a degradation region.

It was mentioned before that the average of residual is zero in normal operation. As a result of failure, this average diverges from zero. By calculating the moving average of residual, start point of degradation region can be estimated. In this study, a fix window is considered. Figure 27 indicates a data window example for feature extraction and T_{fd} calculation. This window size guarantees that the data window includes the information of pre-fault occurrence time instant, fault occurrence time data, and post-fault occurrence time data. In this manner, unique fault characteristics of each fault operation condition are considered.

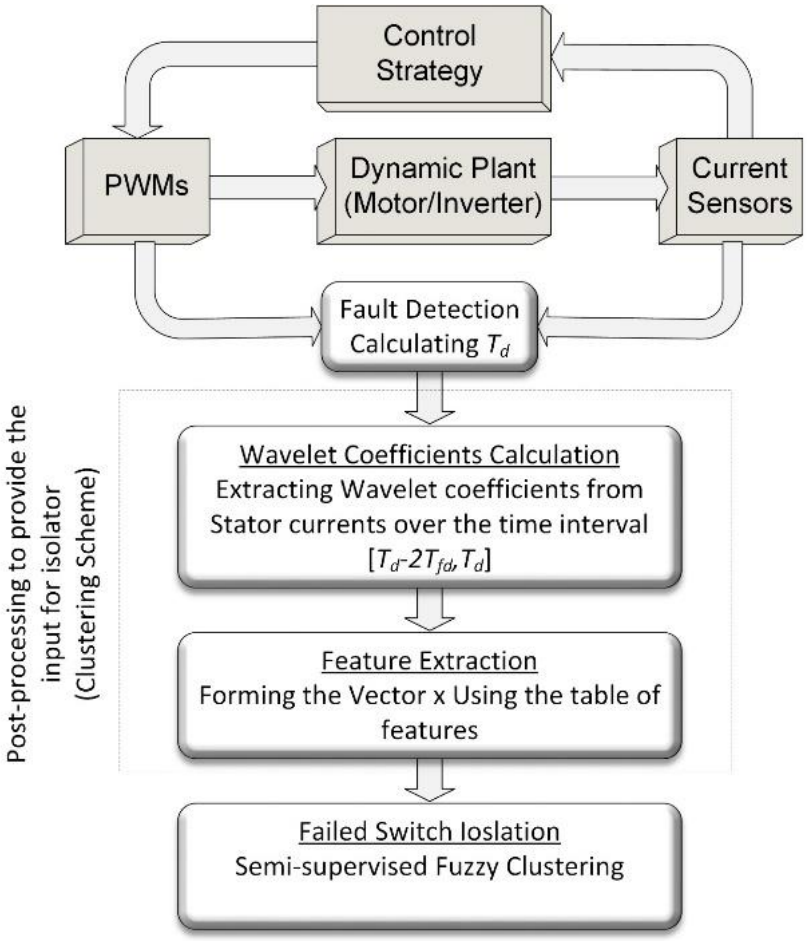

Figure 28 Steps to provide the input for clustering model.

Eventually, the extracted feature vector is given to the model which is adding extracted patterns from the experiment, as constraints, to unsupervised clustering model, based on the extracted patterns from simulation, to enhance the performance of fault isolator. Figure 28 summarizes the steps for providing the input of clustering model. Also, the overall block diagram of the FDI scheme for a multilevel multiphase inverter is indicated in Figure 29.

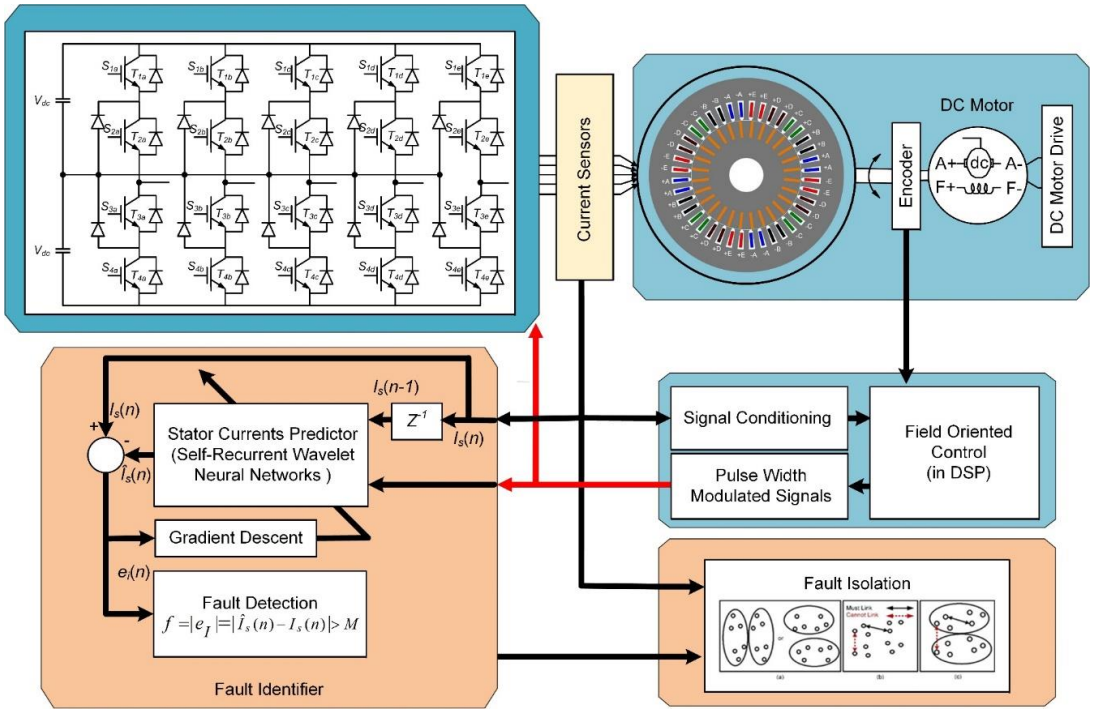

Figure 29 The block diagram of the proposed fault detection and isolation scheme.

5.4 Experimental Results

The configuration of the lab prototype, which has been developed to evaluate the proposed fault detection and isolation scheme, is shown in Figure 30. The experimental testbed includes a five-phase induction motor coupled to a dc motor is used to load profile change. Figure 30 illustrates the three-level NPC used in this experiment. Two three-phase three-level NPC inverters are used as a five-phase three-level inverter. The indirect field-oriented control (IFOC) strategy with sinusoidal pulse-width modulation is implemented on a Texas Instruments TMS320F28377 Delfino microcontroller as well.

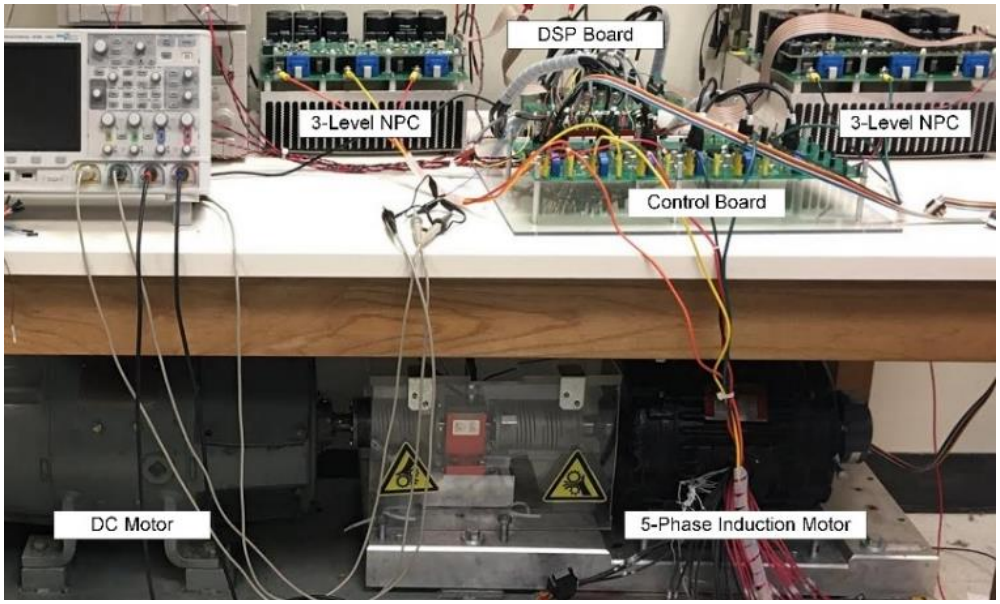

Figure 30 Experimental setup for five-phase three-level NPC.

The proposed approach is validated in controlled current and speed operation. The speed of the motor is desired to remain as 0.3pu, which is a low-speed condition. Due to symmetric system, it is sufficient to investigate the fault scenarios in one phase leg. For that reason, considered fault scenarios are open switch failure in $T_{1a,}$ $T_{2a,}$ T_{3a} and $T_{4a.}$ The supremum of error in normal operation is about 0.05, and thus the threshold is selected as 0.06 to avoid false alarms. Figure 31 and Figure 32 demonstrate actual current and generated fault residual for fault incidence in the switch T_{1a} and $T_{2a.}$ The proposed approach can detect the failure in the upper pair switches T_{1a} and T_{2a} in $T_{fd} \leq 500(\mu s)$, where the threshold is defined as M=0.06A. Likewise, the results of the fault detection scheme for the open switch failure in the lower pair is shown in Figure 33 and Figure 34, respectively. Also the actual current and generated fault residual for fault incidence in the

switch T_{3a} and T_{4a} are shown in Figure 33 and Figure 34.

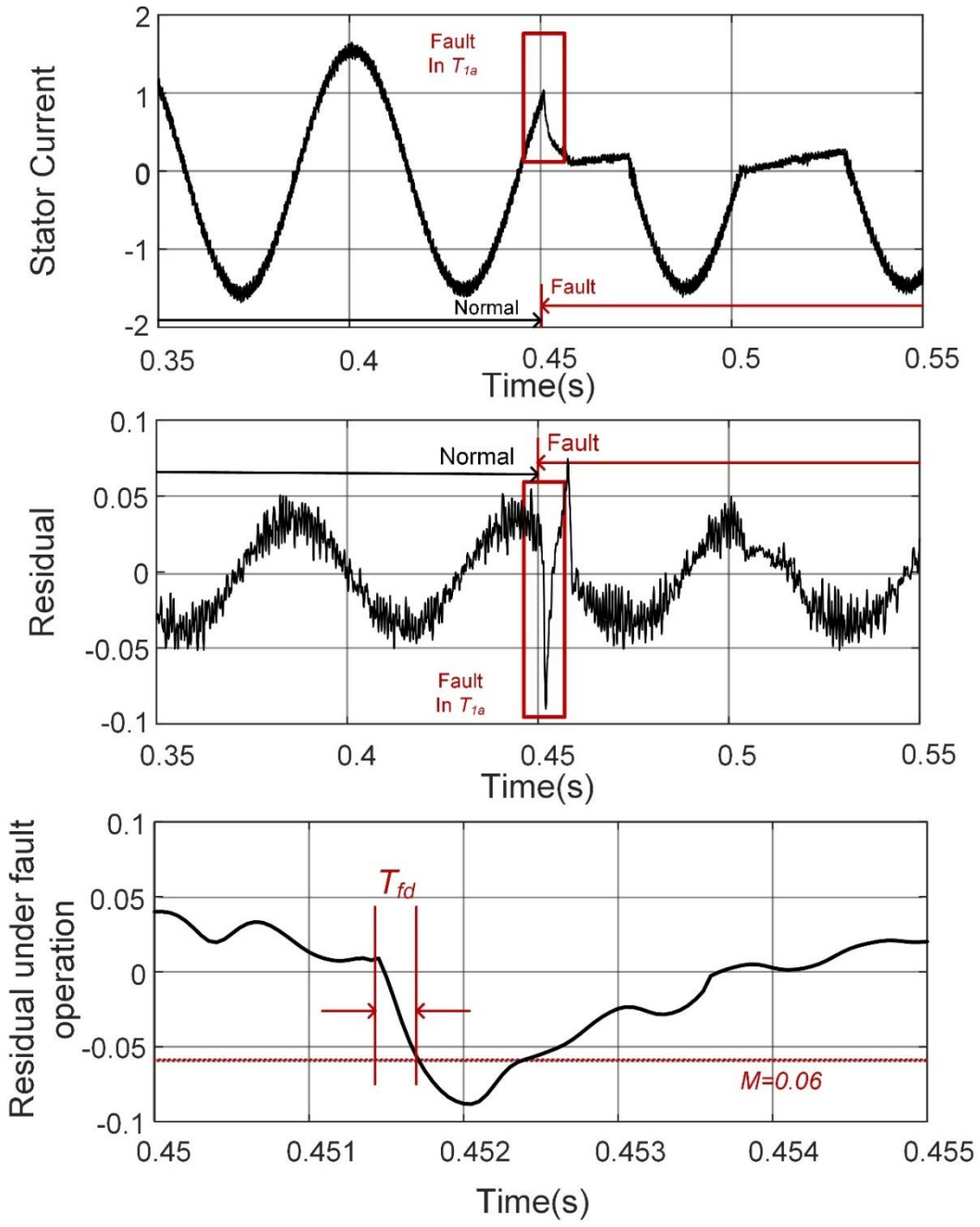

Figure 31 The actual current of phase a and the residual for fault in the switch T_{1a}.

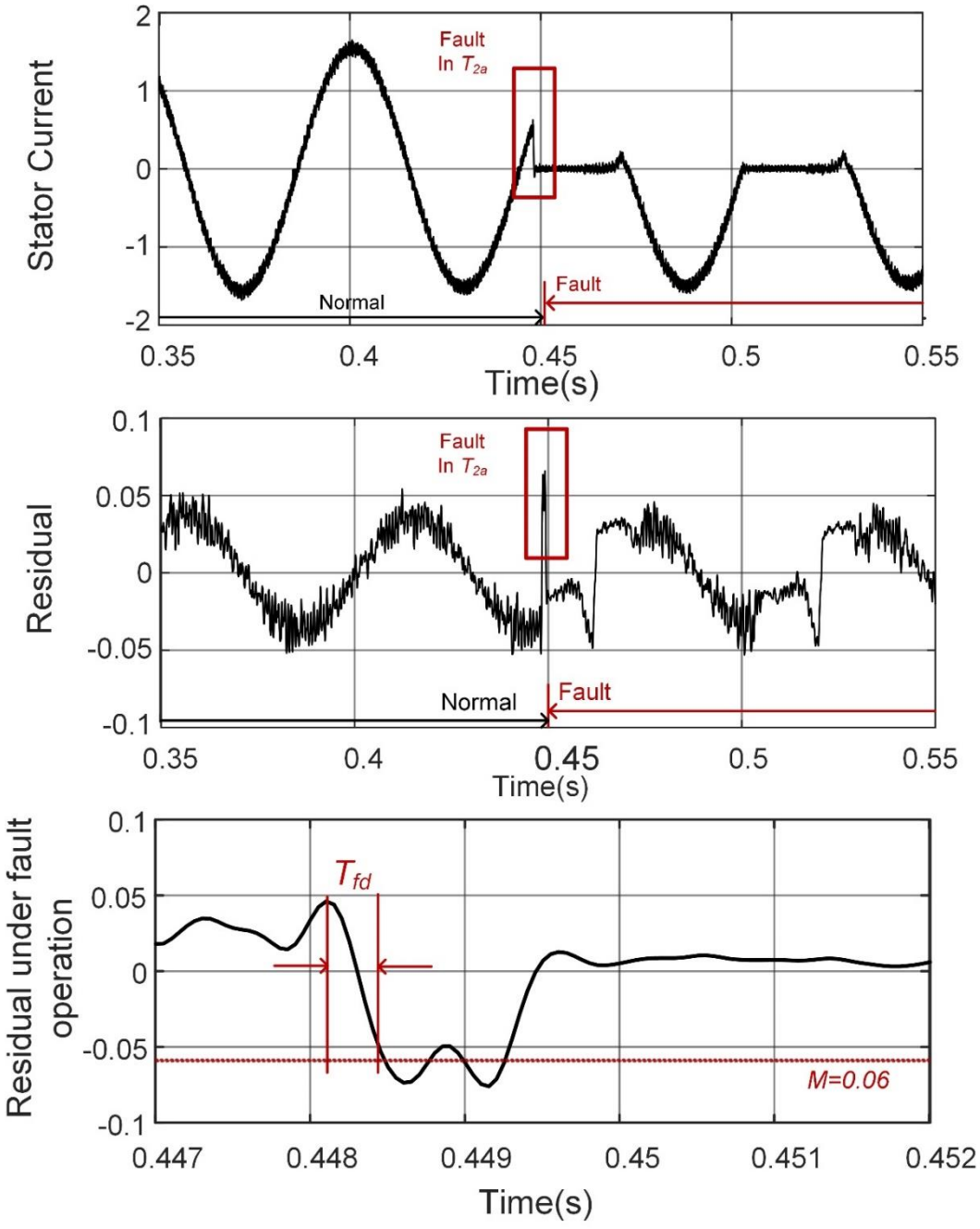

Figure 32 The actual current of phase a and the residual for fault in the switch T_{2a}.

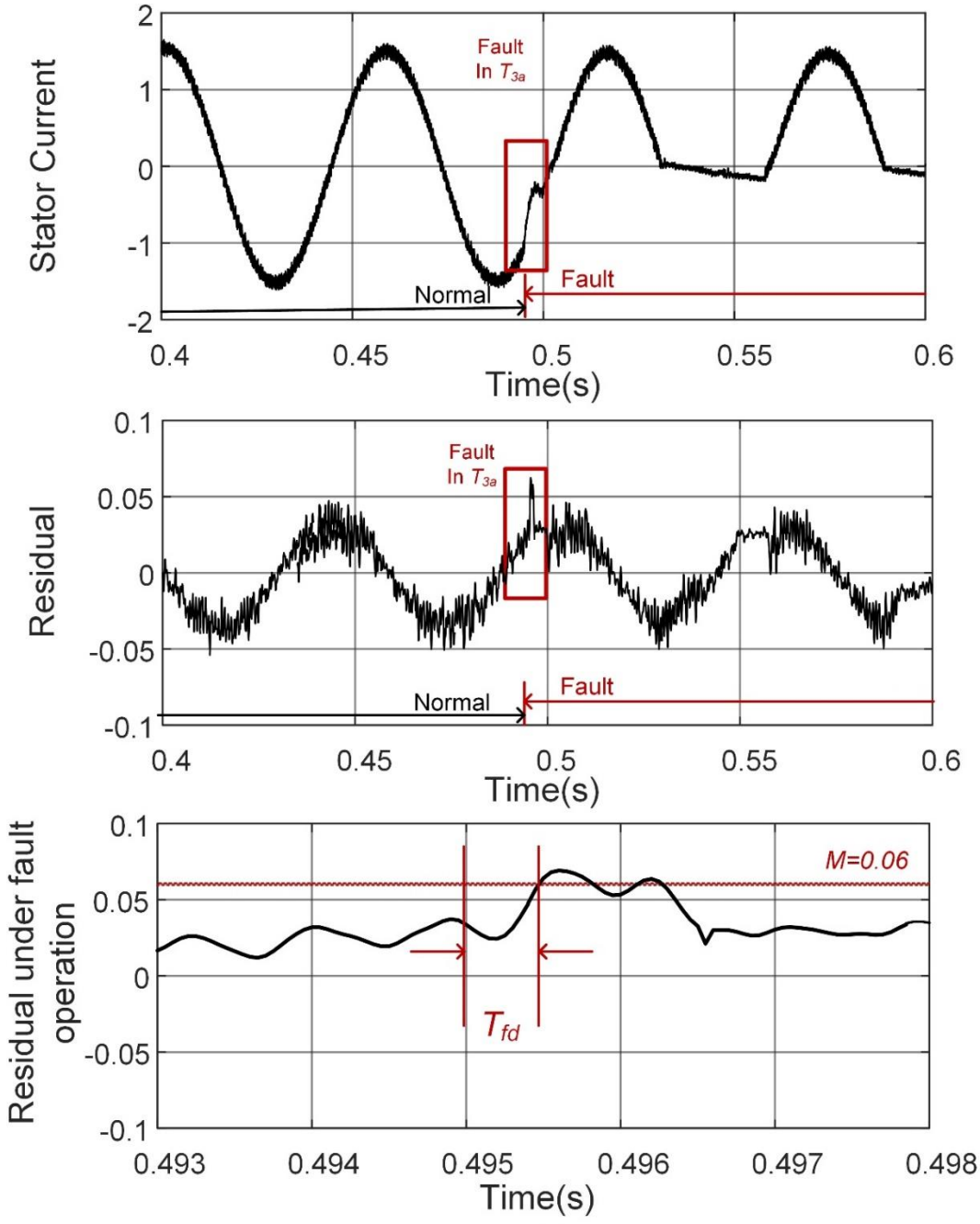

Figure 33 The actual current of phase a and the residual for fault in the switch T_{3a}.

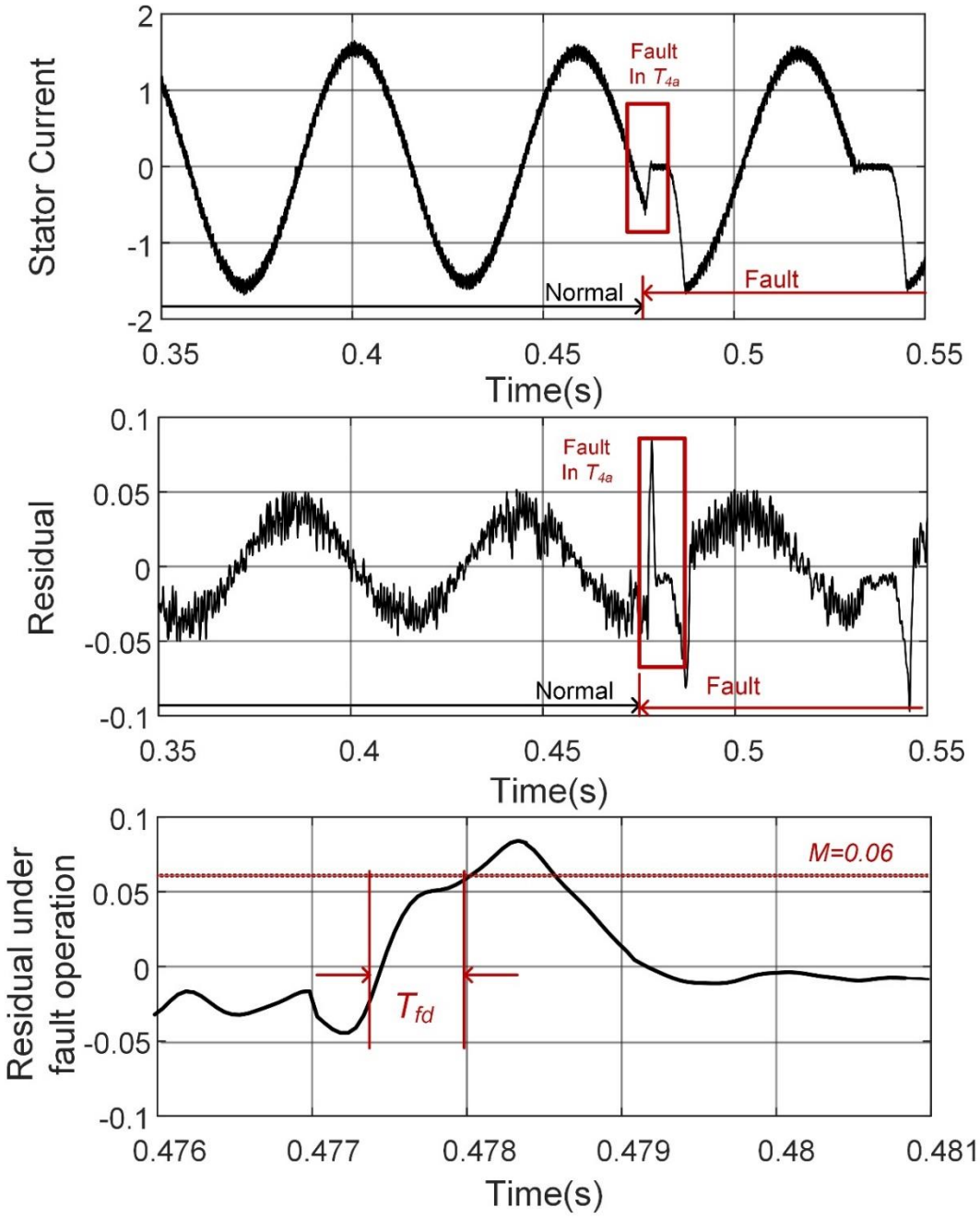

Figure 34 The actual current of phase a and the residual for fault in the switch T_{4a}.

Afterwards, features are extracted from the window of data containing transient individualities of the fault operation condition. Then the extracted feature vector from a sliding window of data with the length of $2\times T_{fd}$ is given to the clustering model. This isolation time is about 1% of the fundamental period of stator current.

The semi-supervised fuzzy clustering method assigns the fault condition feature vector to the corresponding switches correctly. The clustering algorithm is calculating a membership operation function for each cluster. Eventually, the maximum membership values indicated in Table 5 and Table 6 show the cluster corresponding to the failed switch.

Table 5 Clustering result for fault in upper part switches.

Isolated Switches	Membership Value	
	Fault in T_{1a}	Fault in T_{2a}
T_{1a}	**0.7449**	0.0227
T_{2a}	0.0390	**0.6532**
T_{1b}	0.0371	0.0330
T_{2b}	0.0407	0.0331
T_{1c}	0.0433	0.0530
T_{2c}	0.0158	0.0515
T_{1d}	0.0256	0.0418
T_{2d}	0.0376	0.0245
T_{1e}	0.0092	0.0526
T_{2e}	0.0067	0.0344
T_{1a}	0.0002	0.0239
T_{2a}	0.0002	0. 0001
T_{1b}	0.0004	0. 0001
T_{2b}	0.0002	0. 0000
T_{1c}	0.0002	0. 0001
T_{2c}	0.0000	0. 0000
T_{1d}	0.0000	0. 0000
T_{2d}	0.0001	0. 0001
T_{1e}	0.0003	0. 0001
T_{2e}	0.0003	0. 0001

Table 6 Clustering result for fault in lower part switches.

Isolated Switches	Membership Value	
	Fault in T_{3a}	Fault in T_{4a}
T_{3a}	**0.5380**	0.0891
T_{4a}	0.0309	**0.5660**
T_{3b}	0.0159	0.1022
T_{4b}	0.0237	0.1037
T_{3c}	0.0309	0.0624
T_{4c}	0.0537	0.0064
T_{3d}	0.0064	0.0255
T_{4d}	0.1160	0.0383
T_{3e}	0.1215	0.0017
T_{4e}	0.0631	0.0047
T_{3a}	0.1117	0.1775
T_{4a}	0.0000	0.00001
T_{3b}	0.0001	0.00001
T_{4b}	0. 0001	0. 0000
T_{3c}	0. 0001	0. 0001
T_{4c}	0. 0001	0. 0002
T_{3d}	0. 0001	0. 0001
T_{4d}	0. 0000	0. 0000
T_{3e}	0. 0001	0. 0002
T_{4e}	0.0000	0. 0000

5.5 Chapter Summary

In this chapter, a novel hybrid fault diagnosis framework is developed for multiphase multilevel inverters. An adaptive nonlinear model predictor, SRWNN, estimates the stator currents using a combination of switching patterns and former states of currents. As the results show, the modeling method handles the nonlinearities of a multilevel inverter and approximates the currents with less than 1% error. Thus, model-based fault detection method detects the abnormality and failure excellently. Then, a partial knowledge-based based technique, using the concept of semi-supervised learning, assigns the transient current response under fault operation to one of the clusters. Due to the limitation in training process and obtaining labeled data, semi-supervised paradigm combines a limited number of labeled data, gained from the experiment, with a more extensive set of labeled data points, acquired from the simulation. Experimental results for a three-level five-phase NPC converter illustrate that the methodology is fast and reliable. The proposed FDI framework offers the ability to detect and isolate the failure in about less than 500 microseconds based on 1% of the fundamental period of the stator current, while existing system-based fault diagnosis methods require at least one cycle of stator current, while the presented methodology reduces it to 1% of the fundamental period. Moreover, condition monitoring, decision making, and diagnosis are accomplished based on measured stator currents and PWM gate signals. Therefore, no extra hardware is required. Real-time fault characteristic identification does not require large training data collected under fault operation conditions. Fault diagnosis (detection, identification, and isolation) is performed under controlled conditions.

6. REAL-TIME FAULT DETECTION AND ISOLATION STRUCTURE FOR MULTILEVEL NEUTRAL POINT CLAMPED INVERTERS USING ANALYTICAL DATA MINING

Managing the information and selecting the informative data is the most effective element to develop an economical and efficient performance industrial process. Due to the most dominant requirements of advanced intelligent systems, speed and practicable data dimension, the power of computing is not the highest priority anymore. Computational capabilities have been substituted by the quality of data and data management. Comprehension, context, and connection of information are the foundation of high-quality data. As a result, technologies need to be fast, simple, powerful, and intelligent.

Machine learning, which is advanced from the concepts of pattern recognition and computational learning theory in AI, is the process of exploring the information and constructing the algorithms that can learn from and make predictions based on the data. Statistical computations as the central core of machine learning are mostly behaving with the systems under consideration as a black box. However, most modern machine learning algorithms are vulnerable to minimal and almost imperceptible perturbations of their inputs. Therefore, gaining an insight into inner layers of learning has attracted attention. Also, the interpretability of the solutions is highly challenging.

It was discussed in the previous sections that the adaptive self-recurrent wavelet neural network identifier estimates the nonlinear model to predict the phase currents and generates a robust residual. It is followed by the residual evaluation stage, which is a logic

decision-making process to transform quantitative knowledge into qualitative, e.g., yes/no statements. The residual evaluation can be seen as a classification/clustering problem.

A machine learning clustering and pattern recognition paradigm was developed to isolate the failed switch. Calculated coefficients, using wavelet multiresolution concept, were given to statistical feature extraction procedure. Ten different statistical feature was utilized in this study, which results in $n \times 10$ feature vector in an *n*-phase system. Although a dimensionality reduction process was employed to reduce the size of data set points in clustering model, the quality of data is still concerned. Due to the importance of the effectiveness of data and computational complexity, the purpose of this chapter is investigation and simplification of inner layers of machine learning paradigm.

As the first step, it was desired to investigate the effect of features selection methods, which reduce the dimension of data set but preserve the characteristics of features. The goal of conducting feature selection is to identify the most critical and relevant subset of existing features for classification/clustering without a transformation. Also, it was preferred to preserve the most useful features to develop the physical validation in the real system and check the feasibility. Feature selection (feature subset selection) requires a search strategy to select the candidate subsets and an objective function to evaluate selected candidates. Search strategy, as an exhaustive evaluation of feature subsets, involves combinations of features. The number of selected featured could be optimized, but to avoid extra complexity, it is selected as 2 in this study. The search strategy explores the space of entire possible combination of features. Eventually, the objective function evaluates candidate subsets with their information content and

calculates a degree of goodness according to the criterion. It should be noted that evaluation is different from classification process.

In this regard, feature selection was performed using a sequential forward selection approach. Following Figure 35, shows the result for the performance of SFS for 190 combinations of two features from a set of ten ($C_2^{20} = 190$). As it is illustrated, the first 45 values are related to the cases that two features are selected from extracted features from stator current subset x^I. The next 45 values represent the performance of the cases that two features are selected from the pulse width modulated switching signals (PWMs) subset x^{PWM}. The last 100 values are for the combination of one feature from stator current subset x^I and one feature from stator current subset x^{PWM}.

Therefore, the combination of the 9th feature of x^I and 6th feature of x^{PWM} is selected as the most informative features. Based on the Table 4 the selected features are skewness of stator current and wavelet energy of PWM switching signals respectively.

It should be noticed that these features are for each phase and the total number of features, given to clustering or classification model, will be multiplied by *n*, the number of phases. Along these lines, the dimension of the data set for an *n*-phase machine is reduced from $(10 \times n) + (number\ of\ independent\ switches \times 10 \times n)$ to $(1 \times n) + (number\ of\ independent\ switches \times 1 \times n)$. The number of independent switches for three-level NPC inverters is 2.

T o have a better understanding of the skewness concept, it should be remarked that skewness is higher-order statistical attributes of a series in time, frequency, or time-frequency domain. It shows the symmetry of the probability density function of the

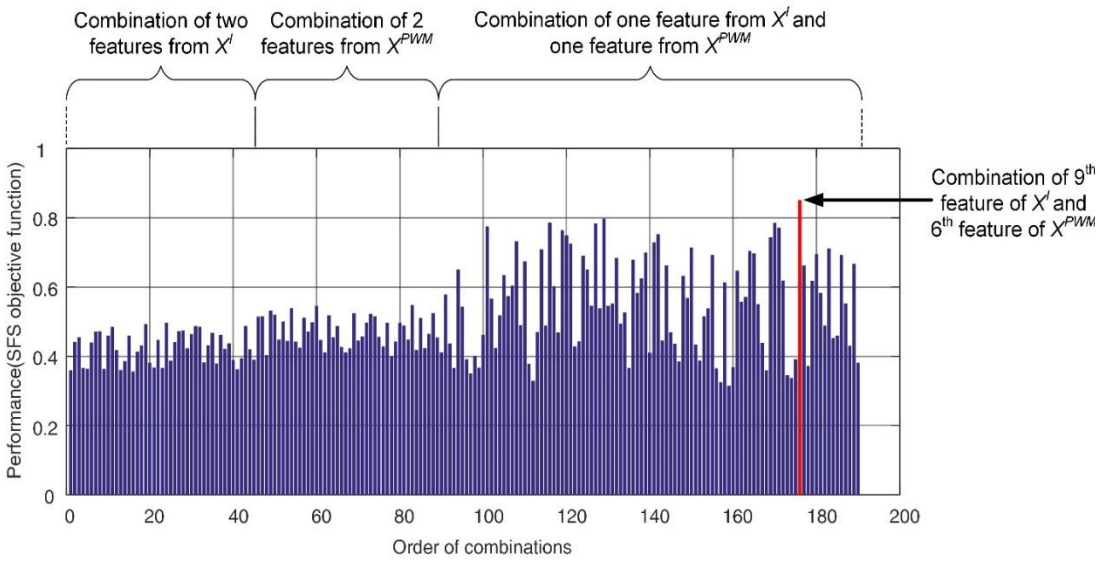

Figure 35 Selecting the features using sequential forward features.

amplitude of the series. Positive or negative skewness represent the systematic changes in the associated physical systems. The Strong influence of stator current skewness leaded us to investigate current symmetrical properties more precisely. Considering the online calculation comlexity and regarding the goal of focusing on symmetrcal behaviour in transient, the slop of the stator currents are scanned and explored. Moreover, Wavelet Entropy (WE) is a metric to quantify the energy distribution in wavelet sub-bands using combining the wavelet decomposition and entropy. In fact, it indicates the degree of disorder in the transient signals with an excellent time-frequency resolution. Here and now, due to the prior knowledge on the effect of selected features, the post-fault operation behavior of PWM switching signals and stator current signals are examined.

6.1 Failed Phase Identification Using Multi-Criteria Wavelet Based Method

Due to the defined objectives, drive systems under closed-loop controlled operation are usually concerned with controlling the voltage, current, flux, torque, or speed. The primary challenge is that there is a conflict of interest between designing a model-based fault diagnosis scheme and a closed-loop control strategy. State of the art for control is to minimize and compensate the deviations of targeted signals from desired behavior, and the effect of failures regard to objectives. While the aim of supervisory function, which includes monitoring, automatic protection, and fault diagnosis is to detect these deviations with the purpose of producing the proper fault indicators to make a decision and take appropriate action as the reconfiguration procedure. It was observed that the controller starts to compensate the influence of switch failures to minimize their consequences instantly. It tries to keep the open-circuit switch closed, however, in fact, it is open continuously. Therefore, as a result of controller's effort, gate signals and switching states patterns change significantly. When the switch is failed (open-circuited), as a result of controller's reaction, the related PWM remains at the highest allowed values (+1) to force the switch to stay closed. PWM signals are high frequency modulated signal which evaluating them to deciding on the plant is challenging. Due to the environmental noise and measurement error, it is essential to select the most informative feature. Under the failure operation, the wavelet energy of modulated gate signal increases remarkably. To investigate the gate signals behaviors under various fault conditions, the wavelet analysis of PWMs related to different switches have been considered. Concerning high switching frequency, it is not practical to apply PWM values directly. Thus, wavelet analysis is

applied to process the energy spectrum of gate signals in different fault scenarios. Wavelet analysis, as a precise tool to present signals, is able to localize the non-stationary behavior of signals both in time and frequency domains efficiently [32]. Several features have been examined, and as a final point, the Wavelet Entropy is chosen which can represent the highest amount of information corresponding to the fault cases. If f^{ij} denote the i^{th} element of the j^{th} level of wavelet decomposition array corresponding to a signal, then the normalized wavelet energy is defined as:

$$WE = \sum_{j=1}^{N} \left| f_{ij} \right|^2 / N \tag{66}$$

where N is the number of samples over a moving window.

However, there is a challenge to discriminate between two switches in the same pair. In Neutral Point Clamped inverters it is not only the gate signals of failed switch that is affected by a fault, but the other witch in the same pair is influenced too. Technically speaking, the controllers try to keep the current error minimized. However, the inner switches in each leg are conducting in the longer period rather than outer switches. This longer time of conduction results in having higher wavelet energy and wavelet Shannon entropy.

Needless to say that in 3-level NPC inverter complementary switching rule is performed for the pairs of (T_{x1}, T_{x3}) and pairs (T_{x2} and T_{x4}).Therefore WE_{2x} indicates the same properties as WE_{4x}, which gives information about the failure in lower part failures indirectly. Taking into account all these factors, the switches gate signals contain discriminable content.

Combining this property with the sign of residual, results in proposing the following fault phase identification criteria:

$$\begin{gathered}\left(\left(sign(R_x)<0\right)\&\left(WE_{1x}=1\right)\right)\\ or\\ \left(\left(ssign(R_x)>0\right)\&\left(WE_{2x}=0\right)\right)\\ x\in\{a,b,c,d,e,\ldots\}\end{gathered} \tag{67}$$

6.2 Failed Switch Isolation Using Analytical Pattern Recognition Method

The next step after identifying failed phase is isolating the failed switch. Considering the physical interpretation of skewness and symmetrical trend in stator currents, the slope of stator current in transient and under post-fault operation were brought into play. From circuit theory perspective, when the open-circuit fault occurs in one of the switches, the current of related phase goes to zero. However, the decay time depends on the location of the failed switch.

Consider the stator current of faulted phase x, using Kirchhoff's law, as below:

$$L_T\dot{I}_{xs}(t)=-R_T I_{xs}(t)+\Delta V_{xz} \tag{68}$$

This differential equation is solved as

$$I_{xs}(t_1)=e^{-(t_1-t_0)/\tau}I_{xs}(t_0)+\Delta V_{xz}\left(1-e^{-(t_1-t_0)/\tau}\right)/R_T \tag{69}$$

where t_0 and t_1 refer to the time instants when the fault occurred and the time instant when at least one of the residuals exceeds the threshold, respectively. τ denotes the time constant as $\tau=L_T/R_T$ for induction machine equivalent circuit parameters L_T and R_T .

By applying Taylor series expansion, the current can be rewritten as:

$$I_{xs}(t_1) = I_{xs}(t_0)\left(1-(t_1-t_0)/\tau\right) - \Delta V_{az}\left(1-\left(1-(t_1-t_0)/\tau\right)\right) \quad (70)$$

Assuming that the $(t_1 - t_0) \ll \tau$, then

$$I_{xs}(t_1) = I_{xs}(t_0)\left(1-\frac{(t_1-t_0)}{\tau}\right) + \frac{\Delta V_{xz}}{R_T}\left(1-\left(1-\frac{(t_1-t_0)}{\tau}\right)\right) \quad (71)$$

It should be noted that V_{xz} and $(t_1 - t_0)$ vary in different fault scenarios because the current path through the faulty leg depends on the location of failed switches. Analyzing the post-fault current behavior illustrates that the rate of changing the current represents which switch is failed. Equation (71) yields in obtaining the slope *m* as

$$m = \frac{\Delta V_{xz}}{L_T} = \left(I_{xs}(t_1) - I_{xs}(t_0)\right)/(t_1 - t_0) \quad (72)$$

m is a slope of staor current. The value of ΔV_{xz} is depended on the fault location. It can be *E*, 2*E*, -2*E*, or -*E*. The sign of ΔV_{xz} demonstrates that the failure is located in the upper half leg or lower half leg. A Positive value is corresponding to lower half leg and a negative value is corresponding to the upper half leg. Furthermore, if $|m \times LT|$ be equal to *E*, the outer switch is failed, while if it be equal to 2*E*, the fault is located in the inner switch.It should be noted that t_1 is not known exactly, but it is possible to consider a fixed time as T and replace the $(t_1 - t_0)$with t_d.

The general fault diagnosis scheme is shown in Figure 36. Self-recurrent wavelet neuro-fuzzy network estimates the five-phase stator currents.

Next, the fault residual is calculated using equation 7. The formed residual carries the information about the current, switching pattern, and deviation values from normal operation. When at least one of the residuals exceeds the pre-defined threshold, fault

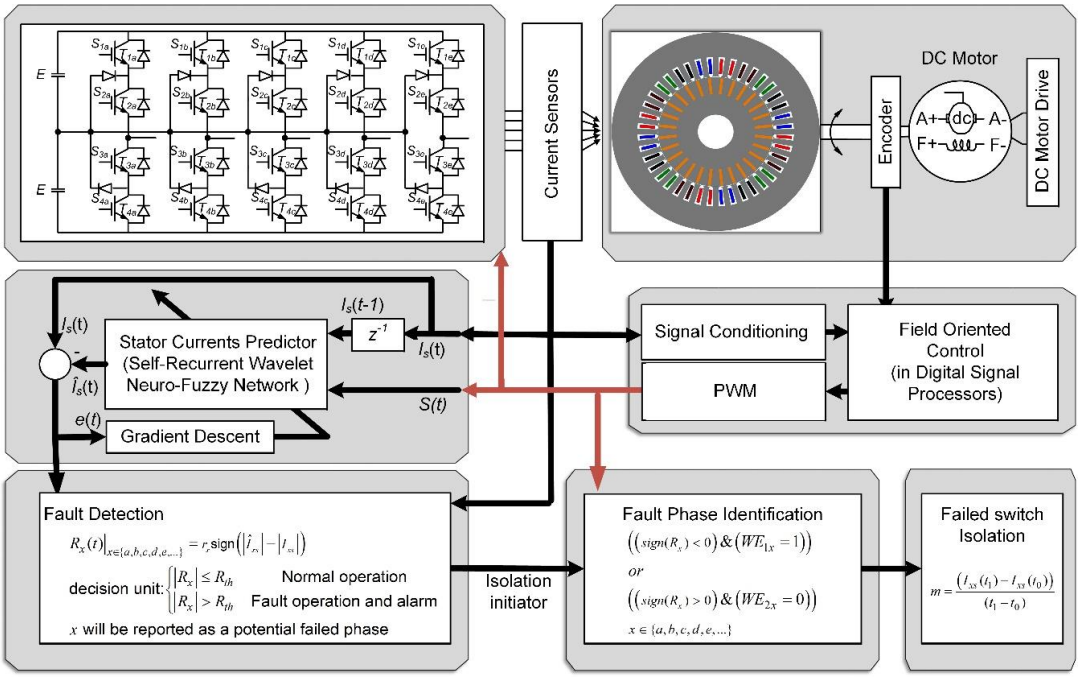

Figure 36 Overal fault diagnosis scheme.

isolation, including the faulted phase identification and failed switch isolation, will be initiated.

6.3 Experimental Results

A lab prototype is designed and implemented to verify the performance of the fault diagnosis scheme, including fault occurrence detection, faulted phase identification, and failed switch isolation. As the setup configuration indicates in Figure 37, the testbed consists of a five-phase induction motor fed by a five-phase three-level NPC. The motor is coupled to a dc motor to load a profile change. Two three-phase three-level NPC

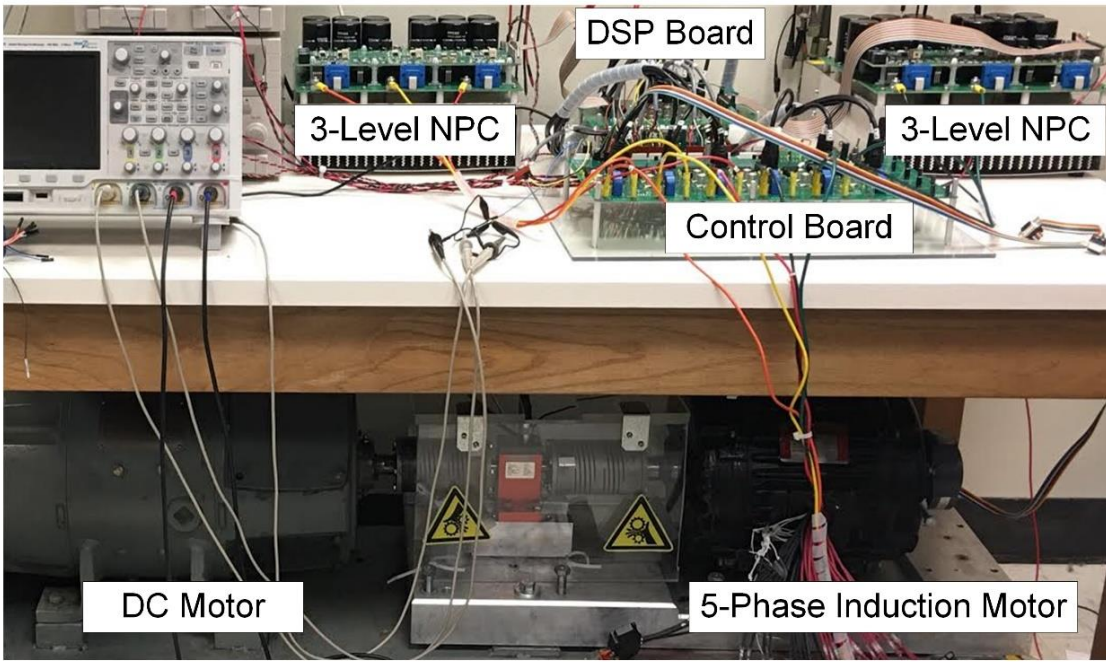

Figure 37 Experimental setup.

inverters (SEMIKRON 3L-SKiiP 28 MLI 07E3V1 Evaluation Inverter) are employed in parallel to build a five-phase inverter. The proposed approach is validated under controlled current and speed operation. The IFOC strategy with SPWM is implemented on a Texas Instruments TMS320F28377 Delfino microcontroller.

The maximum of error in normal operation is about 0.05, and thus the threshold is selected as 0.06 to consider the uncertainties and disturbances and to avoid false alarms. T_d is considered as 500μs. Due to symmetry of the system, it is sufficient to investigate the possible fault scenarios in one phase. Moreover, top and bottom pairs of switches and are symmetric. For this reason, only open switch failure scenarios in T_{1a} and T_{3a} are considered. The investigated fault scenarios, corresponding to T_{1a} and T_{3a}, are presented

in Figure 38 and Figure 39, respectively. Figure 38 shows that switch T_{1a} is failed at t_0=0.4s (the gate signal given to switch is constant zero). Then, the residuals corresponding to two phases, particularly, phase *a* and *e*, are reported as potential faulted phases, because they exceed the limit of 0.06. Although phase *e* plays a role in the fault detection step and initiates the isolator, it does not fulfill (67), the phase identification condition. The residual for phase *e* is negative, but the wavelet energy of inner top switch goes to zero. It means that there is an over positive current in phase *e* and the current is increasing, and thus the controller is compensating it with an updated switching pattern to keep the top switches off and the bottom switches on. On the other hand, in phase *a*, the residual is negative, which means that the current is reduced. Besides, under the effect of current controllers, the generated PWM signals of phase *a* are aimed to turn on the top switches on and bottom switches off. It yields to an increase in the wavelet energy in the top switch of phase *a*. Therefore, phase *a* is presented as the faulted phase and the failure is located in the top switches.

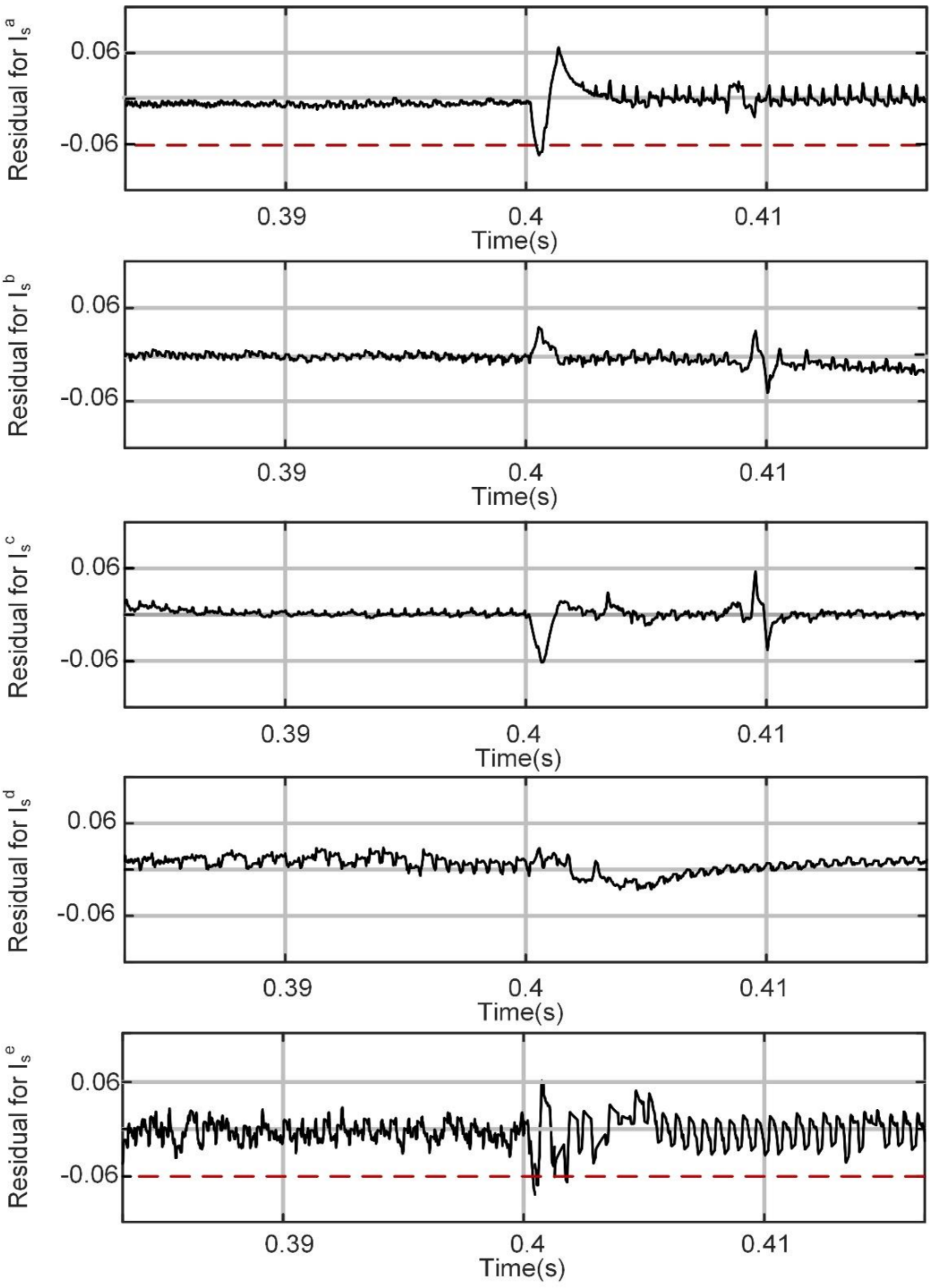

Figure 38 Five phase residuals for T_{1a} open switch.

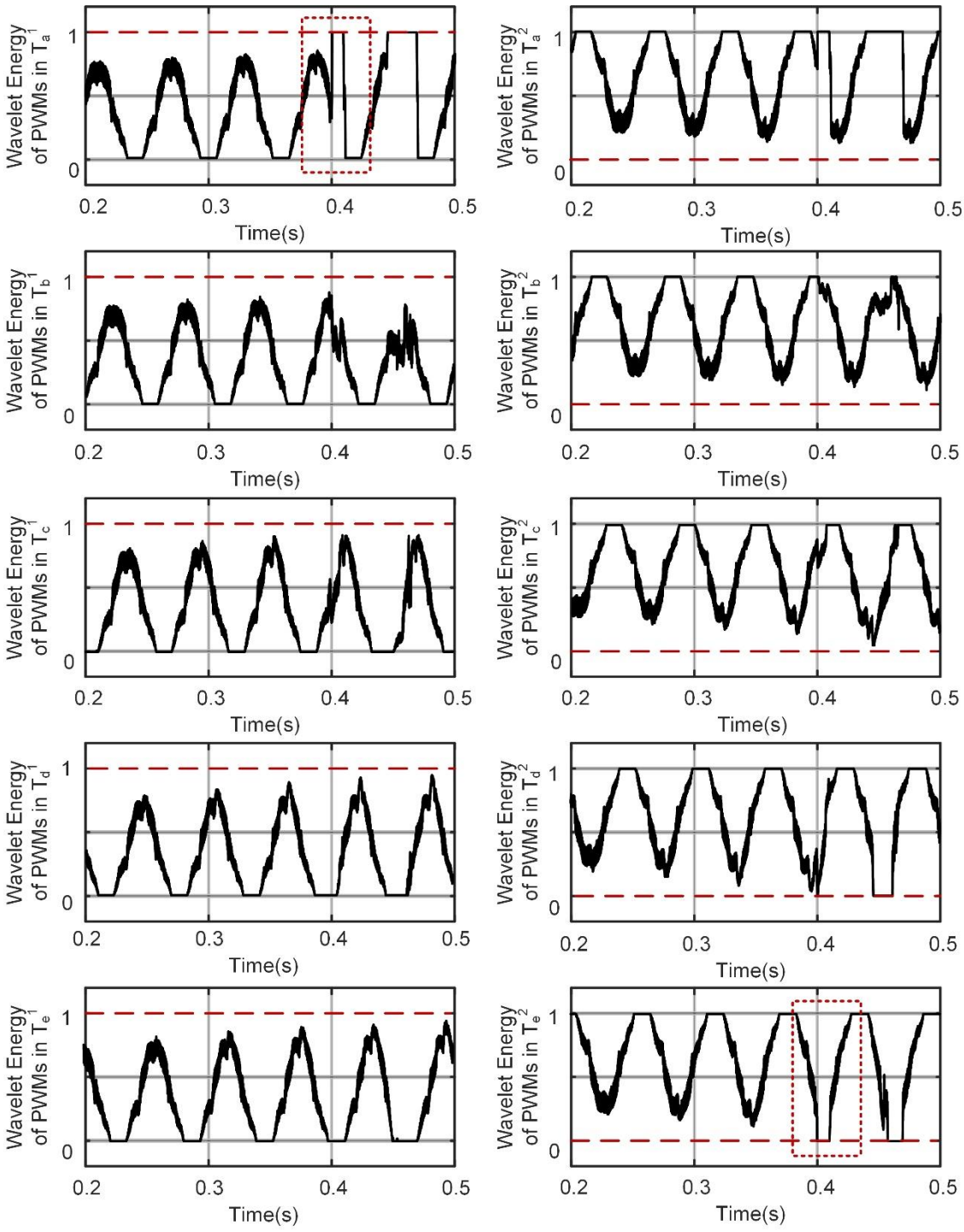

Figure 39 Wavelet energy of PWMs corresponding to top pair switches in five phase for T_{1a} open switch.

Similarly, Figure 40 and Figure 41 the failure in switch T_{3a} at t_0=0.42s. The corresponding residuals to phase *a*, and *e* exceed the threshold. Again, the terms of phase identification condition are not satisfied for phase *e.* The residual of phase *d* is positive representing that top switches are failed, but the wavelet energy of top switch goes to one, which means that that the failure is in the bottom switches. Due to this contrast, phase *e* could not be failed. However, in phase *a*, the residual is positive and energy of the second switch drops to zero, both of which mean the failure is located in the bottom switches. Therefore, bottom switches of phase *a* are reported as failed in this case.

The next level of the hierarchical fault isolation scheme is locating the failed switch. Figure 42 illustrates the wavelet energy of T_{a1} and stator current of phase *a,* which is the phase known to be faulted in this case. The fault happens at *t*=0.4s. The wavelet energy of the PWM switching signal corresponding to T_{a1} goes to +1 at *t*=0.4005s [see Figure 42(a)]. Referring to this time interval, a linear curve, d_1, is fitted to stator current and is shown in Figure 42(b). The calculated slope of this curve is -250. Based on equivalent circuit, L_T is $0.1297H$, which results in $\Delta V_{az}=|m \times L_T|$ =32.42v. ΔV_{az} is close to *E*, which implies that the failure occurs in the outer switches, and negative slope shows that the failure is at the top pair. Considering the aforementioned, the top switch, T_{a1}, is isolated as failed switch.

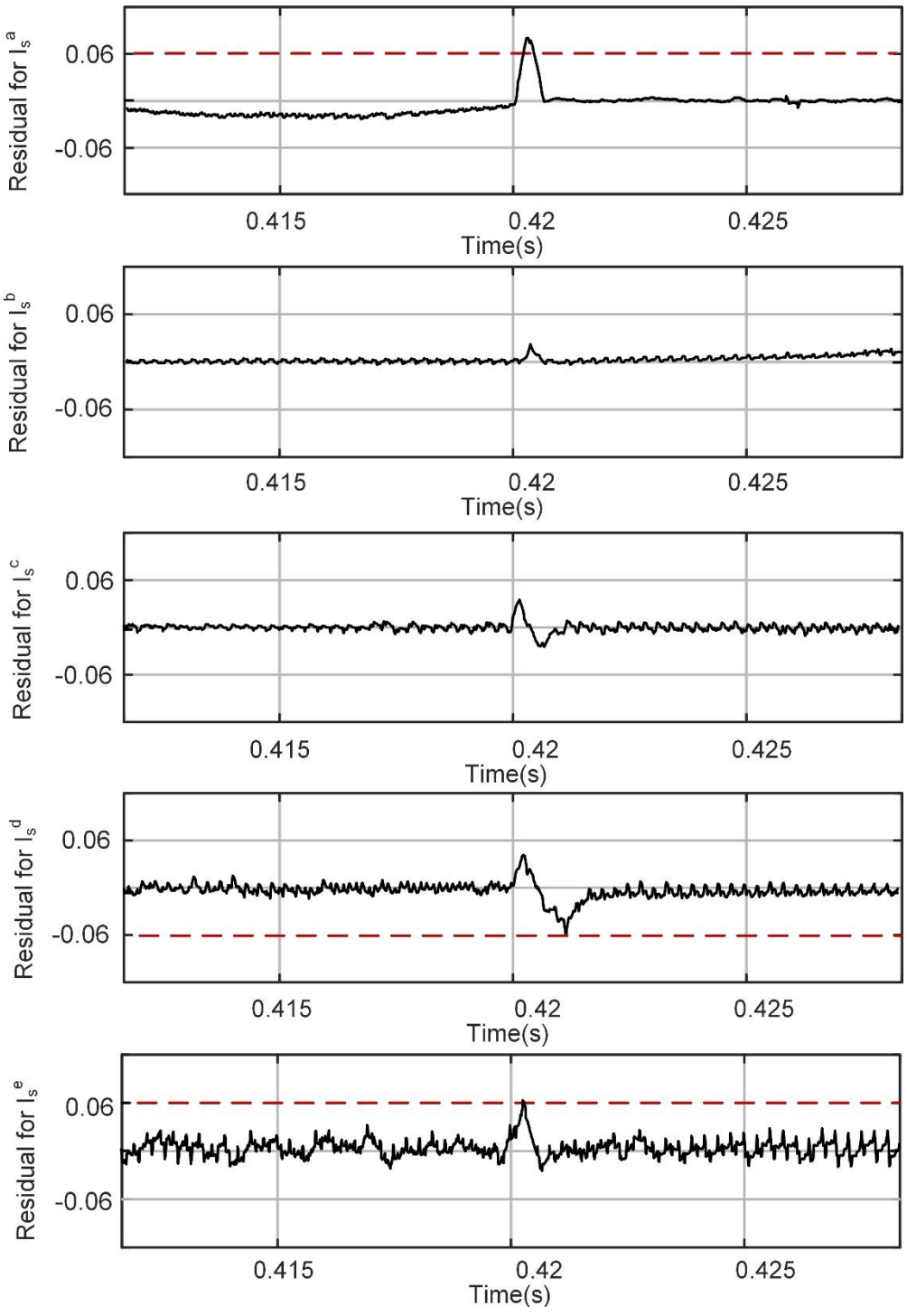

Figure 40 Five phase residuals for T_{2a} open switch.

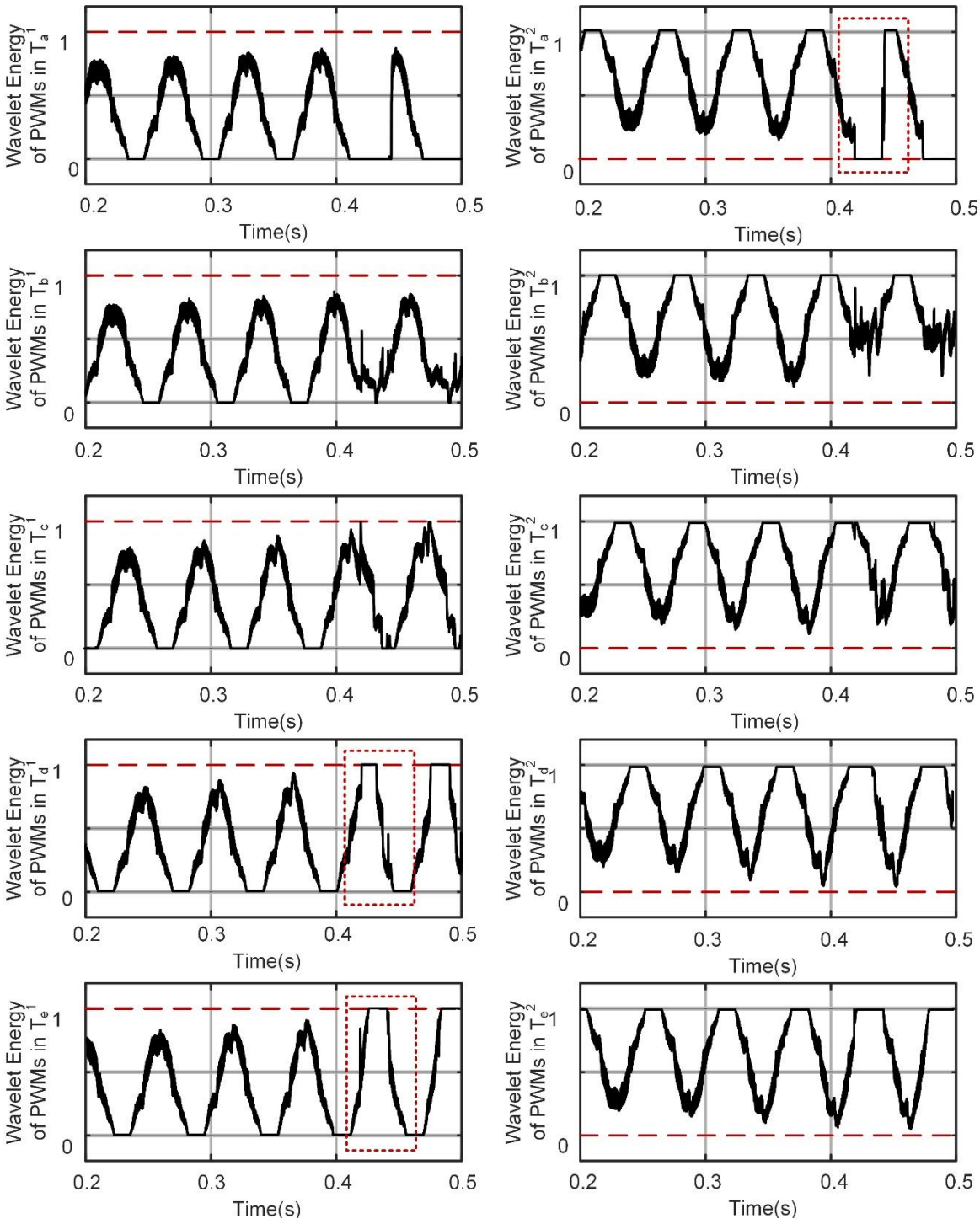

Figure 41 Wavelet energy of PWMs corresponding to top pair switches in five phases for T_{3a} open switch.

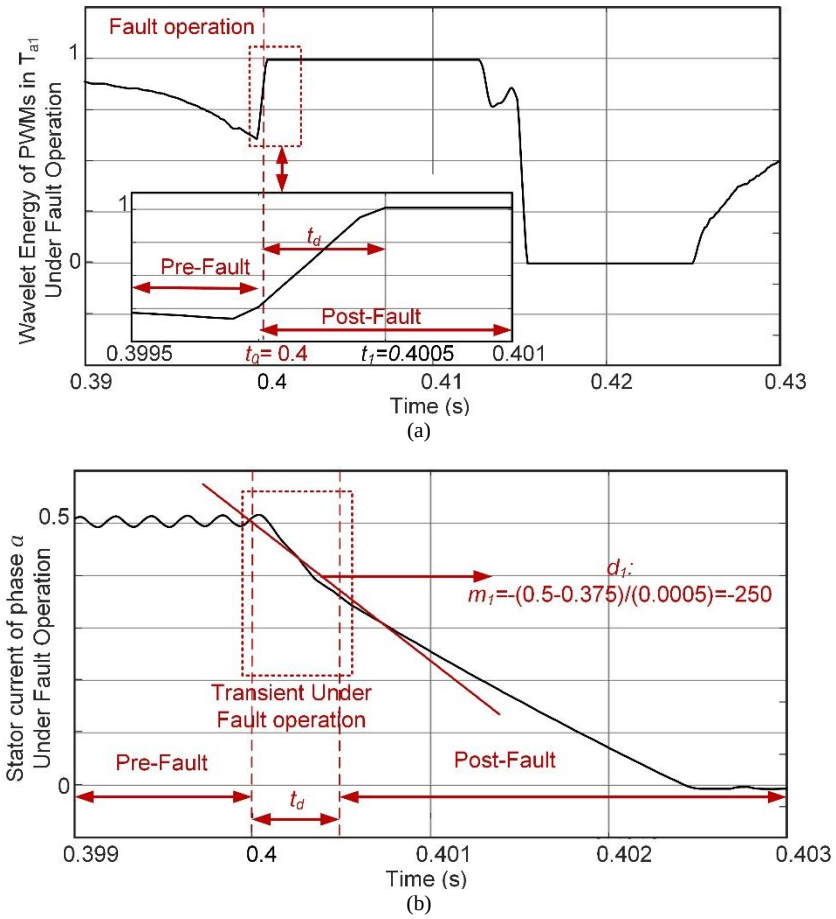

Figure 42 Wavelet energy of T_{a1}, and stator current under failure operation at T_{a1}.

Similarly, Figure 43 represents the wavelet energy of T_{a2} and the stator current of phase a, when T_{a3} is faulted. It was noted before that the PWM of T_{a2} indicates the information about the bottom pair of switch operation. The fault happens at t=0.42s. The wavelet energy of PWM switching signal corresponding to T_{a2} goes to 0 at t=0.4205s [see Figure 43(a)]. The slope of the fitted linear curve to the data over this time interval, shown in Figure 43(b), has the slope equal to 520. The calculated $\Delta V_{az}=|m \times L_T|$=70.03, is close to $2E$, and indicates the failure at T_{a3}. Although in induction machine, motor parameters varies by temperature, the approximation of ΔV_{az} indicates the location of failed switch.

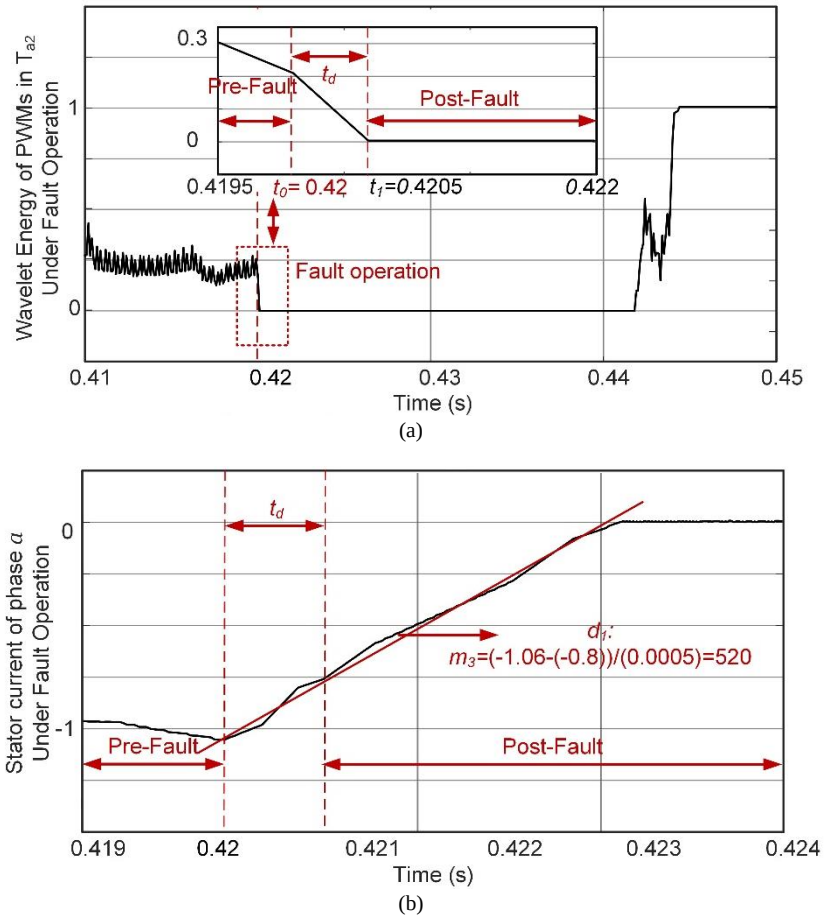

Figure 43 Wavelet energy of T_{a2}, and stator current under failure operation at T_{a3}.

6.4 Chapter Summary

A real-time hybrid fault diagnosis framework, using computational intelligence and the data mining methods, is presented in this paper. The proposed approach can detect and isolate the open-circuited switch in multiphase multilevel neutral point clamped inverters under closed control loop. The scheme consists of a supervisory system, fault detection scheme and fault isolation scheme. Monitoring and supervisory system uses the self-recurrent wavelet-based neuro fuzzy networks for

dynamic system identification. The nonlinear identifier estimates the model of system and predict the stator current values, based on combination of PWM switching signals and previous stator current values, to generate the residual. Then, a decision unit based the concept of threshold detects the fault occurrence and initiates the isolation scheme. Further, the hierarchical fault isolation scheme identifies the faulted phase, and isolates the failed switch. The presented FDI scheme can perform detection and isolation in about 500μs, which is 1% of fundamental period of stator current, and does not require accommodating extra supervisory hardware. Besides, using a data-driven model based fault detection method and analytical data mining methodology make it possible to detect and isolate the failure without prior knowledge of system under fault operation. Also, the intelligent methodology is independent from designed controller, and reduce the computational complexity.

7. CONCLUSION AND FUTURE WORK

7.1 Conclusion

The major focus of this dissertation is on advancements in fault diagnosis methods for multilevel multiphase inverters. A novel fault detection method using data-driven adaptive model-based structures is proposed in this study. A hybrid framework brings the module-level signals (pulse width modulated switching signals), which are generated by control strategy, into play to estimate the system-level signals (stator currents), which are observed by the supervisory system. Taking advantage of PWM signals in predicting the model on entire system enhances the performance of model predictor and yields to an efficient nonlinear model identifier. The adaptive high accuracy identifier estimates the stator currents online and generates the robust fault residuals which are insensitive to noise and is satisfying the invariant relation and fault detectability conditions. Then, two different fault isolation methods are proposed. The first of which applies to all sorts of multilevel inverters is based on machine learning methods. Due to cost, safety, deficient prior knowledge of the system under failure condition, and underlying condition characteristics of the systems are likely to change over time as a function of the operating conditions or aging of the equipment, a semi-supervised clustering method with pairwise constraints is employed. This method combines the labeled data (captured from the condition monitoring system) and unlabeled data (captured from simulation), to enhances the performance of fault analysis system. The second method, which is an analytical pattern recognition method, is specifically for the neutral point clamped inverter. An

analytical procedure is presented to introduce two main criteria; for failed phase identification and failed switch isolation. The faulted phase is identified considering the closed loop controllers performance. The concept of wavelet energy is employed to reflect the fault characteristics. Also, the transient of stator current is investigated to locate the failed switch.

Presented method reduced the detection time to 1% fundamental of stator current, while existing system-based methods, which are using system-level output waveforms, require at least half of 50% fundamental of stator current. Moreover, the suggested approach does not need accommodating extra hardware to monitor semiconductors. The performance of controllers are considered, and the suggested method is applicable for close-loop controlled systems. Last, by not least, fault detection and isolation are performed with minimum prior knowledge of the system.

7.2 Future Work

The works presented in this thesis mainly covered the fault detection and isolation structures and focused on reducing the time and computational complexity. In this section, the major areas where the present work can be extended are identified.

One of the proposed future directions focuses on real-time reconfiguration, Combining intelligent reconfiguration solutions with fault detection and isolation frameworks, based on machine learning methods and computational intelligence, allows the inverters and the electric machine to have a seamless and concurrent post-fault operation.

Also, extending the number of levels in multilevel inverters results in scaling up the complexity and increases the number of redundant states. It is suggested to evaluate the effect of the number of levels on the computational burden and efficiency of fault diagnosis scheme.

Besides, studying the performance of presented fault detection method for gallium-nitride (GaN) and silicon-carbide (SiC) is included in future work studies, because the effect of switching frequency plays a role in identifier directly. Therefore, it worth to evaluate the convergence of model predictor considering the measurements limitations.

REFERENCES

[1] E. Levi, R. Bojoi, F. Profumo, H. A. Toliyat, and S. Williamson, "Multiphase induction motor drives— a technology status review," IET Elect. Power Appl., vol. 1, no. 4, pp. 489–516, Jul. 2007.

[2] E. Levi, "Multiphase electric machines for variable-speed applications," IEEE Trans. Ind. Electron., vol. 55, no. 5, pp. 1893–1909, May 2008.

[3] M. J. Duran and F. Barrero, "Recent advances in the design, modeling, and control of multiphase machines—Part II," IEEE Trans. Ind. Electron., vol. 63, no. 1, pp. 459–468, Jan. 2016.

[4] S. Nandi, H.A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical machines-a review," IEEE Trans. Energy Convers., vol. 20, pp. 719-729, no. 4, Dec. 2005.

[5] L. Parsa, H.A. Toliyat, "Fault tolerant permanent magnet motor drives," in Proc. of IEEE-IAS 2004 Annual Meeting, Seattle, WA, pp. 1048 – 1054, Oct. 3-7, 2004.

[6] P. Lezana, J. Pou, T. A. Meynard, J. P. Rodriguez, S. Ceballos, and F. Richardeau, "Survey on fault operation on multilevel inverters," IEEE Trans. Ind. Electron., vol. 57, no. 7, pp. 2207–2218, Jul. 2010.

[7] R. Isermann, Fault Diagnosis Systems. SpringerVerlag, New York, 2006.

[8] R. Isermann, Fault-diagnosis applications: model-based condition monitoring: actuators, drives, machinery, plants, sensors, and fault-tolerant systems. Springer Science & Business Media, Berlin, 2011.

[9] T.J. Kim, W.C. Lee, and D.S. Hyun, “Detection method for open-circuit fault in Neutral-point clamped inverter systems,” IEEE. Trans. Ind. Electron, vol. 56, no. 7, pp. 2754–2763, July 2009.

[10] J. Rodríguez, J. S. Lai, and F. Z. Peng, “Multilevel inverters: A survey of topologies, controls, and applications,” IEEE Trans. Ind. Electron. , vol. 49, no. 4, pp. 724–738, Aug. 2002.

[11] J. Rodriguez, S. Bernet, P. Steimer, and I. Lizama, “A survey on neutral point-clamped inverters,” IEEE Trans. Ind. Electron., vol. 57, no. 7, pp. 2219–2230, Jul. 2010.

[12] B. Mirafzal, “Survey of fault-tolerance techniques for three-phase voltage source inverters,” IEEE Trans. Power Electron., vol. PP, no. 99, pp. 1–1, 2014.

[13] S. Kim. Hwang, , and Y. Kim, “A survey on fault detection, isolation and reconfiguration methods,” IEEE Trans. Contr. Syst. Tech., vol. 18, no. 3, pp. 636–653, May 2010.

[14] J.S. Lee, K.B. Lee, F. Blaabjerg,“Open-switch fault detection method of a back-to-back converter using NPC topology for wind turbine systems,” IEEE Trans. Ind. Appl., vol. 50, no. 1, pp. 325–335, Jan./Feb. 2015.

[15] R. Peuget, S. Courtine and J. P. Rognon, ‘‘Fault detection and isolation on a PWM inverter by knowledge-based model,’’ IEEE Trans. Ind. App., vol. 34, no. 6, pp. 1318--1326, Nov./Dec. 1998.

[16] U. Choi, J. S. Lee, F. Blaabjerg, and K. B. Lee, “Open-circuit fault diagnosis and fault-tolerant control for a grid-connected NPC inverter,” IEEE Trans. Power Electron., vol. 31, no. 10, pp. 7234-7247, Oct. 2016.

[17] M. B. Abadi, A. M. S. Mendes, and S. M. A. Cruz, “Three-level NPC inverter fault diagnosis by the average current Park’s vector approach,” in Proc. Int. Conf. Electr. Mach., 2012, pp. 1893–1898.

[18] M. S. Mendes, M. B. Abadi, and S. M. A. Cruz, “Fault diagnostic algorithm for three-level neutral point clamped AC motor drives based on average current Park’s vector,” IET Power Electron., vol. 7, no. 5, pp. 1127–1137, May 2014.

[19] M. B. Abadi, A.M.S. Mendes, S.M.A. Cruz, “Method to diagnose open-circuit faults in active power switches and clamp-diodes of three-level neutral-point clamped inverters,” IET Electric Power Applications, vol. 10, no. 7, pp. 623-632, 2016.

[20] U. M. Choi, H. G. Jeong, K. B. Lee, and F. Blaabjerg, “Method for detecting an open-switch fault in a grid-connected NPC inverter system,” IEEE Trans. Power Electron., vol. 27, no. 6, pp. 2726–2739, Jun. 2012.

[21] L. M. A. Caseiro and A. M. S. Mendes, “Real-time IGBT open-circuit fault diagnosis in three-level neutral-point-clamped voltage-source rectifiers based on instant voltage error,” IEEE. Trans. Ind. Appl., vol. 62, no. 3, pp. 1669–1678, Mar./Apr. 2015.

[22] J. He, , N. Demerdash, N. Weise, R. Katebi, “A fast on-line diagnostic method for open-circuit switch faults in SiC-MOSFET based T-Type multilevel inverters,” IEEE Trans. Ind. App., vol. 53, no. 3, pp. 2948 – 2958, May/Jun. 2017.

[23] S. Khomfoi and L. M. Tolbert, “Fault diagnosis and reconfiguration for multilevel inverter drive using AI-based techniques,” IEEE Trans. Ind. Electron., vol. 54, no. 6, pp. 2954–2968, Dec. 2007.

[24] R. Peuget, S. Courtine and J. P. Rognon, ‘‘Fault detection and isolation on a PWM inverter by knowledgebased model,’’ IEEE Trans. Ind. App., vol. 34, no. 6, pp. 1318--1326, Nov./Dec. 1998.

[25] M. Salehifar, R. S. Arashloo, J. M. Moreno-Equilaz, V. Sala, and L. Romeral, “Fault detection and fault tolerant operation of a five phase PM motor drive using adaptive model identification approach,” IEEE J. Emerging Sel. Top. Power Electron, vol. 2, no. 2, pp. 212–223, Jun. 2014.

[26] N. Torabi, V. M. Sundaram, H. A. Toliyat, “On-line fault diagnosis of multi-phase drives using self-recurrent wavelet neural networks with adaptive learning rates." in Proc. IEEE APEC, 2017, pp. 570-577.

[27] N. Torabi, F. Naghavi, H. A. Toliyat, “Real-time fault isolation in multiphase multilevel NPC converters using active semi-supervised fuzzy clustering algorithm with pairwise constraints." in Proc. IEEE IEMDC, 2017, pp. 1-7.

[28] K. B. Lee, J. S. Lee. Reliability Improvement Technology for Power Converters. Springer, Singapore, 2017.

[29] T. A. Lipo. Vector control and dynamics of AC drives. vol. 41. Oxford university press, 1996.

[30] H.A. Toliyat, and H. Xu, "A novel direct torque control (DTC) method for five-phase induction machines". Proc. IEEE APEC, New Orleans, LA, pp.162-168, 2000.

[31] H. Xu, H.A. Toliyat, and L. J. Petersen, "Five-phase induction motor drives with DSP-based control system," IEEE Trans. Power Electron., vol. 17, no. 4, pp. 524–533, Jul. 2002.

[32] H. Xu, H. A. Toliyat, and L. J. Petersen, "Rotor field oriented control of five-phase induction motor with the combined fundamental and third harmonic currents," in Proc. IEEE APEC, 16th Annu. Appl. Power Electron. Conf. Expo., Mar. 4–8, 2001, vol. 1, 392–398.

[33] H.A. Toliyat, "Analysis of Concentrated Winding Induction and Reluctance Motors for Adjustable Speed Drive Applications," Ph.D. Dissertation, Department of. Electrical and Computer Engineering, University of. Wisconsin-Madison, 1991.

[34] L. Parsa, "Performance improvement of permanent magnet AC motors," Ph.D. Dissertation, Department of. Electrical and Computer Engineering, University of. Texas A&M University, 2005.

[35] V. Meenakshi Sundaram," Reliable Multiphase Induction Motor Drives," Ph.D. Dissertation, Department of. Electrical and Computer Engineering, University of. Texas A&M University, 2016.

[36] Y. Zhao. T. A. Lipo, “Space vector PWM control of dual three-phase induction machine using vector space decomposition." IEEE Trans. Ind. App., vol. 31, no.5, pp. 1100-1109,1995.

[37] R. Kianinezhad, B. Nahid-Mobarakeh, L. Baghli, F. Betin, and G. A. Capolino, “Modeling and control of six-phase symmetrical induction machine under fault condition due to open phases,” IEEE Trans. Ind. Electron., vol. 55, no. 5, pp. 1966–1977, May 2008.

[38] H. A. Toliyat, “Analysis and simulation of five-phase variable speed induction motor drives under asymmetrical connections,” IEEE Trans. Power Electron., vol. 13, no. 4, pp. 748–756, Jul. 1998.

[39] Wu, Bin, and Mehdi Narimani. High-power converters and AC drives. John Wiley & Sons, 2017.

[40] R. Isermann, "Supervision fault-detection and fault-diagnosis methods- an introduction", Control engineering practice, vol. 5, pp. 639-652, 1997.

[41] X. Ge, J. K. Pu, B. Gou, Y. C. Liu, "An open-circuit fault diagnosis approach for single-phase three-level neutral-point-clamped converters," IEEE Trans. Power Electron., vol. 33 , no. 99, pp.3179-3184, 2017.

[42] Technical Explanation 3L Evaluation Inverter – SEMIKRON.

[43] K. V.Kumar, P. A.Michael, J. P. John, S. S. Kumar, “Simulation and comparison of SPWM and SVPWM control for three phase inverter”. ARPN Journal of Engineering and Applied Sciences, vol. 5, no. 7., pp. 61-74, 2010.

[44] D.V. Schrick, "Remarks on terminology in the field of supervision, fault detection and diagnosis," in Proc. IFAC Symp. Fault Detection, Supervision Safety for Techn. Process. (SAFEPROCESS), pp. 959-964, 1997.

[45] R. Isermann, Fault-Diagnosis Systems. Springer, Berlin, 2006

[46] R. Isermann, "Trends in the application of model-based fault detection and diagnosis of technical processes," Control Eng. Practice, vol. 5, pp. 709-719, 1997.

[47] Z. Lu, S. Jing, R. B. Kennet, "Linear programming support vector regression with wavelet kernel: A new approach to nonlinear dynamical systems identification." Mathematics and Computers in Simulation 79, no. 7, 2051-2063, 2009.

[48] I. Hwang"A survey of fault detection, isolation, and reconfiguration methods." IEEE Trans. Cont. Sys., vol. 18, no. 3, 636-653, 2013.

[49] R. Isermann, "Process fault detection based on modeling and estimation methods–a survey", Automatica, vol. 20, no. 4, 1984.

[50] R. Isermann,"Model-based fault detection and diagnosis-status and applications", Annual Reviews in Control, vol. 29, Is. 1, 2005.

[51] P.M. Frank, "Survey of robust residual generation and evaluation methods in observer-based fault detection systems." Journal of process control, vol.7, no. 6, 403-424, 1997.

[52] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," IEEE Trans. Pattern Anal. Mach. Intell. vol. 2, no. 7, pp. 674 – 693. July, 1989.

[53] C. Torrence, P. C. Gilbert, "A practical guide to wavelet analysis." Bulletin of the American Meteorological society, vol. 79, no. 1, pp.61-78, 1998.

[54] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis." IEEE trans. on inf. theory, vol. 36, no. 5, pp. 961-1005, 1990.

[55] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matchingpursuit: Recursive function approximation with applications to wavelet decomposition," in Proc. 27th Annu. Asilomar Conf. Signals Syst. Comput., vol. 1, pp. 40–44, Nov. 1993.

[56] S. El Safty, A. El-Zonkoly, "Applying wavelet entropy principle in fault classification." International Journal of Electrical Power & Energy Systems, vol. 31, no. 10, pp. 604-607, 2009.

[57] N. Torabi, M. Karrari, M.B. Menhaj, and S. Karrari, "wavelet based fault classification for partially observable power systems." In proc. APEEC, 2012, pp. 1-6. Mar. 27-29.

[58] N. Torabi, M. Karrari, M. B. Menhaj, "A novel wavelet fuzzy fault location method for partial observable transmission networks based on WAMS/PMU," International Journal on Communications Antenna and Propagation, vol. 1, no. 6, Dec. 2011

[59] S. Pittner and S. V. Kamarthi, "Feature extraction from wavelet coefficients for pattern recognition tasks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 1, pp. 83–88, Jan. 1999.

[60] S. J. Yoo, J. B. Park, and Y. H. Choi, Indirect adaptive control of nonlinear dynamic systems using self recurrent wavelet neural networks via adaptive learning rates, Information Sciences, vol. 177, no. 15, pp.3074-3098, 2007.

[61] F. F. M. El-Sousy and K. A. Abuhasel, “Self-organizing recurrent fuzzy wavelet neural network-based mixed h2/hf adaptive tracking control for uncertain two-axis motion control system,” IEEE Trans. Ind. Inform., vol. 52, no. 6, pp. 5139-5155, Nov./Dec. 2016.

[62] Lin, Cheng-Jian, and Cheng-Chung Chin, “Recurrent wavelet-based neuro fuzzy networks for dynamic system identification." Mathematical and computer modelling, vol. 41, no. 2-3, pp.226-239, 2005.

[63] L. Behera, S. Kumar, and A. Patnaik, “On adaptive learning rate that guarantees convergence in feedforward networks,” IEEE Trans. on Neural Networks, vol. 17, pp. 1116–1125, 2006.

[64] H.Luo, Plug-and-Play Monitoring and Performance Optimization for Industrial Automation Processes. Springer, Fachmedien Wiesbaden, 2017.

[65] S. X. Ding ,"Statistical methods based residual evaluation and threshold setting." Model-Based Fault Diagnosis Techniques. Springer, London, 315-338, 2013.

[66] A. Q. Khan, S. X. Ding, “Threshold computation for fault detection in a class of discrete-time nonlinear systems." International journal of adaptive control and signal processing vol. 25, no. 5, 407-429, 2011.

[67] V.N. Vapnik. The Nature of Statistical Learning Theory. Springer, New-York, 1995.

[68] M.A. Hearst, S.T. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines", IEEE Intelligent Systems, vol. 13, no. 4, pp. 18-28, July-Aug. 1998.

[69] K.R. Muller, S. Mika, G. Rtsch, K. Tsuda, and B. Schlkopf, “An introduction to kernel-based learning algorithms,” IEEE Trans. on Neural Networks, vol. 12, no. 2, pp. 181–201, 2001.

[70] S. Mika, G. Rätsch, J. Weston, B. Schölkopf, and K.-R. Müller, “Fisher discriminant analysis with kernels,” Neural Netw. Signal Process., vol. IX, pp. 41–48, 1999.

[71] E.I. Altman, G. Marco, F. Varetto, “Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks”,Journal of banking & finance vol. 18, no. 3, 505-529, 1994.

[72] W.R. Klecka. Discriminant analysis, volume 19. Sage,1980.

[73] W.J. Krzanowski, Principles of Multivariate Analysis: A User's Perspective. New York: Oxford University Press, 1988.

[74] Z. Xiukuan, M. Li, J. Xu, and G. Song, “Multi-class semi-supervised learning in machine condition monitoring." International Conference on Information Engineering and Computer Science, ICIECS, pp. 1-4, 2009.

[75] C. Olivier, B. Scholkopf, A.Zien, “Semi-supervised learning” ,IEEE Transactions on Neural Networks vol. 20, no. 3, 542-542, 2009.

[76] W. Xiaogang, H. Feng, Y. Fan, “Fault detection and classification for complex processes using semi-supervised learning algorithm." Chemometrics and Intelligent Laboratory Systems, vol. 149, pp. 24-32, 2015.

[77] M. Jianping, J. Jiang, “Semisupervised classification for fault diagnosis in nuclear power plants." Nuclear Engineering and Technology vol. 47, no. 2, pp. 176-186, 2015.

[78] S. Zhong, Q. Wen, Z. Ge, “Semi-supervised Fisher discriminant analysis model for fault classification in industrial processes." Chemometrics and Intelligent Laboratory Systems, vol. 138, pp.203-211, 2014.

[79] Z. Zhi-Hua, M.Li, “Semi-supervised learning by disagreement." Knowledge and Information Systems vol. 24, no. 3, 415-439, 2010.

[80] B. Mikhail, S. Basu, and R. J. Mooney, “Integrating constraints and metric learning in semi-supervised clustering." Proceedings of the twenty-first international conference on Machine learning. ACM, 2004

[81] S. Basu, A. Banerjee, R. J. Mooney, "Active semi-supervision for pairwise constrained clustering," SIAM Int. Conf. on Data Mining, pp. 333-344, Florida, USA, 2004.

[82] N. Grira, M. Crucianu, N. Boujemaa, "Semi-supervised fuzzy clustering with pairwise-constrained competitive agglomeration," Proceedings of the IEEE International Conference on Fuzzy Systems, Reno, pp. 867-872, 2005.

[83] A. Jain, Z. Douglas, "Feature selection: Evaluation, application, and small sample performance." IEEE transactions on pattern analysis and machine intelligence, vol. 19, no. 2, pp. 153-158,1997.