الجمهورية الجزائرية الديمقراطية الشعبية
DEMOCRATIC AND POPULAR REPUBLIC OF ALGERIA
زارة التعليم العالي والبحث العلمي
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

University Kasdi Merbah Ouargla
Faculty of Applied Sciences
Department of Electrical Engineering

جامعة قاصدي مرباح ورقلة
كلية العلوم التطبيقية
قسم الهندسة الكهربائية

Photovoltaic system, control, analysis and fault diagnosis

Presented to the Faculty of Applied Sciences in Fulfillment of the

Requirements for the Degree of Doctorate (Es-Science) in Electrical Engineering

Field: Electrical Engineering

Option: Electrical Engineering

Presented by:

Ali Bouhafs

Supervised by:

Dr. Mohamed Lakhdar Louazene

Co supervised by:

Dr. Mohamed Redouane Kafi

Publicly defended on: 23/11/2022

Chair	Dr. Samai Djamel	KM UOuargla
Examiner	Dr. Amieur Toufik	LTU Tebessa
Examiner	Dr. Meghni Billel	BMU Annaba
Examiner	Dr. Berrahal mokhtar	ENS AA Oran
Supervisors	Dr. Louazene Mohamed Lakhdar	UKM Ouargla
Co supervisor	Dr. KAFI Mohamed Redouane	UKM Ouargla

Academic Year: 2022 – 2023

Abstract

This work deals with fault diagnosis using machine learning algorithms of the inverter used in photovoltaic systems that supply an insulated electrical load and how to safely transfer the current to devices. This thesis discusses the multicellular inverter and describes how this is affected in cases of faults on the load current. And use two modes of control In order to compare in terms of functionality under failures, load current save and Smoothest, and in terms of accuracy built classification model To use sliding mode control mode and exact linearization mode, this is for Purpose of comparison in terms of system performance during failure And the extent of its impact on the load current by examining the shape of its signal And the robustness analysis of the two controls was not significantly affected by defects and their explanation.

Keywords : Machine learning, multi cellular power converter, photovoltaic system, sliding mode control, exact linearization control

الملخص

تتعامل هذه الأطروحة مع تشخيص عطل محول الطاقة، باستخدام خوارزميات التعلم الآلي ، المستخدم في الأنظمة الكهروضوئية التي توفر حملاً كهربائيًا في منا طق معزولة وكيفية نقل التيار بأمان إلى الأجهزة. تناقش هذه الرسالة العاكس متعدد الخلايا ويصف كيف يتأثر ذلك في حالات الأعطال في تيار الحمل. واستخدم طريقتين للتحكم من أجل المقارنة من حيث الوظيفة في ظل حالات الفشل ، وحمل التيار وحفظه ، ومن حيث الدقة ، ونموذج التصنيف المدمج لاستخدام وضع التحكم في الوضع المنزلق ووضع الخطي الدقيق ، وهذا لغرض المقارنة من حيث من أداء النظام أثناء الأعطال ومدى تأثيره على تيار الحمل من خلال فحص شكل إشارته ولم يتأثر تحليل متانة الضابطين بشكل كبير بالعيوب وتفسيرها.

الكلمات المفتاحية: التعلم الآلي ، محول الطاقة متعدد الخلايا ، النظام الكهروضوئي ، التحكم بالوضع المنزلق ، التحكم الخطي الدقيق

List of Figures

Figure 1.1 The photovoltaic effect 04

Figure 1.2: I-V characteristic for an ideal PV cell 05

Figure 1.3 : PV and IV characteristic for real PV cell on T=25℃ 06

Figure 1.4: PV and IV characteristic for a real PV cell on G=1000w/m2 06

Figure 1.5: IV characteristic for ns PV cells 07

Figure 1.6: IV characteristic for np PV cells 07

Figure 1.7: Double exponential PV cell model 08

Figure 1.8: simplified PV cell model 09

Figure 1.9: Application for PV system 10

Figure 1.10 Schematic of boost converter. 11

Figure 1.11 Circuit diagram of boost converter during Mode 1 12

Figure 1.12 Circuit diagram of boost converter during Mode 2 12

Figure 1.13 Boost converter waveforms at CCM. 13

Figure 1.14 : Solar panel characteristic showing MPP and operating points A and B 13

Figure 1.15. Flowchart of P&O method 14

Figure 1.16 Basic Flowchart of HC method. 15

Figure 1.17. The flow chart of INC method [38]. 16

Figure 1.18: Two stages topology 17

Figure 1.19: Single stages topology 17

Figure 1.20: Classification of conventional inverters interfaced with grid connected PV system, based on the PV-module arrangement (Kouro et al., 2015). 19

Figure 1.21: Three leg inverter without a neutral point (balanced output). 20

Figure 1.22: Three-leg inverter with a neutral point. 20

Figure 1.23: Four-leg inverter. 21

Fig.1.24. Diode-clamped multilevel inverter.[53] 21

Fig.1.25. Three-level flying capacitor multilevel inverter topology 23

Figure 1.26:. Multicellular arm having 3 switching cells 24

Fig.2.1: Different convergence modes for state trajectory 29

Fig. 2.2. Representation of the sign function. 39

Fig. 2.3. Linearization of the nonlinear system. 40

Fig .2.4 Exact linearization of the multicellular inverter. 43

Fig. 2.5 Simulink of control by exact linearization of the multicellular inverter with Matlab 44

Fig. 2.6 Scheme of control of the multicellular converter in Matlab 44

Figure 2.7:the voltages of the capacitors and the supply voltage Vdc of the multicellular converter. 45

Fig. 2.8 Three-cell three-phase multi-cellular inverter load current for exact linearization control 45

Fig. 2.9 Simulink the sliding mode control of the multicellular converter with Matlab 46

Figure 2.10: dc and capacitors voltages for sliding mode control 46

Figure. 2.11: sliding mode current load 47

Figure 3.1: supervised learning process 48

Figure 3.2: example of Binary classification representation 49

Figure 3.3: example of Multiclass classification representation 50

Figure3.4: training data needs 52

Figure 3.5: the classification error matrix 54

Figure 3.6: An example of k-nearest neighbor classification. 58

Figure 3.7: The Euclidean distance 59

Figure 4.1. Error voltage C_1 and in C_2 healthy mode both control 65

Figure 4.2. The voltage of capacitor C_1 and C_2, failure C_1 in both controls. 66

Figure 4.3. Error voltage of capacitor C_1 and C_2, failure C_1 in both controls. 66

Figure 4.4 Voltage of capacitor C_1 and C_2, failure C_2 in both controls 67

Figure 4.5. Error voltage of capacitor C_1 and C_2, failure C_2 in both controls. 67

Figure 4.6. Voltage of capacitor C_1 and C_2, failure C_2 and C_2 in both controls 68

Figure 4.7. Error voltage of capacitor C1 and C2, failure C2 and C2 in both controls. 68

Figure 4.8. Feature space during failure of capacitor using sliding mode control 71

Figure 4.9. 3D feature space during failure of capacitor using sliding mode control 71

Figure 4.10. Feature space during failure of capacitor using exact linearization control 72

Figure 4.11. 3D feature space during failure of a capacitor using exact linearization control 72

Figure 4.12. the KNN workflow. 73

Figure 4.13 Histogram represent the comparison between the two modes of control 73

Acknowledgements

We are deeply indebted to our thesis co supervisor Louzene Med lakhdar whose unlimited stead- fast support and inspirations have made this project a great success. I also would like to express my sincere gratitude to my co supervisor Kafi Med Redouane and my friend boubakeur Rouabah for their support, assistance and guidance during all the working period. their suggestions helped me in the whole time of research and for the thesis writing. their manner to simplify complex problems, thinking outside the box and troubleshooting improves my technical skills and pushed me to give all my best. Special thanks go to our friends and families who have contained the hectic moments and stress we have been through during the course of the research project. We thank the University for giving us the grand opportunity to work as a team which has indeed promoted our team work spirit and communication skills. We also thank the individual group members for the good team spirit and solidarity.

My appreciation also extends to the jury's members for evaluating my work.

Contents

Introduction 1
Chapter 1. Photovoltaic system 3
1.1-Introduction 3
1.2-Solar energy 3
1.3- The photovoltaic effect 4
1.4- Influence of irradiance 5
1.5 Influence of temperature 6
1.6-Series connection of photovoltaic cells 7
1.7-Parallel connection of photovoltaic cells 7
1.8- Photovoltaic cell equivalent circuit models 8
1- The Accurate Model 8
2- The Simplest Model 9
1.8.3-Fill factor 9
1.9-Application for photovoltaic system 10
1.10-Boost converter 11
1-Boost Converter Analysis 11
1.11-Maximum Power Point Tracking (MPPT) : 13
1-Perturb And Observe 14
1.11.2- Hill Climbing 14
1.11.3- Incremental Conductance 15
1.12-Photovoltaic system topologies 16
1- Two stages topology 16
2- Single stage topology 16
1.13-inverter topologies for PV system 17
1-Centralized converter topology 17
2-Series-connected string converter topology 18
3- Multi-string parallel-connected topology 18
4-Cascaded dc-dc converters topology 18
5- Three-Phase Inverter Topologies 20
6-Multilevel inverter topologies 21
7-Multicellular inverter Topology 23
7.1-The Modeling of Multicellular inverter 24
1.14Conclusion 26
Chapter 2. Control of Multicellular Power inverter 27
2.1 Introduction 27
2.2. Modeling of the three-cell multicellular inverter 27
2.2.1 sliding mode control 28
1 Definition 28
2 Principle of sliding modes control. 29
3. Choice of sliding surface. 30
4. Convergence and existence conditions. 30
5 Determination of the control law. 31
6. Sliding mode control of a multicellular converter: 34
2.3 Exact linearization Control mode. 37
1. Definition 37

2. Regulation loop and mathematical equations of a multicellular converter 38
3. Control by exact linearization of a three-phase multicellular inverter. 41
2.4. Simulation of a multicellular inverter with the exact linearization control 43
2.4. Simulation of a multicellular inverter with the sliding mode control. 45
2.5 Conclusion 47
Chapter 3. Machine Learning Fault Detection Methodologies. 48
3.1-Introduction 48
3.2 Types of machine learning 48
3.2.1 Supervised learning 48
3.2.2 Unsupervised learning 49
3.2.3 Semi-supervised learning 49
3.2.4 Reinforcement learning 49
3.3 Classification 50
3.3.1 Types of Classification 50
3.3.1.1 Binary Classification 50
3.3.1.2 Multiclass Classification 50
3.4-Dataset 51
3.4.1- Learning data in supervised and unsupervised learning 51
3.4.1.1 Training Data 51
3.4.1.2 Labeled Data 52
3.4.2 Datasets for training, validation and testing 53
3.4.3 Data Quantitative Requirements 53
3.4.4 Improving the Quality of Training Data 54
3.4.5 Metrics in classification problems Metrics 54
3.4.5.1 Accuracy, precision, recall and specificity 55
3.4.5.2 Accuracy: 55
3.4.5.3 Precision, recall and F-measure 56
3.4.6 Training Data Preparation 56
3.4.6.1 Data cleaning 56
3.5 The k Nearest Neighbors (kNN) 58
3.5.1 Theoretical component of the k-NN algorithm 59
3.5.1.1 Euclidean metric 59
3.5.1.2 Feature selection 60
3.5.1.3 Z-normalization 61
3.5.2 Min-Max normalization 61
3.6 Spectral analysis 62
3.6.1 Fourier Transform 62
3.7 Conclusion 64
Chapter 4. Implementation of machine in photovoltaic system 64
Introduction 64
3.1-- Simulation Results 64
3.2. Data Processing and Feature Space Analysis 70

3.3-. K-Nearest Neighbor (KNN) implementation 72
3.4.Conclusion 74
Conclusion 75
Bibliography 76

Introduction

Because it is a clean energy source that does not emit any gas and is friendly to the environment, solar photovoltaic (PV) energy has recently been in high demand for the production of electric power [01]. This has led to a wide development in many areas, such as avoiding new electric transportation projects, lowering electric energy bills, and using PV energy to provide water in remote areas. In order to convert the direct current (DC) voltage produced by photovoltaic panels to alternating current (AC) voltage, the power converter is used as an interface between the photovoltaic panels and the electric load. The power converter accounts for between 43% and 70% of PV power plant service requests [02]. However, when the power converter fails, all of the advantages may be negatively impacted, including no water extraction in isolated areas, high costs, increased maintenance intervention time, and a significant impact on the reliability of the PV systems [There are a number of factors that prevent power converters from functioning properly, particularly power switches and flying capacitors. Multicellular power converter control in [03] and [04], a solar PV system based on Step-Up Boost converter in order to extract the maximum power from solar panels [05], in addition to treating wind energy as in [06], wind turbine converters in [07] and [08], the evolution also included the controlling systems using exact linearization control in [09] and active filtering in [10] in [11], including the optimization in conversion energy efficiency in [12], the contribution of multicellular power converter as mentioned The fault diagnosis of a four-level multicellular power converter is included in this thesis.

The fault diagnosis of a four-level multicellular power converter is included in this thesis. It focuses on capacitor failures that cause distortions in the shape of the load current, including the appearance of harmonics that affect electric loads by shortening devices' lifespans, increasing mechanical vibrations, and overheating machine windings. Depending on the environment, these flaws may also result in catastrophic damage. Harmonics are the disturbances that cause the signal waveform to spread out. A few of the effects of harmonics on electrical installations and equipment include distortions of waveforms that cause malfunctions, rises in peak values that cause dielectric breakdowns, which in turn cause overheating and additional losses in voltage and current, and a spectral spread that causes vibrations and mechanical fatigue. Oversizing, decreased productivity, decreased energy efficiency, and increased costs are just a few of the factors that have a significant negative impact on the economy. This work proposes fault detection methods for tracing the formation

of faults in multicellular power converter systems and identifying the fault's source so that faults can be addressed directly without affecting the healthy components. Due to its accuracy in automatic classification, such as in [13], the machine learning method has received a lot of use. It topped the list of scientific research concerns, medical [14], electrical [15].By following the decision tree [16], it will be able to provide us with the correct classification through simple symbols that are humanly understandable. It uses algorithms that are similar to human thinking to distinguish between types. Classification is achieved by a series of logical processes. These are capable of modeling even the most complex problems with high efficiency when provided with sufficient data during the model's training [17].

To ensure the effectiveness of this method used in detecting faults regardless of the circumstances, two methods were used to control the sliding mode and exact linearization mode to determine the extent to which the fault detection algorithm is affected by the control model.

This thesis consists of three chapters. The first chapter is devoted to the available photovoltaic systems, their method of operation, their characteristics, and the method of connecting them, their components from the photovoltaic cell, the DC/DC, the inverters topologies, and control mode of them. The chapter is also concerned with machine learning and its use in classifying faults .In the third chapter, the machine learning method was applied to classify the faults of multicellular inverter capacitors for the sliding and exact linearization control modes .Finally, we will end with a general conclusion.

Chapter 1

Photovoltaic system

1.1-Introduction

In this chapter, we will discuss solar energy and its use in the production of electricity in all parts of the world, then we will talk about the photovoltaic system in all its parts, and we will focus on the main part of our work, which is the multicellular inverter, the operating principle and modes of control.

1.2-Solar energy

Due to the growth demand of global energy and adverse effects of conventional energy such as pollution caused by fossil fuel and nuclear fission sources, the exploration of the renewable energy sources (RES) is increased [17]. Renewable energy, as clean and alternative energy, is based on self-renewing energy sources including the solar energy, wind energy, wave energy, tidal energy, ocean thermal energy, hydropower, the geothermal energy, and biomass energy [18-20]. With increasing concern about the environmental issues, RES are paid more and more attention. The growth rate of renewable power installation has exceeded that of the fossil fuel and nuclear power capacity combined [21-24]. For example, in 2021, the installed renewable power capacity is increased more than 290 GW, mostly are PV which has the largest growth rate ever and enhanced the global total to 2,588 GW by the end of year [24]. The Renewables Market Report said growth will accelerate to average 305 GW per year for 2021 through 2026. By the end of 2026, global renewables capacity will reach 4,800 GW, or 60% greater than at end 2020. Renewables will represent about 95% of the new power capacity installed worldwide from now through 2026 [25].

Since the electricity generated from RES is more cost effective than that from the coal-fired power plants [18], it is cheaper to build new wind or PV plants than utilizing existing coal-fired power plants [19]. Besides, renewables also beat new natural gas power station on cost in many locations and has become the cheapest sources of new electricity generation on the earth (excluding Antarctica [20]).

Among the clean energy technologies, PV has significantly grown in recent years [22],[24]. Not only are the efficiencies of the most domestic solar panels low, i.e., around 10-20%, but the performance of other components such as inverters and batteries are limited as well. Battery, which can provide fast response for balancing the power between the generation and consumption [26], is becoming a good candidate for the electrical energy storage system (ESS). However, the initial installation costs are still high [27],[28].

1.3- The photovoltaic effect

The solar energy is the energy that comes from the sun; the photovoltaic cells convert that energy into electricity. This phenomenon is called the photovoltaic effect [29], the light enters a PV cell and imparts enough energy to some electrons (negatively charged atomic particles) to free them. A built-in-potential barrier in the cell acts on these electrons to produce a voltage (the so-called photo voltage) which can be used to drive a current through a circuit. The figure 1.1 present the photovoltaic effect [30].

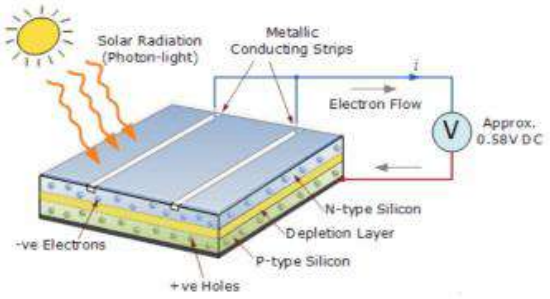


Figure 1.1 The photovoltaic effect

A PV cell is a p-n semiconductor junction fabricated from semiconductor (usually silicon). Solar radiation is made up of particles called photons. A photon is characterized by its Wavelength (λ) and energy (E):

$$E = h * \frac{C}{\lambda} \qquad (1.1)$$

Where,

E: photon energy[J]

h: Plank's constant (6, 62607004 $*10^{-34}$ [J.s])

c: the speed of light (299 792 458 [m/s])

λ: wavelength of photon [m]

For generating a current, the energy of the photon must be greater than the gap of the material [31].

To get the IV characteristic, a resistive load is used and each time the resistance value is changed from very small resistance (***Isc***, V=0) to very high resistance (I=0, V=***Voc***). The figure 1.2 indicates the IV characteristic for an ideal pv cell.

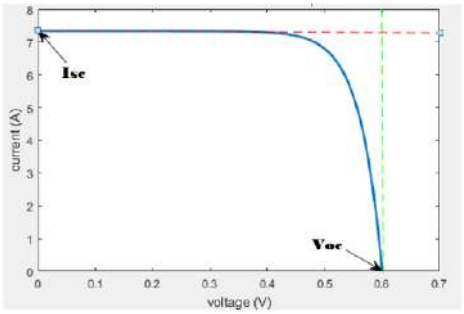


Figure 1.2: I-V characteristic for an ideal PV cell

1.4- Influence of irradiance

The irradiance is an external parameter that depends on the exposure of the cell to the sun; it is the most influencing factor on the parameters of the cell. The photo-current is proportional to the irradiance and opposed to the diode current.

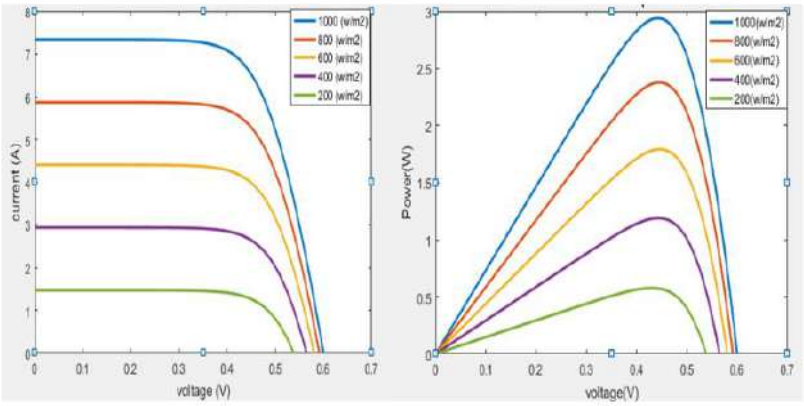


Figure 1.3 : PV and IV characteristic for real PV cell on T=25°C

By varying the irradiance G, is noticed that the open circuit voltage Voc has a small variation, contrary to the short-circuit current Isc which varies a lot in function of the irradiance, which causes a huge variation in the power of the photovoltaic cell.

1.5 Influence of temperature

Temperature is the second most important parameter influencing the characteristics of the PV cell. The figure 1.4 shows IV and PV characteristics for a real PV cells under constant irradiance (G= 1000 w/m^2):

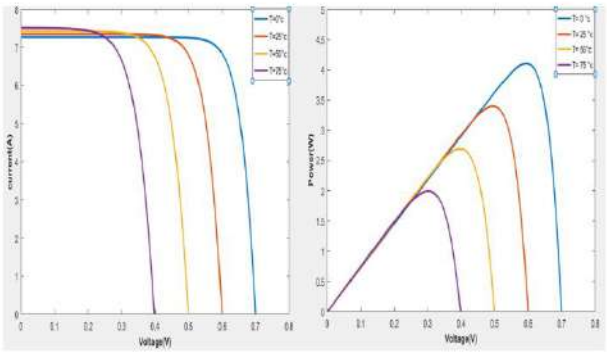


Figure 1.4: PV and IV characteristic for a real PV cell on G=1000w/m^2

Unlike the variation in irradiance, the variation in temperature mainly affects the open circuit voltage Voc . When the temperature increases, the open circuit voltage decreases and vice versa. The maximum power is achieved when we have a maximum of irradiance and a minimum of temperature.

1.6-Series connection of photovoltaic cells

An association of ns cells in series increases the voltage; the cells will be crossed by the same current. The output voltage is the addition of every PV cell voltage as can be noticed from the next equation and figure 1.5.

$$\begin{cases} V_{OC\ ns} = n_s * V_{OC} \\ I_{SC\ ns} = I_{SC} \end{cases} \quad (1.2)$$

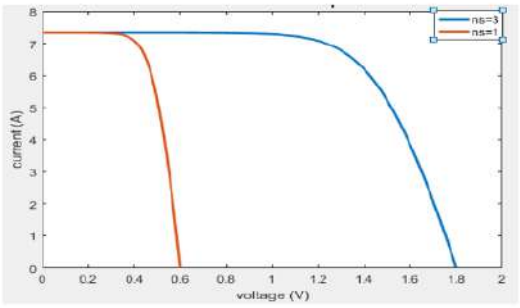


Figure 1.5: IV characteristic for ns PV cells

1.7-Parallel connection of photovoltaic cells

On the other side, a parallel association of np cells increases the current provided by the generator. A number of cells connected in parallel are under the same voltage. The output current is the addition of every PV cell current as presented in the equation (1.3) and the figure 1.6.

$$\begin{cases} V_{OC\ np} = V_{OC} \\ I_{SC\ np} = n_p * I_{SC} \end{cases} \quad (1.3)$$

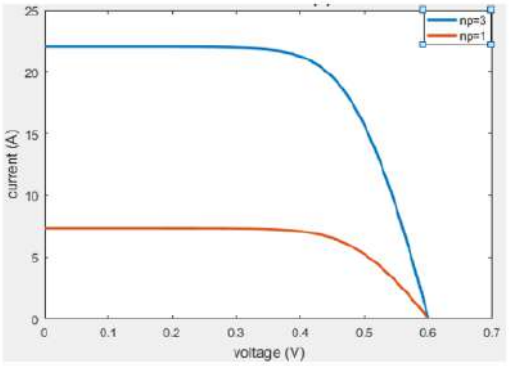


Figure 1.6: IV characteristic for np PV cells

The operating voltage depends on the configuration of the power supply system, and the surface area of the module is variable depending on the constructor. The assembly of the modules in series and/or in parallel will make it possible to adjust different voltages and powers [32], [33].

1.8- Photovoltaic cell equivalent circuit models

1- The Accurate Model

The equivalent circuit model of a PV cell is needed in order to simulate its real behavior. One of the models proposed in literature is the double exponential model depicted in figure 1. Using the physics of p-n junctions, a cell can be modeled as a DC current source in parallel with two diodes that represent currents escaping due to diffusion and charge recombination mechanisms. Two resistances, Rs and Rp , are included to model the contact resistances and the internal PV cell resistance respectively. The values of these two resistances can be obtained from measurements or by using curve fitting methods based on the I-V characteristic of the cell.

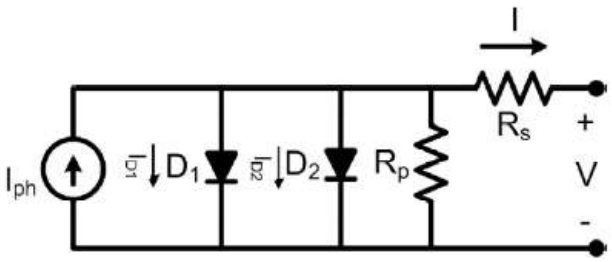


Figure 1.7: Double exponential PV cell model

The relationship between the PV cell output current and terminal voltage is governed by:

$$\begin{cases} I = I_{ph} - I_{D1} - I_{D2} - \frac{V+IR_S}{R_P} \\ I_{D1} = I_{01}\left[exp\left(\frac{q(V+IR_S)}{akT}\right) - 1\right] \\ I_{D2} = I_{02}\left[exp\left(\frac{q(V+IR_S)}{akT}\right) - 1\right] \end{cases} \quad (1.4)$$

Where,

Iph is the PV cell internal generated photocurrent,

ID1 and ID2 are the currents passing through diodes D1 and D2,

a is the diode ideality factor, k is the Boltzmann constant ($1.3806503 \times 10^{-23}$ J/K),

T is the cell temperature in degrees Kelvin, q is the electron charge ($1.60217646 \times 10^{-19}$ C),

I_{01} and I_{02} are the reverse saturation currents of each diode respectively.

2- The Simplest Model

Assuming that the current passing in diode D2 due to charge recombination is small enough to be neglected, a simplified PV cell model can be reached as shown in figure 2.

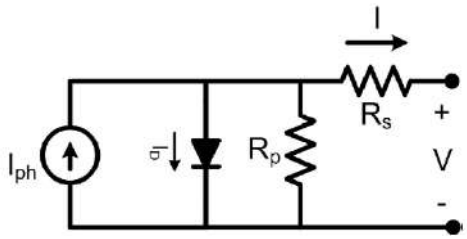


Figure 1.8: simplified PV cell model

This model provides a good compromise between accuracy and model complexity and has been used in several previous works. In this case, current I_{D2} can be omitted from (1.4) and the relation simplifies to:

$$I = I_{ph} - I_0 \left[exp\left(\frac{q(V+IR_S)}{akT}\right) - 1 \right] - \frac{V+IR_S}{R_p} \qquad (1.5)$$

It is clear that the relationship between the PV cell terminal voltage and output current is nonlinear because of the presence of the exponential term in 1 and 2. The presence of the p-n semiconductor junction is the reason behind this nonlinearity. The result is a unique I-V characteristic for the cell where the current output is constant over a wide range of voltages until it reaches a certain point where it start dropping exponentially [34].

3-Fill factor

The fill factor (FF) as expressed in Eq 1.5 may be calculated using the voltage and current at MPP, VMPP and IMPP , the short circuit current (ISC), and the open-circuit voltage (VOC).

$$FF = \frac{I_{MPP}V_{MPP}}{I_{SC}V_{OC}} \qquad (1.6)$$

It is a broadly utilized indicator of the all-inclusive condition of a solar cell. It is the proportion of maximum power (IMPP VMPP) to conceptual maximum power (ISC VOC), which is not achievable. Due to the general series and shunt resistances, as well as the diode illustrated in diagram in figure 1.8, the MPP voltage and current are less than the open-circuit voltage and short-circuit current, respectively [35].

1.9-Application for photovoltaic system

Photovoltaic energy sources may be used in both isolated and grid-connected applications as shown in fig 1.9. Battery charging, Water pumping, refrigeration, residential power supplies, street lighting, and satellite power systems. telecommunications, heating systems, hybrid vehicles, military space, swimming pools, and hydrogen manufacturing are all applications for photovoltaic energy sources [36] ,[37]

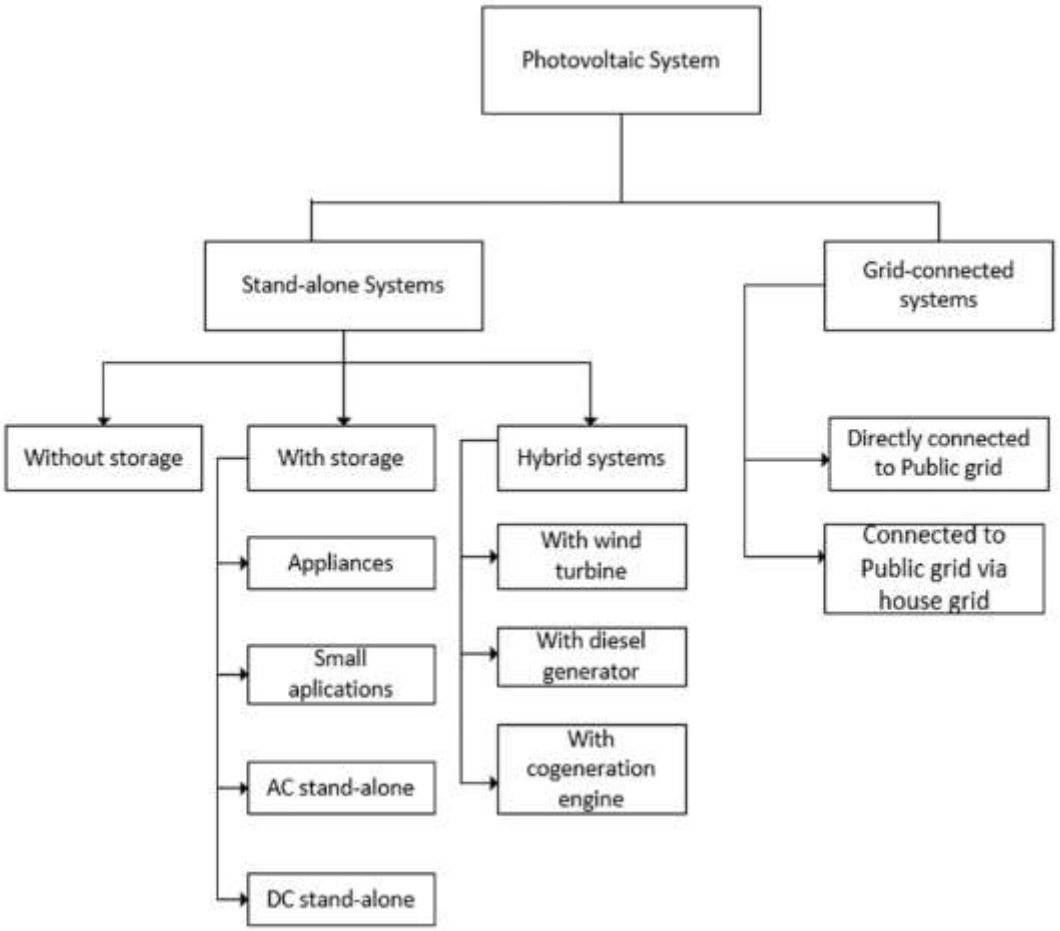


Figure 1.9: Application for PV system

1.10-Boost converter

The boost converter is a medium of power transmission to perform energy absorption and injection from solar panel to grid-tied inverter. The process of energy absorption and injection in boost converter is performed by a combination of four components which are inductor, electronic switch, diode and output capacitor. The connection of a boost converter is shown in Figure 1.10. The process of energy absorption and injection will constitute a switching cycle. In other word, the average output voltage is controlled by the switching on and off time duration. At constant switching frequency, adjusting the on and off duration of the switch is called pulse-width-modulation (PWM) switching. The switching duty cycle, k is defined as the ratio of the on duration to the switching time period. The energy absorption and injection with the relative length of switching period will operate the converter in two different modes known as continuous conduction mode (CCM) and discontinuous conduction mode (DCM).

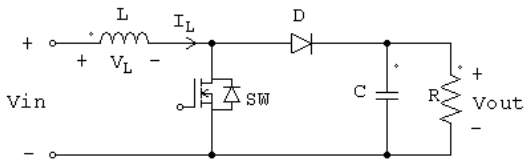


Figure 1.10 Schematic of boost converter.

1-Boost Converter Analysis

For this case, it is considered that the continuous conduction mode. Continuous Conduction Mode Under CCM, it is divided into two modes. Mode 1 begins when the switch SW is turned on at t = 0 as shown in Figure 3. The input current which rises flows through inductor L and switch SW. During this mode, energy is stored in the inductor and load is supplied by capacitor current. Mode 2 begins when the switch is turned off at t = kT. The current that was flowing through the switch would now flow through inductor L, diode D, output capacitor C, and load R as shown in Figure 1.11. The inductor current falls until the switch is turned on again in the next cycle. During this time, energy stored in the inductor is transferred to the load together with the input voltage. Therefore, the output voltage is greater than the input voltage and is expressed as

$$V_{out} = \frac{1}{1-k} V_{in} \qquad (1.7)$$

where V_{out} is the output voltage, k is duty cycle, and V_{in} is input voltage .

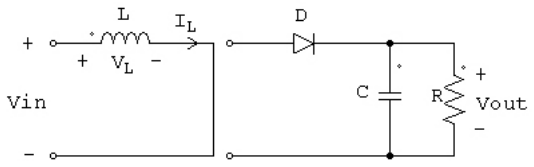


Figure 1.11 Circuit diagram of boost converter during Mode 1.

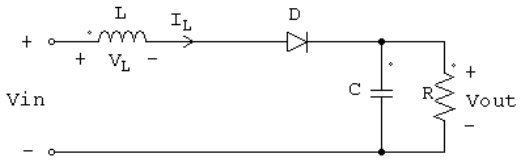


Figure 1.12 Circuit diagram of boost converter during Mode 2.

In order to operate the converter in CCM, the inductance is calculated such that the inductor current I_L lows continuously and never falls to zero as shown in Figure 1.12. Thus, L is given by

$$L_{min} = \frac{(1-k)^2 kR}{2f} \qquad (1.8)$$

where L_{min} is the minimum inductance, R is output resistance, and f is the switching frequency of switch SW . The output capacitance to give the desired output voltage ripple is given by

$$C_{min} = \frac{k}{RfV} \qquad (1.9)$$

Where C_{min} is the minimum capacitance and Vr is output voltage ripple factor. Vr can be expressed as

$$V_r = \frac{\Delta V_{out}}{V_{out}} \qquad (1.10)$$

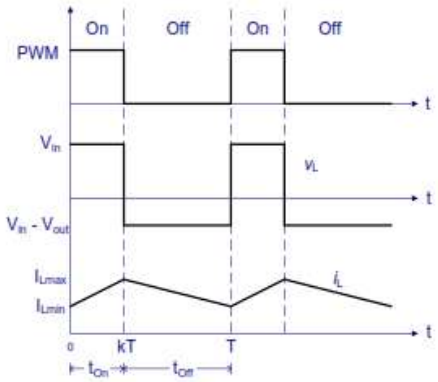


Figure 1.13 Boost converter waveforms at CCM.

1.11-Maximum Power Point Tracking (MPPT)

A typical solar panel converts only 30 to 40 percent of the incident solar irradiation into electrical energy. Maximum power point tracking technique is used to improve the efficiency of the solar panel. According to Maximum Power Transfer theorem, the power output of a circuit is maximum when the Thevenin impedance of the circuit (source impedance) matches with the load impedance. Hence our problem of tracking the maximum power point reduces to an impedance matching problem. In the source side we are using a boost convertor connected to a solar panel in order to enhance the output voltage so that it can be used for different applications like motor load. By changing the duty cycle of the boost converter appropriately we can match the source impedance with that of the load impedance [34].

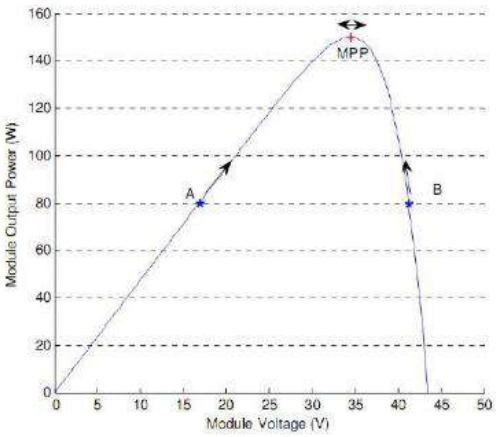


Figure 1.14 : Solar panel characteristic showing MPP and operating points A and B

Over the years, numerous MPPT algorithms have been proposed to track the maximum power of PV panel. Even though these algorithms are proposed for the same purpose, they differ tremendously regarding efficiency, tracking speed, steady state oscillations, complexity, hardware implementation, track global MPP or not and cost. Moreover, each method may work effectively in certain conditions while not in others. For instance, some MPPT methods yield better performance under stable irradiance. Conversely, under fast change of irradiance the results are found to be unsatisfactory

1-PERTURB AND OBSERVE

P&O is the most widely used MPPT method due to its simplicity. The P&O operation principle is presented by the flowchart in Fig. 1.15. As its name suggests, it works by introducing a perturbation (offset) in the PV panel's operating voltage or current according to the variation in operating power that is observed using the samples of voltage (V(k)) and current (I(k)) (Ishaque et al., 2014).

The amount of perturbation value 'offset' depends upon the nature of the algorithm. It can be constant or varying. Pertaining to this fact, two further groups of P&O method are available in the literature, namely fixed-step P&O and variable step P&O methods.

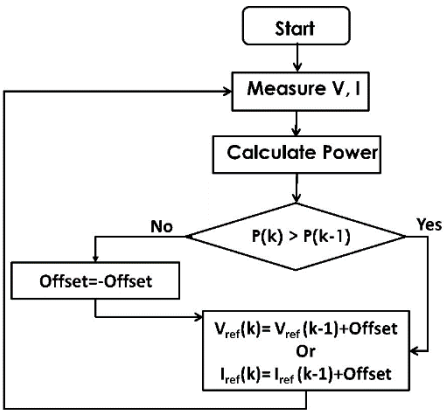


Figure 1.15. Flowchart of P&O method.

1.11.2- Hill Climbing

The HC method is shown in Fig. 1.16. Its principle operation is like P&O, but the duty cycle is perturbed rather than perturbing the current or voltage to update the operating point of the

PV panel (Ishaque and Salam, 2013). The HC method periodically adjusts the duty cycle α(k) by a fixed perturbation (offset) with the direction of increasing power. The perturbation direction is reversed when the change of power is negative. This implies that the PV system is not moving towards the MPP (Ishaque and Salam, 2013). Since it works based on the perturbation mechanism, this algorithm can be based on the fixed or variable step. Nevertheless, under partial shading condition (PSC), the P-V curve contains several peaks (global and local MPPs) (Bayrak et al., 2017). Therefore, based on the principle of P&O and HC, they fail to find Global MPP (GMPP) (Ishaque, 2012).

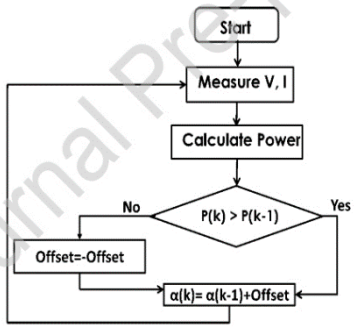


Figure 1.16 Basic Flowchart of HC method.

1.11.3- Incremental Conductance

The P&O method has two major limitations as mentioned in (Wasynczuk, 1983). (1) Due to a fixed amount of perturbation at the steady-state, a small power variation around the MPP is always there which contributes to some power losses. (2) Under rapid fluctuations of environment, the operating point is most likely to diverge from the true MPP. The INC method was proposed in (Wasynczuk, 1983) to circumvent the disadvantage mentioned above. The INC is based on the fact that the derivative of (dP/dV) is zero at MPP.

Hence, the basic idea of INC method is incrementally comparing the instantaneous conductance with the ratio of derivative of conductance (Wasynczuk, 1983). The basic flow chart of INC algorithm is depicted in Fig. 1.17. [38]

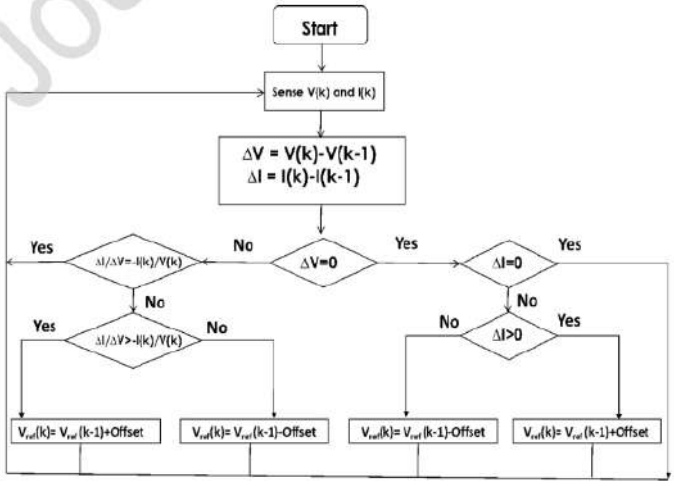


Figure 1.17. The flow chart of INC method [38].

1.12-Photovoltaic system topologies

Now there are mainly two types of power electronic topologies being used in photovoltaic generation field, i.e., single stage topology and two stages topologies. The two stages topology is composed of former DC/DC part and latter inverter part. One stage consists of the inverter only. The merit of two stages topology lies in the convenience of designing its control scheme, but has to burden more power loss than that of single stage [39]. Single stage, on the contrary, can achieve relatively higher power efficiency, but the control scheme is more complex since the inverter alone must achieve all of the control objectives: grid current following, power factor constant and MPPT function [40, 41]. About these two topologies, presently there are not many literatures investigating the difference of their design details and little conclusion refer to their most suitable applying occasion respectively [42].

1- Two stages topology

Commonly, the output voltage of the PV array is not high enough to connect to the grid. Moreover, the voltage source inverter (VSI) usually has a voltage-down property, which causes the “PV array + Inverter” topology to output a lower voltage, thus two stages topology is suggested. This topology adds a voltage-up link part, usually configured as Figure 1.18 The DC/DC part often adopts a Boost circuit or some other derived versions, like Buck-boost, isolated Boost, etc.[43,44]. Besides voltage-up function, the Boost circuit can also offer a more stable input voltage for the inverter. The main advantage of the two stages topology is

the flexibility of designing its control scheme since it has a higher freedom degree, i.e. more controllable variables, which means multiple control objectives (MPPT, grid connecting, var compensating, active filter etc.) can share by two stages respectively simultaneously [41, 46].

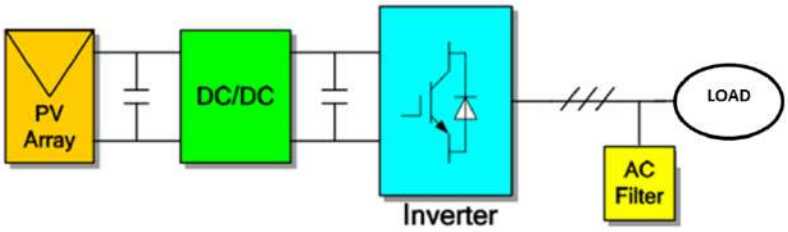


Figure 1.18: Two stages topology

2- Single stage topology

Although two stages topology has advantages in controller design, it also has some deficiencies [45, 47]. With the circuit stages increasing, the power loss rises as well that makes the holistic energy transferring efficiency decrease; more stages also adds system complexity, thereby reduce the system reliability. To enhance system efficiency, system only relying on the inverter, i.e. so-called single stage topology has been suggested, as shown in Figure 1.19.

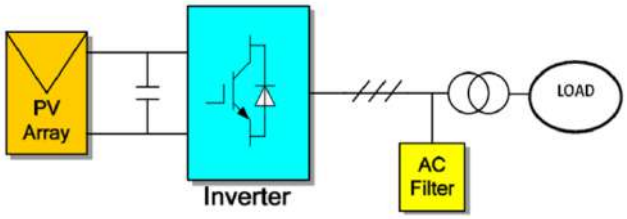


Figure 1.19: Single stages topology

1.13-inverter topologies for PV system

1-Centralized converter topology

In this topology, the PV panels are wired in a single common array (in series and or parallel) and the array is connected to a single inverter. This topology is economic due to the small number of inverters, but the partial shadowing of the one panel will affect the whole array power output. This topology has two kinds of connection:

2-Series-connected string converter topology

Here the array of the PV panels is split to the several strings in series. Losses due to electric conductors are important when the voltage is low [48]. This topology is characterized by the high voltage and low current, or in other words it allows reduction in the cost of wires, decrease in the voltage drops and thus power losses the conductors. Strings are connected one by one to inverters with MPPT and the outputs of the inverters are connected to the transmission lines. Figure 1.20 shows this topology.

When one or several panel is shadowed in a string, other string are not affected and irradiance level remains same for other string. Each PV modules is equipped by parallel diode to bypasses it if its output voltage is low. The topology is appropriate for installation where the irradiance is regularly altered for complete panel string, while irradiation of other panel strings is not disturbed.

3- Multi-string parallel-connected topology

This topology as shown in the Figure 1.20 is similar to the series connected string topology with the difference of having separate inverter and dc-dc converter with MPPT. Here the converters are simpler and thus are cheaper and at the time more efficient. Every string of the PV panels is connected to dc-dc converter and converters connected to common inverter that is in turn connected to the grid.

The separate MPPT and inverter control reduces the cost of the PV system. The other advantage of this topology is that the inverter provides high power rate and the system efficiency is higher.

4-Cascaded dc-dc converters topology

MPPT function in this topology, as shown in Figure 1.20, is performed for each of the PV panels. Each PV module is equipped by its own converter and MPPT controller. Converters are then connected in parallel or series; however the series connection is more common. In the series connection if a panel is shadowed the converter is gaining maximum available power, but other panels are not affected, the output current of each converter must have the same value.

This topology assures the optimal power generation process and maximal total yield if efficiency is not taken into account.

The cascaded converter topology with central inverter provide better distribution level and performance during the partial shadowing and thus is recommended for the building integrated photovoltaic. Higher distribution level implies higher number of converters that allows better performance and increases the harvested power.

But the large number of converters decreases the overall efficiency because of the decrease of power rate of the converters (and small power rate converters are relatively less efficient than high power converter). But this topology is able to reach higher total energy than other systems when the shadowing is significant. On the other side centralized and multi string inverter topologies are common for the solar plants and roof-top systems, which are not affected by surrounding, since high efficiency can be reached at higher power rates.

To sum up, less distributed systems in case of low irradiation disturbance presents higher performance due to higher efficiency that could be achieved relatively easy. Systems that are more distributed have better performance in cases of the frequent partial shading close to the location of the PV panel and thus can shade the panels partially [49].

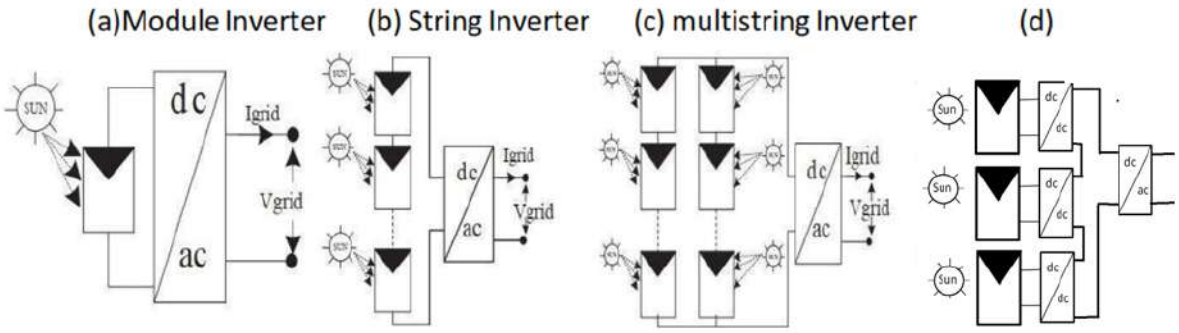


Figure 1.20: Classification of conventional inverters interfaced with grid connected PV system, based on the PV-module arrangement (Kouro et al., 2015).

5- Three-Phase Inverter Topologies

Two configurations able to generate three-phase asymmetrical signals are existed. These are the three-leg neutral point build by capacitors and the four-leg inverter with a controlled neutral point by the fourth leg. Three-phase inverters with neutral point are an evolution from the single-phase ones. Three half-bridge single-phase inverters joined together can be seen as a three-phase neutral point inverter, see Fig. 1.21, where each output feeds one phase. This topology can be used to feed balanced or unbalanced loads. In case of unbalanced loads, the sum of the output currents ia , ib , and ic will not be zero and the neutral current will flow in the connection between the neutral point and the mid-point of the capacitive divider [50,51,52]. To maintain a symmetrical voltage across the two capacitors an adequate power electronic and a voltage stage management are needed, this will not be taken further into discussion.

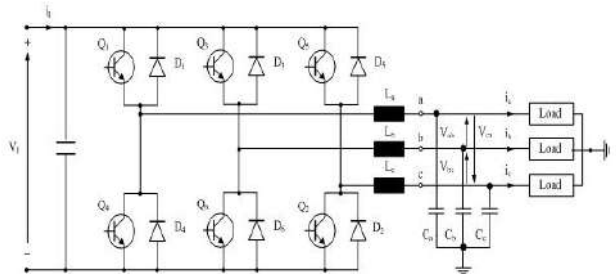


Figure 1.21: Three leg inverter without a neutral point (balanced output).

The general power electronic topology of the four-legged inverter is shown in Figure. 1.22. the goal of the three-phase four-leg Inverter is to supply a desired sinusoidal output voltage waveform to the load for all load conditions and transients. Compared with the four leg inverter, the three leg inverter has a lower number of semiconductor switches and the control function can be built like three individual single line inverters.

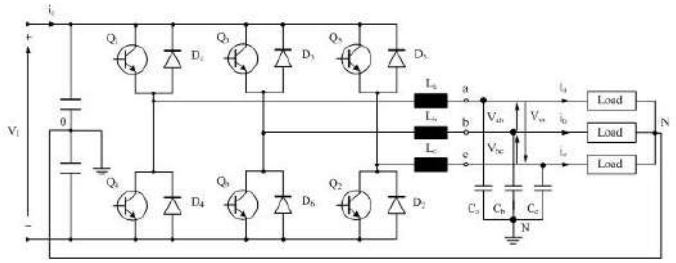


Figure 1.22: Three-leg inverter with a neutral point.

However the four-leg inverter still have the advantages of higher utilization of the DC link voltage, small DC link capacitor as no zero sequence current flow across the DC link capacitor and an additional degree of freedom due to the 4th leg [52].

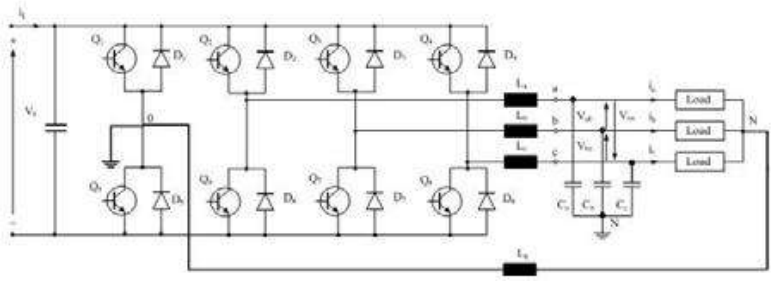


Figure 1.23: Four-leg inverter.

6--Multilevel inverter topologies

Nowadays, multilevel inverters are widely used in power industry. It starts from three level inverter. Voltage unbalance problem is one of the major issue in working of multilevel inverter. So to reduce it, there are main three types of multilevel inverter are used. These are as follows:

• Diode-Clamped Multilevel Inverter

The diode-clamped multilevel inverter is shown in fig.1.24. It requires (m-1) capacitors on dc bus to produce m levels of voltage.[53]

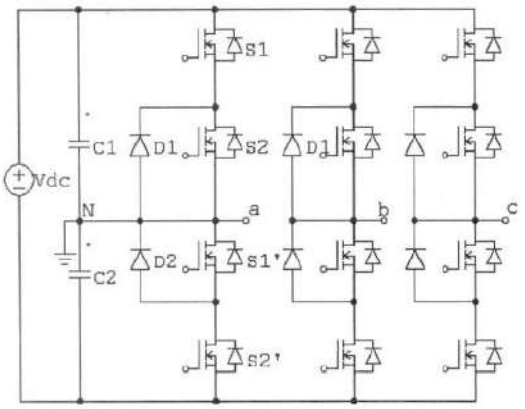


Fig.1.24. Diode-clamped multilevel inverter.[53]

It has different features like, high-voltage rating required or blocking diodes, unequal device rating and capacitor voltage unbalance.

Advantages:

a. High efficiency.

b. No need of filters to reduce harmonics.

c. Reactive power flow can be controlled.

Disadvantages:

a. For high levels, more number of diodes are required.

b. Real power flow control for individual converter is difficult.

• Flying Capacitors Multilevel Inverter

Figure 7 shows the flying capacitors based multilevel inverter topology. It requires (m-1) capacitors on dc bus form level converter.

Advantages:

a. Extra ride through capability during power outage.

b. No need of filters to reduce harmonics.

c. It gives proper switching combination to balance different voltage levels.

d. Real and reactive power flow can be controlled.

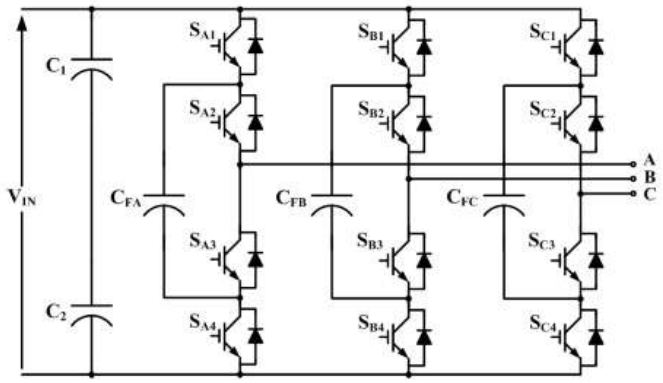


Fig.1.25. Three-level flying capacitor multilevel inverter topology

7-Multicellular inverter Topology

The general scheme for three-phase multicellular converter is illustrated in Fig. 1.26 The multilevel inverter consists of pairs of semiconductor switches separated by floating capacitors. The two switches in each pair must always be complementary in order to avoid shorting the voltage sources. Each pair of switches represents a switching cell. The principle of this topology is to divide the DC bus voltage into several basic voltage sources. The operation of each switching cell is similar to a two-level inverter with a voltage source equal to v/N (N is the number of cells and v_{dc} dc is the supply voltage) and a current source. The maximum voltage of the IGBTs switching are achieved by Vmax= v/N. The first advantage of these converters is the reduced volt-dc age requirements on the switches. It is necessary to identify all the converter possible states, the voltage across the floating capacitors and the converter's output voltage level for all states (equal to v j /N, j = 1,...N). Multicellular series converters also improve the waveform of the output voltage and allow greater flexibility for different voltage levels as compared to the NPC structure.

However, These topologies have received more interest because of the several advantage such as small dv/dt stress, small rating of power switches, can be used in high dc voltage application low power loss, low voltage stress in power switches and the dc side voltage being divided on flying capacitors naturally [54-55].

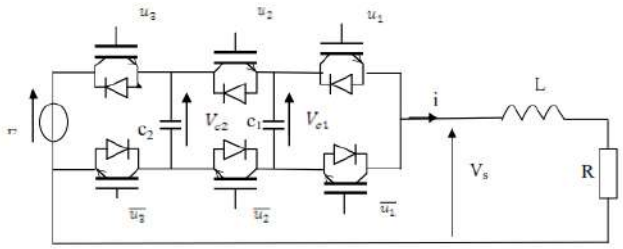


Figure 1.26:. Multicellular arm having 3 switching cells

Table 1 States of 4-level inverter and its output voltage

States	S_3	S_2	S_1	Output voltage	Level
1	0	0	0	0	1
2	0	0	1	$(1/3)v_{dc}$	2
3	0	1	0	$(1/3)v_{dc}$	2
4	0	1	1	$(2/3)v_{dc}$	3
5	1	0	0	$(1/3)v_{dc}$	2
6	1	0	1	$(2/3)v_{dc}$	3
7	1	1	0	$(2/3)v_{dc}$	3
8	1	1	1	v_{dc}	4

7.1-The Modeling of Multicellular inverter

The power converter, used in this work, has three cell-multicellular topologies which includes floating capacitors. The modeling of this type of converter involves the mathematical equations of floating capacitors. In the latter, the instantaneous variations of current and voltage are given by equations 1.11 and 1.1 2.

$$i_{Ci} = C_i \frac{d}{dt} V_{Ci} \tag{1.11}$$

$$\frac{d}{dt} V_{Ci} = \frac{1}{C_i} [S_{(i+1)} - S_i] i_L \tag{1.12}$$

S_i : the switching functions of the multicellular converter.

i_{Ci} : Current of flying capacitor C_i i=1,2

V_{Ci} : Voltage of flying capacitor C_i

i_L : Load current.

The application of Kirchhoff's law from the mid-point point of Vdc and three-phase load gives the following equations.

$$L_L \frac{di_L}{dt} = V_S - R_L i_L - \frac{V_{dc}}{2} \tag{1.13}$$

$$V_S = S_1[V_{C1}] + S_2[V_{C2} - V_{C1}] + S_3[V_{dc} - V_{C2}] \tag{1.14}$$

$$\frac{di_L}{dt} = \frac{1}{L_L}(S_1[V_{C1}] + S_2[V_{C2} - V_{C1}] + S_3[V_{dc} - V_{C2}]) - R_L i_L - \frac{V_{dc}}{2L_L} \tag{1.15}$$

With:

Switching functions S1, S2, and S3 are the input of multicellular converter.

Flying capacitors C_1 and C_2 are traversed by a currents i_{C1}, i_{C2} and the voltages at their terminals are V_{C1}, V_{C2} respectively.

V_{dc}: DC voltage source.

R_L: load resistor

L_L: Load inductor

The nonlinear model of multicellular converter topology used in the photovoltaic system is given by equation 1.16.

$$\begin{bmatrix} \dot{V}_{C1} \\ \dot{V}_{C2} \\ \dot{i}_L \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \dfrac{-R_L}{L_L} \end{bmatrix} \begin{bmatrix} V_{C1} \\ V_{C2} \\ i_L \end{bmatrix} + \begin{bmatrix} \dfrac{-i_L}{C_1} & \dfrac{i_L}{C_1} & 0 \\ 0 & \dfrac{-i_L}{C_2} & \dfrac{i_L}{C_2} \\ \dfrac{V_{C1}}{L_L} & \dfrac{V_{C2} - V_{C1}}{L_L} & \dfrac{V_{dc} - V_{C2}}{L_L} \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ S_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \dfrac{-V_{dc}}{2L_L} \end{bmatrix} \tag{1.16}$$

The state space of proposed topology can be expressed by the following equation:

$$\dot{x} = f\,(x) + g\,(x)u + H \qquad (1.17)$$

With,

State vector is expressed by x=[V_{C1}, V_{C2}, i_L]T . the reference vector is

x_{ref}=[$\frac{V_{dc}}{3}$, $\frac{2V_{dc}}{3}$, i_L]^T.

$$f(x)=\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & \frac{-R_L}{L_L}\end{bmatrix},\ g(x)=\begin{bmatrix}\frac{-i_L}{C_1} & \frac{i_L}{C_1} & 0\\ 0 & \frac{-i_L}{C_2} & \frac{i_L}{C_2}\\ \frac{V_{C1}}{L_L} & \frac{V_{C2}-V_{C1}}{L_L} & \frac{V_{dc}-V_{C2}}{L_L}\end{bmatrix},\ u=[S_1 S_2\ S_3]^T \text{ input vector,}$$

$$H=\begin{bmatrix}0\\ 0\\ \frac{-V_{dc}}{2L_L}\end{bmatrix} \text{ Constant vector.}$$

1.14.Conclusion

In this part, we presented detailed types of the photovoltaic systems and its components, and in the last we focused on the use of multicellular topology which the main object of the study.

Chapter 2

Control of Multicellular Power inverter

2.1 Introduction

A nonlinear controlled system is a set of nonlinear equations describing the temporal evolution of the constituent variables of the system under the action of a finite number of independent variables called inputs or control variables [56], The theory of nonlinear control has more than one asset allowing its application in the analysis and the control of the various physical systems. Among other things, there are two different ways, often more efficient are exact linearization and sliding mode control. The first consists in introducing before and around the controlled system non-linear organs so that the whole is linear then to apply on this new system one of the policies linear. The drag control brings the point representing the state of the system to a surface over which this state slides in a linear behavior and moves towards the point desired balance. The chapter ends with the notion of passivity and its applications in the control of nonlinear systems [57].

2.2. Modeling of the three-cell multicellular inverter

The instantaneous value model representing one phase of the N-cell multi-cell inverter supplying any load is given by the following system of equations:

$$\begin{cases} \dfrac{d}{dt}v_{C1} = \dfrac{1}{C}[S_2 - S_1]i_f \\ \dfrac{d}{dt}v_{C2} = \dfrac{1}{C}[S_3 - S_2]i_f \\ \vdots \\ \dfrac{d}{dt}v_{CN-1} = \dfrac{1}{C}[S_N - S_{N-1}]i_f \\ \dfrac{d}{dt}i_f = -\dfrac{[S_2 - S_1]}{L_f}v_{C1} - \dfrac{[S_3 - S_2]}{L_f}v_{C2} - \cdots - \dfrac{[S_N - S_{N-1}]}{L_f}v_{CN-1} \\ \qquad -\dfrac{R_f}{L_f}i_f + \dfrac{S_N}{L_f}V_{dc} - \dfrac{V_{dc}}{2L_f} - \dfrac{v_s}{L_f} \\ i_f = i_{ch} - i_s \end{cases} \quad 2\text{-}1$$

$$\begin{bmatrix} \dot{v_{C1}} \\ \dot{v_{C2}} \\ \dot{i_f} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{-R_f}{L_f} \end{bmatrix} \begin{bmatrix} v_{C1} \\ v_{C2} \\ i_f \end{bmatrix} + \begin{bmatrix} \frac{-(i_{ch}-i_s)}{C} & \frac{(i_{ch}-i_s)}{C} & 0 \\ 0 & \frac{-(i_{ch}-i_s)}{C} & \frac{(i_{ch}-i_s)}{C} \\ \frac{v_{C1}}{L_f} & \frac{v_{C2}-v_{C1}}{L_f} & \frac{V_{dc}-v_{C2}}{L_f} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{-V_{dc}}{2L_f} - \frac{v_s}{L_f} \end{bmatrix} \quad 2\text{-}2$$

The matrix state representation at the instantaneous values of one phase of the three-cell multi-cellular inverter applied to the multicellular inverter is given by equation (2- 1).

From the nonlinear matrix form, we can use the nonlinear controls. In this chapter, we use the control by sliding mode then the control by exact linearization

2.2.1 Sliding mode control

1 Definition

Conventional control laws of the PID type are very effective in the case of systems linear with constant parameters. For nonlinear systems or systems having parameters not constants, these control laws can be insufficient because they are not robust especially when requirements on accuracy and other dynamic characteristics of the system are strict [58]. it is necessary to use control laws insensitive to the variations of the parameters control laws insensitive to parameter variations, disturbances and non-linearity. The control laws known as variable structure (CSV) also known as control by sliding mode constitute a solution to these problems [59]. In practice, the use of this control technique has long been limited by the oscillations linked to the switching of the command and which can occur on the controlled quantities. Since then, many solutions have been proposed to reduce these oscillations: increase in frequency switching, continuous control in a band around the sliding variety or breakdown of the command into a low-frequency DC component and a discontinuous high frequency control and the boundary layer method where the component discontinuous of the command is replaced in the vicinity of the hyper surface of sliding by a continuous function. In this

chapter we will give some basic concepts of variable structure control, and some basic concepts of sliding mode theory [58].

2. Principle of sliding modes control.

The principle of control by sliding modes and constraining the system to achieve the given surface called the sliding surface, depending on the purposes of the control fixing the dynamics in closed loop: it is the mode of convergence. Then by synthesis a discontinuous control that allows the trajectories of the system to reach and stay on this surface: it is the sliding surface and the movement that occurs along it is called sliding moves.

The trajectory in the phase plane consists of three distinct parts[60]:

1- **The "CM" convergence mode:** This is the mode during which the variable to be regulated moves from any initial point in the phase plane and tends towards the switching surface S(x)=0 . This mode is characterized by the control law and the convergence criterion.

2- The "SM" sliding mode: This is the mode during which the state variable has reached the sliding surface and tends towards the origin of the phase plane. The dynamics of this mode is characterized by the choice of the sliding surface S(x)=0.

3- The steady state mode "SSM": This mode is added for the study of the response of the system around its point of equilibrium (origin of the phase plane), it is characterized by the quality and performances of control.

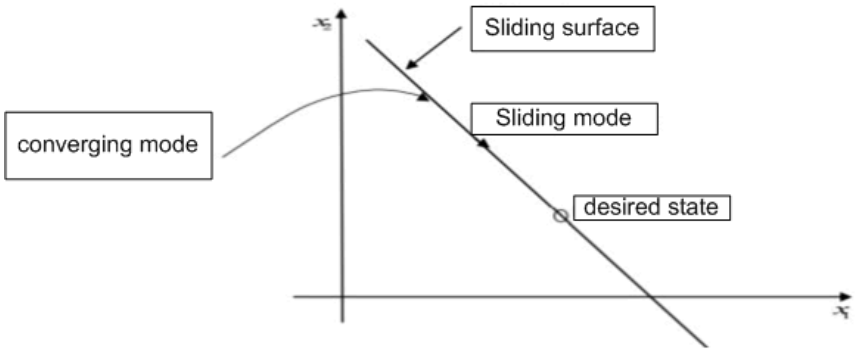


Fig.2.1: Different convergence modes for state trajectory

Implementation of the control law can be carried out in three main steps very dependent on each other [61]:

- The choice of surface.

- The establishment of the conditions for the existence of convergence.

- The determination of the control law.

3. Choice of sliding surface.

In order to ensure the convergence of a state variable x towards its reference value xref , different shapes of the sliding surface have been proposed, each surface has different best performance for a given application. In general, we choose a surface not linear.

The nonlinear form is a function of the error on the controlled variable, denoted e(x). She is given by the equation:

$$s(x) = (\frac{d}{dt} + \gamma)^{r-1}\tilde{x} \qquad 2\text{-}3$$

Consider a class of nonlinear systems the surface vector has the same dimension than the control vector u.

$$\dot{x} = f(x,t) + g(x,t).u \qquad 2\text{-}4$$

e(x) = x_{ref}- x. is the error between the controlled variable x and its reference $xref$

λ: is a positive constant.

r: is the relative degree.

4. Convergence and existence conditions.

The conditions of existence and convergence are the criteria that allow the different dynamics of the system to converge towards the sliding surface and stay there, regardless of the disturbance. There are two considerations for ensuring the mode of convergence [62].

1. The discrete commutation function: It gives the convergent surface dynamics towards zero.

$$\begin{cases} \dot{S}(x) > 0 & if\ S(x,t) < 0 \\ \dot{S}(x) < 0 & if\ S(x,t) > 0 \end{cases} \quad 2.5$$

$$\dot{S}(x).S(x) < 0 \quad 2.6$$

2. The Lyapunov function: The Lyapunov function is a positive scalar function V(x)>0 for system state variables. The control law must decrease this function V(x) < 0.

We define the Lyapunov function as follows:

$$V(x) = \frac{1}{2}S^T(x).S(x) \quad 2.7$$

The derivative of this function is:

$$\dot{V}(x) = \frac{1}{2}S(x).\dot{S}(x) \quad 2.8$$

5 Determination of the control law.

It is necessary to determine the command necessary to attract the state trajectory towards the surface of sliding S(x) =0 and then towards its point of equilibrium while maintaining the conditions existence of the sliding mode [63].

In the presence of a disturbance, the discontinuous part is essentially intended to check the conditions of attractiveness. In this case, the structure of a controller by sliding mode S(x) consists of two parts equivalent control (u_{eq}) and discrete control (u_n) We write:

$$u = u_{eq} + u_n \qquad 2.9$$

Such as:

Ueq: The control vector proposed by Utkin, it is used to maintain the variable to be controlled on the sliding surface S(x)=0, the equivalent control vector is deduced by considering that the derivative of the surface is zero $S(x)$ =0. It can be interpreted as a status return individual playing the role of a control signal applied to the system to be controlled, it can be like an average value that the control takes when switching fast between the values $umax$ and $umin$.

Un: is determined to check the convergence condition.

To highlight the computation of the order, we consider a system defined in the state space by the equation (II.10). This is to find the expression of the control vector U:

$$\dot{S}(x) = \frac{dS}{dt} = \frac{dS}{dx}\frac{dx}{dt} \qquad 2.10$$

This gives the following result:

$$\dot{s}(x) = \frac{ds}{dt} = \frac{ds}{dx}\left[f(x) + g(x)u_{eq}\right] + \frac{ds}{dx}g(x)u_n \qquad 2\text{-}11$$

During the sliding mode the sliding surface is null, then, its derivative and the discontinuous part are also null. From where, one deduces the expression of the equivalent control vector.

$$u_{eq} = -\left[\frac{ds}{dx}g(x)\right]^{-1}\left[\frac{ds}{dx}f(x)\right] \qquad 2.12$$

For the equivalent control to take a finite value, the matrix system must be reversible:

$$\frac{ds}{dx}g(x) \neq 0 \qquad 2.13$$

By replacing the equivalent control vector by its expression in equation 2.9, we obtain the new expression of the derivative of the surface:

$$\dot{s}(x) = \frac{ds}{dx}g(x)u_n \qquad 2\text{-}14$$

And the attractiveness condition $S(x).\dot{S}(x) < 0$ becomes:

$$S(x)\frac{ds}{dx}g(x)u_n < 0 \qquad 2.15$$

To satisfy the attractiveness condition (the derivative of the sliding surface is negative).

The sign of U_n must be opposite that of $S(x)\frac{ds}{dx}g(x)$.

The simplest form of the discrete command is the sign function.

$$u_n = k_x sign\, S(x) \qquad 2\text{-}16$$

The sign of kx must be different from that of $\frac{ds}{dx} g(x)$

Fig. 2.2 represents the sign function

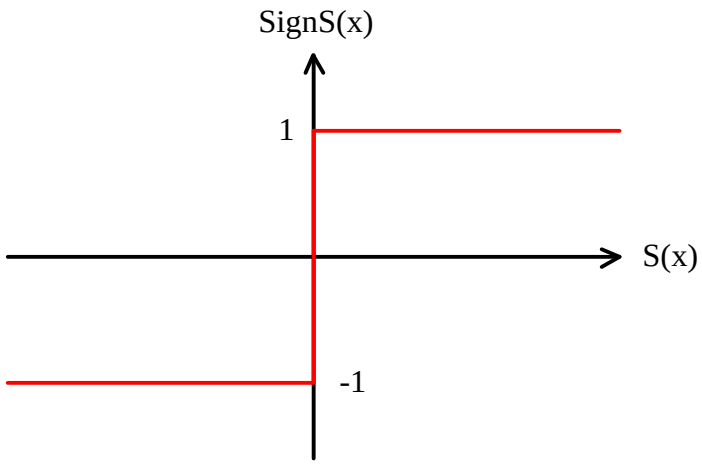


Fig. 2.2. Representation of the sign function.

6. Sliding mode control of a multicellular converter:

Nonlinear state representation of the three-phase multicellular inverter with three cells is in the form:

$$\dot{x} = Ax + B(x)u + H \qquad 2\text{-}17$$

Choose the sliding surface as follows:

$$S(x) = x - x_{ref} \quad 2\text{-}18$$

With :

$x = [v_{C1}, v_{C2}, i_f]^T$ is the state vector, $X_{ref} = \left[\frac{V_{dc}}{3}, \frac{2V_{dc}}{3}, i_{f-ref}\right]^T$ is the reference vector.

And to verify the convergence condition, choosing the Lyapunov function as follows:

$$V = \frac{1}{2}S^2(x) \quad 2\text{-}19$$

$$\dot{V} = S(x)\dot{S}(x) \quad 2\text{-}20$$

The derivative of the sliding surface:

$$\dot{S}(x) = \dot{x} - \dot{x}_{ref} \quad 2\text{-}21$$

By replacing the state form of the multicellular inverter in (II.22):

$$\dot{S}(x) = A(x) + B(x)u + H - \dot{x}_{ref} \quad 2\text{-}22$$

The command is equivalent for the system to slide on the sliding surface where the derivative of the sliding surface is zero:

$$u_{eq} = -\big(B(x)\big)^{-1}\big(A(x) + H - \dot{x}_{ref}\big) \quad 2\text{-}23$$

The total control vector U is the sum of the two control vectors, equivalent and discontinuous as the equation (2-9) indicates it

By replacing equations (2-9) and (2-23) in equation (2-22), we obtain the equation of the derivative of the following sliding surface:

$$\dot{S}(x) = B(x)u_n \tag{2-24}$$

Then, the derivative of the Lyapunov function which must be negative is given by:

$$S(x)\dot{S}(x) = S(x)B(x)u_n < 0 \tag{2-25}$$

For the state representation of the multicellular converter

$$[u_{n1}u_{n2}u_{n3}]^T \tag{2-26}$$

$$\begin{aligned} S(x)\dot{S}(x) = S(x)\Bigg[&\left(\frac{-(i_{ch}-i_s)}{C}+\frac{v_{c1}}{L_f}\right)u_{n1} \\ &+\left(\frac{(i_{ch}-i_s)}{C}-\frac{(i_{ch}-i_s)}{C}+\frac{v_{c2}-v_{c1}}{L_f}\right)u_{n2} \\ &+\left(\frac{(i_{ch}-i_s)}{C}+\frac{(v_{dc}-v_{c2})}{L_f}\right)u_{n3}\Bigg] \end{aligned} \tag{2-27}$$

To ensure stability according to Lyapunov's theorem we need $\dot{V}(x) < 0$

So

$$u_{n1} = -sign\left[S(x)\left(\frac{-(i_{ch} - i_s)}{C} + \frac{v_{c1}}{L_f}\right)\right] \quad 2\text{-}28$$

$$u_{n2} = -sign\left[S(x)\left(\frac{(i_{ch} - i_s)}{C} - \frac{(i_{ch} - i_s)}{C} + \frac{v_{c2} - v_{c1}}{L_f}\right)\right] \quad 2.29$$

$$u_{n3} = -sign\left[S(x)\left(\frac{(i_{ch} - i_s)}{C} + \frac{(v_{dc} - v_{c2})}{L_f}\right)\right] \quad 2.30$$

$$u_{n1} = -sign\left[S(x)\left(\frac{-(i_{ch} - i_s)}{C} + \frac{v_{c1}}{L_f}\right)\right] \quad 2\text{-}31$$

$$u_{n2} = -sign\left[S(x)\left(\frac{(i_{ch} - i_s)}{C} - \frac{(i_{ch} - i_s)}{C} + \frac{v_{c2} - v_{c1}}{L_f}\right)\right] \quad 2.32$$

$$u_{n3} = -sign\left[S(x)\left(\frac{(i_{ch} - i_s)}{C} + \frac{(v_{dc} - v_{c2})}{L_f}\right)\right] \quad 2.33$$

2.3 Exact linearization Control mode.

1. Definition.

The linear method is one of the nonlinear control methods, Its mechanism of operation is to algebraically transform a nonlinear system into a system linear, so that linear control

techniques can be applied, After this transformation, all linear correction techniques are used [64] [65].

The main idea of this method is to make an accurate conversion without going through estimates.

2. Regulation loop and mathematical equations of a multicellular converter

To know the results of this method, we apply it to the inverter three-celled multicellular.

Consider the nonlinear multi-input multi-output system presented in the equation

$$\begin{cases} \dot{x} = f(x) + g(x)u + H \\ y = d(x) \end{cases} \qquad 2.34$$

$[x_1,\ x_2,\ x_3, \dots, x_N]^T$ is the state vector.

$[y_1,\ y_2,\ y_3, \dots, y_N]^T$ is the output vector.

$[u_1,\ u_2,\ u_3, \dots, u_N]^T$ is the input vector.

$[H_1,\ H_2,\ H_3, \dots, H_N]^T$ is the disturbing component.

And whose vector representations are:

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}, g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_n(x) \end{pmatrix}, d(x) = \begin{pmatrix} d_1(x) \\ \vdots \\ d_n(x) \end{pmatrix}$$

The functions f, g and d are considered as smooth functions which are indefinitely derivable with respect to each of their arguments

The idea of exact linearization is to complete the system by introducing a new command w such that $u = R(x, w)$.

To perform this linear operation, successive derivatives are expressed with the symbol yi, when the inputs start to interfere with the expression of the derivative, we stop the differentiation, we set the following equation

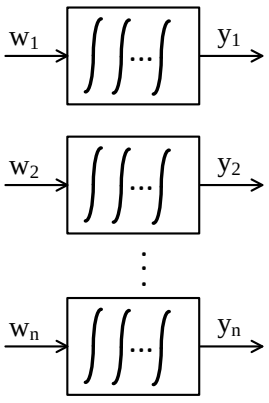
$$\begin{pmatrix} y_1^{(r1)} \\ \vdots \\ y_m^{(rm)} \end{pmatrix} = \Delta(x)u + \Delta_0(x) + H \qquad 2.35$$

Where the new input w is given by the equation

$$\begin{cases} y_1^{r1} = w_1 \\ \vdots \\ y_n^{rn} = w_n \end{cases} \qquad 2.36$$

Where ri denotes the smallest derivative so that the entries u appear in the $rième$ derivative of the output yi.

The linearized system is given by:



Using the matrix property $\Delta(x)$ is invertible, the following loops are performed

$$u(x) = \Delta^{-1}(x)w - \Delta^{-1}(x)\Delta_0(x) - \Delta^{-1}(x)H \qquad 2.38$$

$$u(x) = \Delta^{-1}(x)[w - H] - \Delta^{-1}(x)\Delta_0(x) \quad 2.39$$

Where w is the new entry to make the system (2-31) linear

We can write the equation (2.36) in the following form

$$u(x) = \alpha(x) + \beta(x)[w - H] \quad 2\text{-}40$$

With:

$$\alpha(x) = -\Delta^{-1}(x)\Delta_0(x) \quad 2\text{-}41$$

$$\beta(x) = \Delta^{-1}(x) \quad 2.42$$

The system represented by equation (2.31) is linear and completely decoupled with the new order w and is easily orderable using classical techniques of the output return command.

The overall scheme of this system with the control vector w is given in the following figure

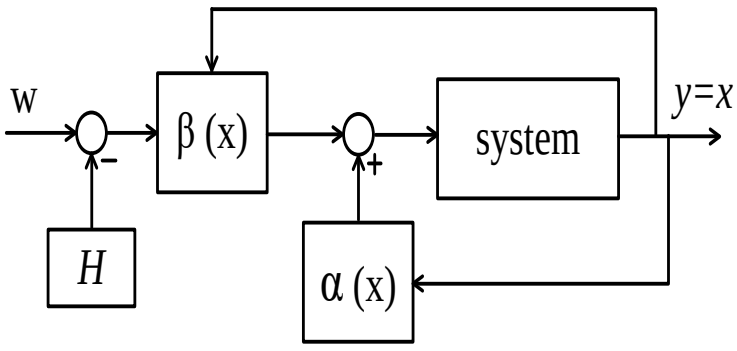


Fig. 2.3. Linearization of the nonlinear system.

3. Control by exact linearization of a three-phase multicellular inverter.

We will study a model with average values with the exact linearization on an inverter three-phase three-cell multicellular with an R-L load, the condition is given by the equations

$$\begin{bmatrix} \dot{V_{c1}} \\ \dot{V_{c2}} \\ \dot{I_{ch}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{-R_{ch}}{L_{ch}} \end{bmatrix} \begin{bmatrix} V_{c1} \\ V_{c2} \\ I_{ch} \end{bmatrix} + \begin{bmatrix} \frac{I_{ch}}{C} & 0 & 0 \\ 0 & \frac{\mathrm{I}_{ch}}{C} & 0 \\ \frac{V_{c1}}{L_{ch}} & \frac{V_{c2}}{L_{ch}} & \frac{V_{dc}}{L_{ch}} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{-V_{dc}}{2L_{ch}} \end{bmatrix} \qquad 2.42$$

The state vector x of the system of each phase is therefore of order three with two voltages of the floating capacitors and a load current.

$$x_1 = V_{C1}, x_2 = V_{C2}, x_3 = I_{ch},\ a = \frac{1}{C},\ b_0 = \frac{R_{ch}}{L_{ch}}, b_1 = \frac{1}{R_{ch}}$$

For the decoupling of the system (II-42) to be possible, the matrix $\Delta(x)$ must be invertible, then, the delt $\Delta(x) = a^2 b_1 V_{dc} {x_3}^2$ is deferent from zero.

For the matrix $\Delta(x)$ to be invertible it is necessary that $V_{dc} \neq 0$ et $x_3 \neq 0$

$$\Delta_0(x) = \begin{pmatrix} 0 \\ 0 \\ -b_0 x_3 \end{pmatrix} \qquad 2.43$$

$$\Delta^{-1}(x) = \begin{pmatrix} \frac{x_1 - V_{dc}}{aV_{dc}x_3} & \frac{x_2 - V_{dc}}{aV_{dc}x_3} & \frac{1}{b_1 V_{dc}} \\ \frac{x_1}{aV_{dc}x_3} & \frac{x_2 - E}{aV_{dc}x_3} & \frac{1}{b_1 V_{dc}} \\ \frac{x_1}{aV_{dc}x_3} & \frac{x_2}{aV_{dc}x_3} & \frac{1}{b_1 V_{dc}} \end{pmatrix} \qquad 2.44$$

The state feedback can therefore be expressed by

$$\alpha(x) = -\Delta^{-1}(x)\Delta_0(x) = \begin{pmatrix} \frac{b_0 x_3}{b_1 V_{dc}} \\ \frac{b_0 x_3}{b_1 V_{dc}} \\ \frac{b_0 x_3}{b_1 V_{dc}} \end{pmatrix} \qquad 2.45$$

By applying the state feedback to the system, we obtain the following linearized system

$$\begin{cases} \dot{y}_1 = \dot{V}_{c1} = w_1 \\ \dot{y}_2 = \dot{V}_{c2} = w_2 \\ \dot{y}_3 = \dot{I_{ch}} = w_3 \end{cases}$$

The nonlinear form $\dot{x} = Ax + B(x)u + H$

With :

$x = [v_{C1}, v_{C2}, i_f]^T$ is the state vector, $X_{ref} = \left[\frac{V_{dc}}{3}, \frac{2V_{dc}}{3}, i_{f-ref}\right]^T$ is the reference vector.

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{-R_{fk}}{L_{fk}} \end{bmatrix}$$

$$B(x) = \begin{bmatrix} \frac{-(i_{ch}-i_s)}{C} & \frac{(i_{ch}-i_s)}{C} & S0 \\ 0 & \frac{-(i_{ch}-i_s)}{C} & \frac{(i_{ch}-i_s)}{C} \\ \frac{v_{C1}}{L_f} & \frac{v_{C2}-v_{C1}}{L_f} & \frac{v_{dc}-v_{C2}}{L_f} \end{bmatrix}$$

$u = [S_1, S_2, S_3]^T$ is the control vector, $H = \begin{bmatrix} 0 \\ 0 \\ \frac{-v_{dc}}{2L_f} - \frac{v_s}{L_f} \end{bmatrix}$ is the constant vector..

The diagram of the control by exact linearization is given in figure 2.4

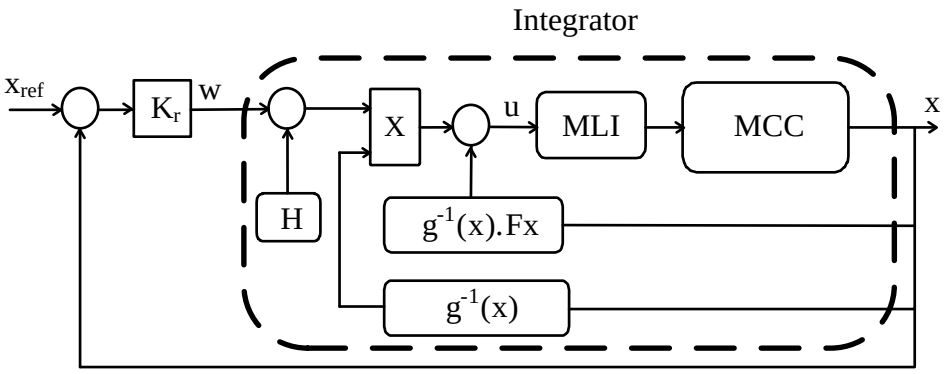


Fig .2.4 Exact linearization of the multicellular inverter.

2.4 Simulation of a multicellular inverter with the exact linearization control.

The simulation parameters are:

Floating capacitors	$4\mu F$
DC bus voltage	$2000V$
Inductor load	$0.4mH$
Resistor load	8Ω

The exact linearization control of the multicellular converter with MATLAB /Simulink is presented in the figure (2.5 and 2.6)

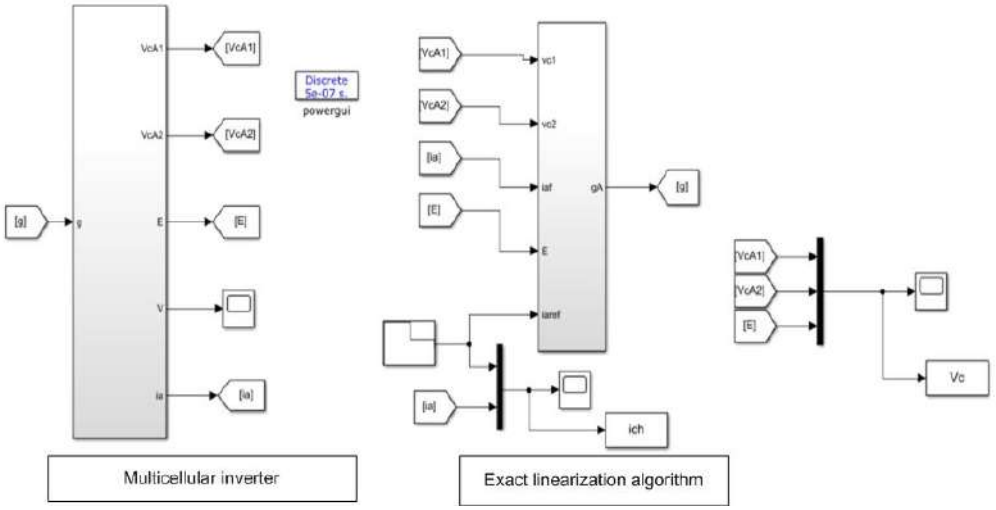


Fig. 2.5 Simulink of control by exact linearization of the multicellular inverter with Matlab

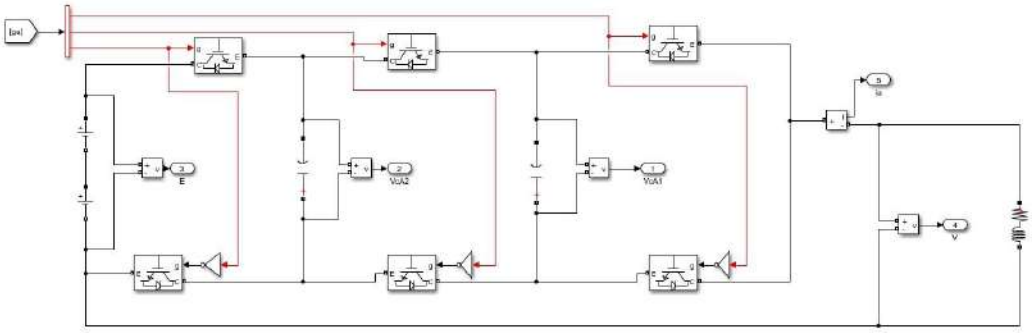


Fig. 2.6 Scheme of control of the multicellular converter in Matlab

Figure 2.7 presents the voltages of the capacitors and the supply voltage Vdc of the multicellular converter. When the load current is varied, the voltages of the capacitors present a ripple author of its reference.

In figure 2.8, the load current always follows the reference in the case of the variation of the latter. This proves the robustness of the control by exact linearization against the variation of the reference load current.

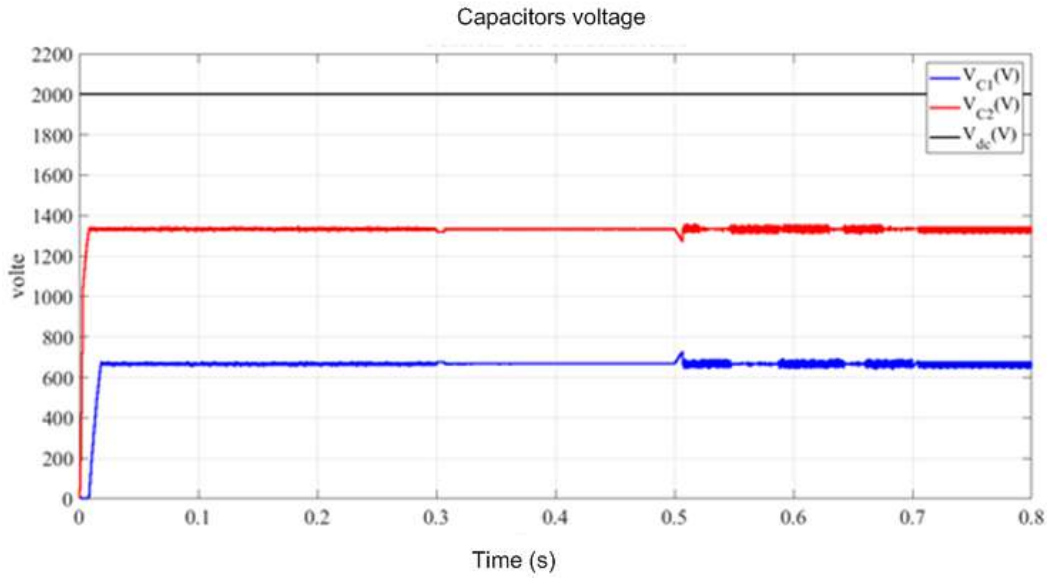


Figure 2.7:the voltages of the capacitors and the supply voltage Vdc of the multicellular converter.

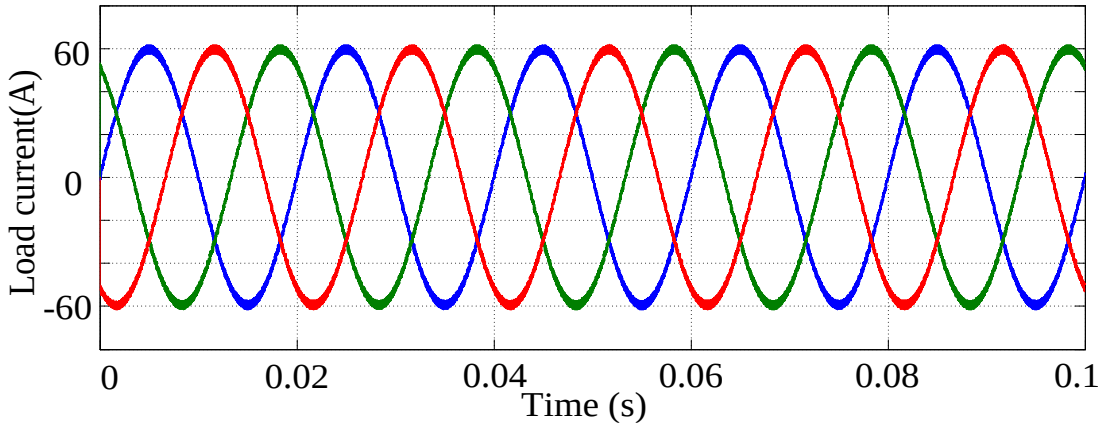


Fig. 2.8 Three-cell three-phase multi-cellular inverter load current for exact linearization control

2.4.Simulation of a multicellular inverter with the sliding mode control.

Figure 2.9 presents the sliding mode control of the multicellular inverter. In the sliding mode control of the three-cell multi-cellular inverter, the simulation parameters are the same used in the exact linearization control.

Figure 2.11 shows that the capacitor voltages V_{C1} and V_{C2} follow their references during steady state after a transient state of 0.03s with reduced ripple when the load current change. The current of the load i_{ch} . In the following figure 2.10 always the reference in the case of the

variation of the latter with a. This proves the robustness of the sliding mode control against the variation of the current of the load.

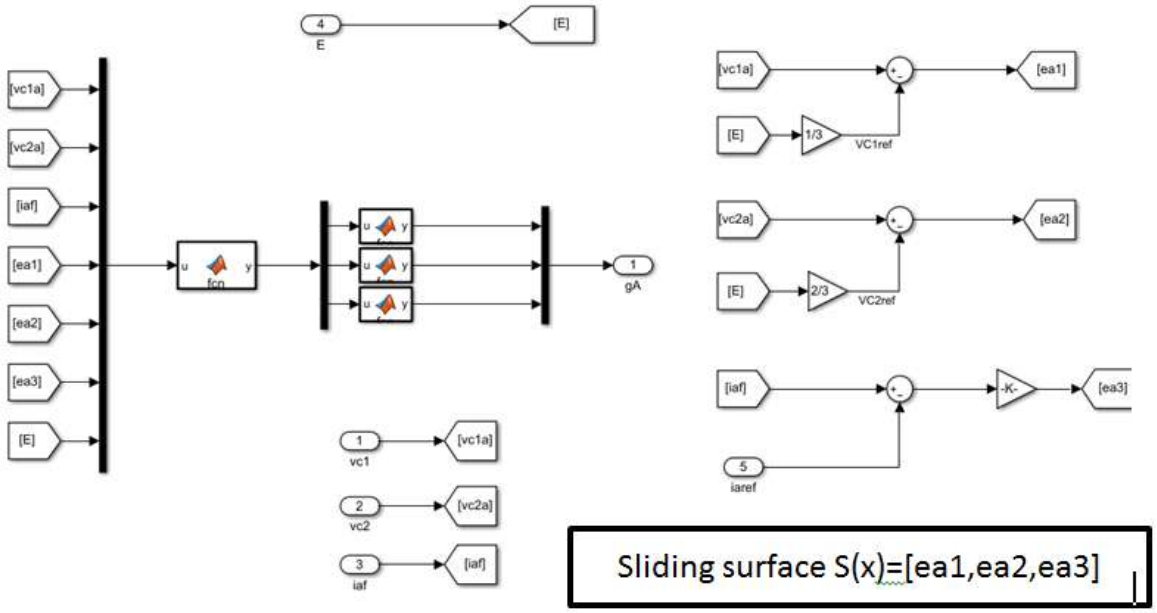


Fig. 2.9 Simulink the sliding mode control of the multicellular converter with Matlab

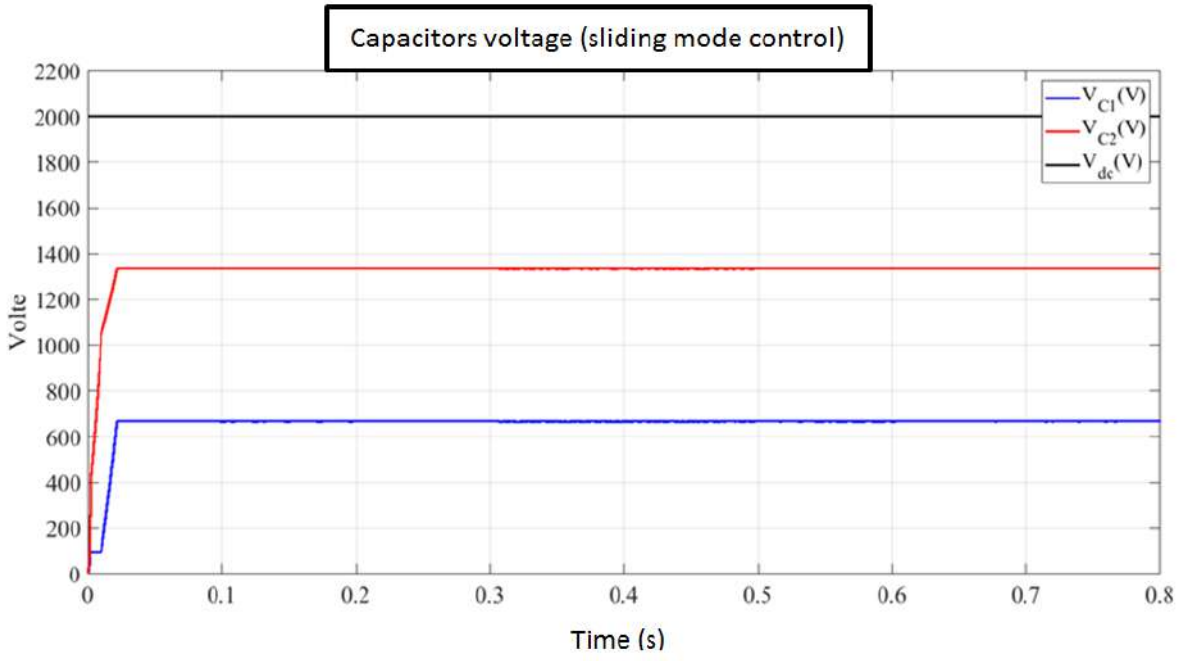


Figure 2.10: dc and capacitors voltages for sliding mode control

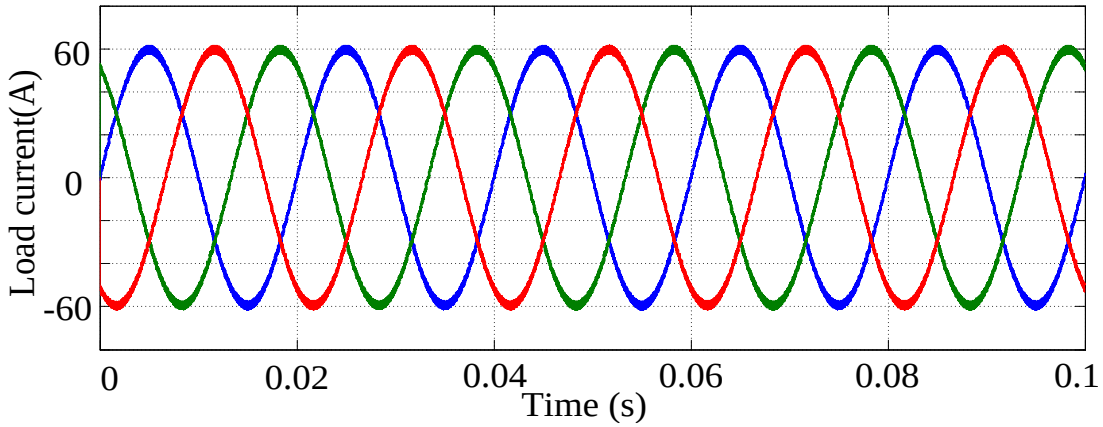


Figure. 2.11: sliding mode current load

2.5 Conclusion

The closed-loop nonlinear controls of the three-cell multicellular converter made it possible to minimize the duration of the transient state, and to minimize the ripple author of the reference value of the load current and capacitor voltages. The control by sliding mode presents a better robustness than the control by exact linearization.

Chapter 3

Machine Learning Fault Detection Methodologies.

Introduction

Diagnosing faults in electrical power systems has become a more important topic recently, Due to the integration of renewable energy sources with existing systems. Machine learning; It is the latest approach that draws attention to the applications of different power systems. In this context, the use of machine learning tools is very clear and logical to deal with Diagnostic challenges in these systems. In this chapter, we will show differently advanced technologies that can automatically identify patterns in data and classify this data into different categories, and then use the exposed patterns to predict future data.

3.2 Types of machine learning

Machine learning is usually divided into two main types.

3.2.1 Supervised learning

Supervised learning refers to a sort of machine learning model that can be trained on a set of samples where the desired outputs (or labels) are already known. The models learn from these later results and make modifications to their internal parameters to accommodate the input data. Once the model is well trained, it can make correct predictions approximately unseen or future data. There are two main applications of supervised learning: classification and regression.

Overview of the overall process:

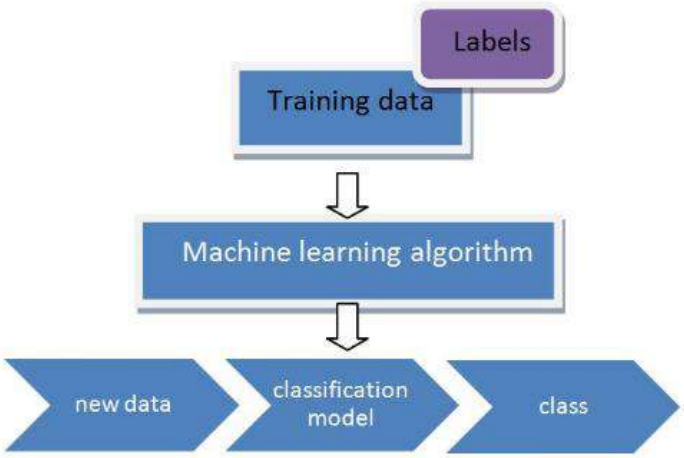


Figure 3.1: supervised learning process

3.2.2 Unsupervised learning

When the records used to train are neither classified nor labeled, unsupervised machine learning algorithms are used. Unsupervised learning investigates how systems can infer a function from unlabeled data to explain a hidden shape. The system does not recognize an appropriate output with certainty under any conditions. Instead, it draws inferences about what the output should be from datasets.

3.2.3 Semi-supervised learning

Semi-supervised learning, where the teacher gives an incomplete training signal: a training set with some (often many) of the target outputs missing. We will focus on unsupervised learning and data clustering in this blog post.

3.2.4 Reinforcement learning

Reinforcement learning employs a positive incentive system for correct behavior and a negative reward system for incorrect behavior. As a result, the technique assigns positive values to desired activities and negative values to unwanted acts in order to influence the agent. This instructs our agent to seek the largest overall reward over the long term in order to arrive at the best option. These long-term objectives prevent the agent from stopping there. The system eventually learns to avoid bad acts and only do positive ones. Trial and error is used to learn through interaction with the environment.

3.3 Classification

In Machine Learning and Statistics, Classification is the problem of identifying to which of a set of categories (subpopulations), a new observation belongs, on the basis of a training set of data containing observations and whose categories membership is known.

3.3.1 Types of Classification

Classification is of two types:

3.3.1.1 Binary Classification

When we have to categorize given data into 2 distinct classes. Example – On the basis of given health conditions of a person, we have to determine whether the person has a certain disease or not [66].

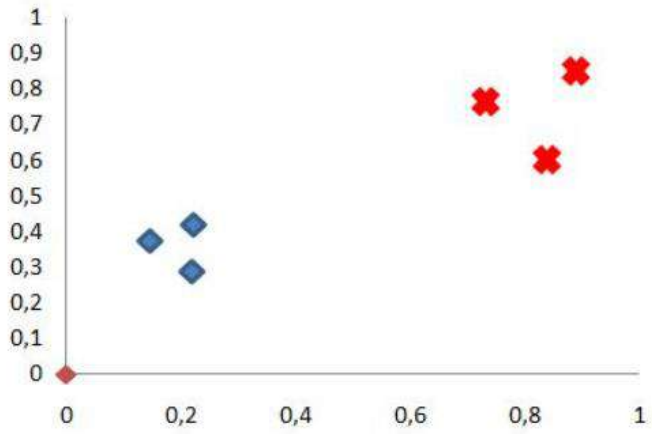


Figure 3.2: example of Binary classification representation

3.3.1.2 Multiclass Classification

The number of classes is more than 2. For Example – On the basis of data about different species of flowers, we have to determine which specie does our observation belongs to.

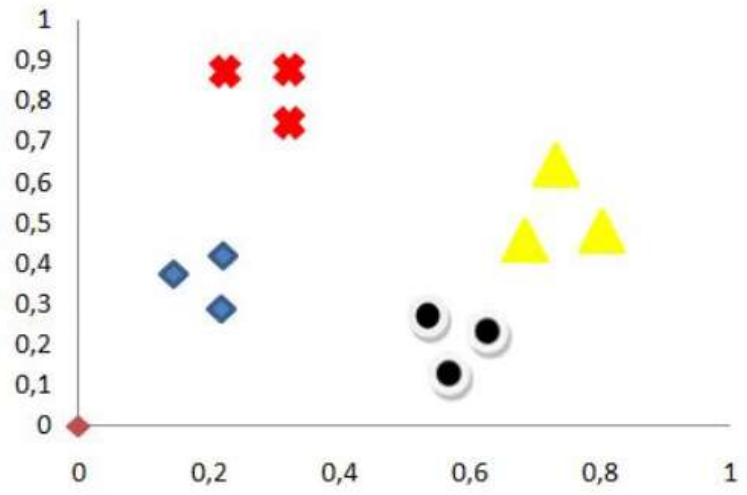


Figure 3.3: example of Multiclass classification representation

3.4-Dataset

3.4.1- Learning data in supervised and unsupervised learning

In supervised learning, experts classify the data and teach the model precisely what it needs to discover. For example, in the area of spam detection, the input is any text, and the label makes it apparent if the message is spam. Supervised learning is stronger since we don't allow the model draw its own inferences from data beyond the limitations labeled with our labels.

In unsupervised Learning, people feed the model raw, unlabeled data, and the model finds patterns in the data. For example, recognizing the level of similarity or difference between two data samples based on common features extracted. This helps the model make inferences and come to conclusions, such as separating similar images or grouping them into clusters.

3.4.1.1 Training Data

Training data refers to the initial data set passed to the machine learning model, on which the model is trained. Humans learn best from examples, and machines also need a set of data to learn patterns from it. In most cases, the training data contains input: annotation pairs collected from various sources that are used to train the model to perform a specific task with a high level of accuracy. They may consist of raw data (images, texts, or sound) containing annotations such as bounding boxes, labels, or links. Machine learning algorithms learn annotations from training data to handle new, unlabeled data in the future. All training

methods start with the collection of raw data from various sources. Raw data can be of any kind: text, images, sounds, videos, etc.

However, to tell the model what to look for in this data, we must add .These annotations help us to regulate learning by verifying that the model focuses on the features we specify, instead of making conclusions from other related (but not conditioned) data items.

Table 3.1: Data in supervised vs unsupervised learning

	Labeled data	Unlabeled data
Supervised data	x	
Unsupervised data		x
Hybrid model that includes supervised learning	x	x

All incoming data must have an appropriate label to allow the machine to move in the direction of what the forecast should look like. Such a processed dataset can be obtained using humans, and sometimes other ML models that are accurate enough for reliable labeling. Once the labeled dataset is ready to be passed to the AI, the training phase begins .On it, the model tries to identify important features that are common to all the examples that we have assigned labels. For example, if we segmented several cars in the imagery, then she will understand that wheels, rear-view mirrors and door handles are features that correlate with a car. The models continually test themselves on the validation data set produced before the training process. After the model is completed, the last check is performed on the test dataset (a set that the model has never seen before); this gives us insight into the performance of the model on relevant new examples.

The training data includes the training, validation, and test datasets. The more training data we have, the more accurate the model will be.

3.4.1.2 Labeled Data

Labeled data is data supplemented with labels/classes containing meaningful information. Here are some examples of labeled data: images tagged with cat/dog, emails/messages tagged as spam, stock market price predictions (labeled as future state), nodule malignancy with polygon highlighting, audio files with information about what words

are spoken in them. Accurately labeled data allows the machine to recognize patterns according to the task, so it is widely used in solving complex problems.

3.4.2 Datasets for training, validation and testing

No AI model can be trained and tested on the same data. The estimate of the model will be biased, because the model is being tested on what it already knows. This is analogous to giving students the same questions in an exam that they have already answered in class. This way we don't know if the student remembered the answers or really understood the topic. The same rules apply to machine learning models.

Here are their percentages of data volumes:

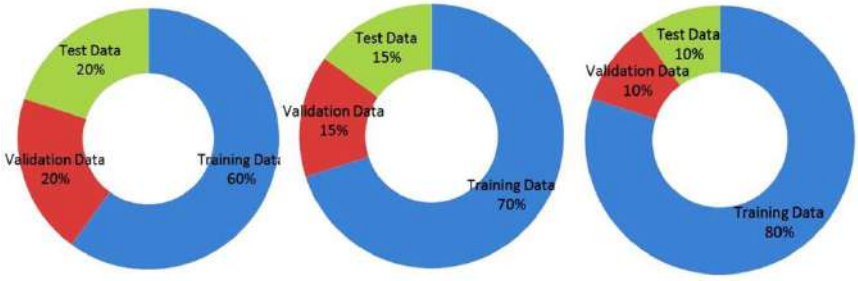


Figure3.4: training data needs

Training data - at least 60% of the data must be used for training. Validation data - a sample (10%-20%) of the total data set used for validation and periodically checked against the model during training. This validation dataset should be a representative sample of the training dataset. Test data - This dataset is used to test the model after it has been fully trained. It is separated from both the training set and the validation set. After training and validation, the model is tested on the test set. The data in the test set should look exactly like the real data will look after the model is deployed. There can be multiple test sets in a shared dataset each test set can be used to check if the model has trained enough for a specific application scenario.

3.4.3 Data Quantitative Requirements

We need to have at least 1000 training examples for every possible class in our scenario. If we utilize 10% of the data as our test set, we should be able to assess class correctness with a margin of error of at least 1%.For reference, 1000 examples is a sufficient data set. 10000 is a great data set, 100 thousand-1 million is an excellent data set. High-

quality models are trained on large amounts of training data, and for good reason – modern neural network architectures work great because they are able to store many weights (parameters) efficiently. However, if we don't have much training data, then we can only use a fraction of the potential of our model. The size of the data set also depends on the scope of our problem and the variance of each class. However, if we require a generic human recognizer then a collection of 10,000 instances will only represent just a portion of the differences in sizes, positions, look, and clothing styles. Therefore, a class with high variance, such as "human", requires much more training data.

3.4.4 Improving the Quality of Training Data

The high quality of dataset markup is required for the machine learning model to operate correctly. The term "qualitative data" refers to cleaned data that contains all the attributes that model training depends on. The consistency and correctness of the labeled data may be used to rate the quality .here is 4 characteristics of qualitative data for training ML models:

1. Relevance - The data set should contain only those features that provide meaningful information to the model. Identifying important features is a complex task that requires knowledge of the area and a clear understanding of which features should be considered and which should be eliminated.

2. Consistency - Similar examples should have similar labels, ensuring that the dataset is homogeneous.

3. Uniformity - The values of all attributes must be comparable across all data. Irregularities or the presence of outliers in datasets adversely affect the quality of training data.

4. Comprehensiveness - The dataset must include enough parameters or characteristics to ensure that no edge situations are missing. The dataset should contain enough samples of these edge cases so that the model can learn them as well.

3.4.5 Metrics in classification problems Metrics are used in machine learning tasks to measure the quality of models and compare different algorithms, and their selection and analysis is an essential element of a data scientist's job. We'll look at several quality criteria in classification issues and talk about what matters when picking a metric and what may go wrong.

3.4.5.1 Accuracy, precision, recall and specificity.

An important concept needs to be introduced to describe these metrics in terms of classification errors, the confusion matrix .Suppose we have two classes and an algorithm that predicts whether each object belongs to one of the classes, then the classification error matrix will look like this:

Here, is the response of the algorithm on the object, and is the true label of the class on this object. Thus, there are two types of classification errors: False Negative (FN) and False Positive (FP).

	$y = 1$	$y = 0$
$\hat{y} = 1$	True Positive (TP)	False Positive (FP)
$\hat{y} = 0$	False Negative (FN)	True Negative (TN)

Figure 3.5: the classification error matrix

3.4.5.2 Accuracy:

An intuitive, obvious and almost unused metric is accuracy - the proportion of correct answers of the algorithm:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.1}$$

This metric is useless in problems with unequal classes, and this is easy to show with an example. Let's say we want to evaluate the performance of a mail spam filter. We have 100 non-spam emails, 90 of which our classifier determined correctly (True Negative = 90, False Positive = 10), and 10 spam emails, 5 of which were also correctly determined by the classifier (True Positive = 5, False Negative = 5). Then accuracy:

$$accuracy = \frac{5 + 90}{5 + 90 + 10 + 5}$$

However, if we just predict all emails as non-spam, we get a higher accuracy:

$$accuracy = \frac{0 + 100}{0 + 100 + 0 + 10}$$

At the same time, our model does not have any predictive power at all, since initially we wanted to identify spam emails. The transition from a common metric for all classes to individual indicators of class quality will help us overcome this.

3.4.5.3 Precision, recall and F-measure

$$precision = \frac{TP}{TP + FP} \tag{3.2}$$

$$recall = \frac{TP}{TP + FN} \tag{3.3}$$

To assess the quality of the algorithm on each of the classes separately, we introduce the metrics precision (accuracy) and recall (completeness).

Precision can be interpreted as the proportion of objects called positive by the classifier and at the same time are really positive, and recall shows what proportion of objects of a positive class out of all objects of a positive class the algorithm found. It is the introduction of precision that does not allow us to write all objects into one class, since in this case we get an increase in the False Positive level. Recall illustrates the algorithm's capacity to recognize a particular class in general, while precision demonstrates the ability to differentiate this class from other classes.

3.4.6 Training Data Preparation

3.4.6.1 Data cleaning

Raw data can be very messy and corrupted in many ways. If not properly cleaned, they can skew results and cause the AI model to produce erroneous results. Data cleaning is the process of repairing or deleting inaccurate, corrupted, or duplicated data from a dataset.. The steps in the data cleaning process depend on the particular data set.

1. Check for duplicates - the same sample data may occur several times in a dataset. This may be generated by gathering data from several sources, resulting in comparable data. They need to be deleted since they may cause the model to over fit certain patterns and provide incorrect predictions.

2. Eliminate outliers - some parts of the data behave differently from the rest of the data. An example would be the Session ID, which is constantly found in the weblog data. This may be

due to some malicious activity that does not need to be passed to our model. As a result, monitoring emissions is one method of removing data that should not be supplied to the machine.

3. Fix Structural Errors - In certain conditions, the dataset may include faulty markup. For example, "Dog" and "dog" are regarded to be separate classes, although "doog" and "dog" are different according to a typo resulting in misclassification.

4. Check for missing values - there may be components in the dataset for which data examples are sorely missing attributes/features. We can overcome this issue by simply not include these components in the training dataset.

3.5.6.2 Data markup

Data markup is the process in which we assign a value to data in the form of a class or a label. Data labeling can be performed by employees, operators in the control loop, or any automated machine that speeds up the labeling process.

1. Establish the gold standard - Data scientists are considered the gold standard in data labeling, labeling raw data with maximum sensitivity and accuracy. Their markup is considered a guide for the annotation team and can be used as responses when screening annotation options.

2. Don’t use too many labels - splitting a dataset into a large number of classes can confuse employees when annotating it. In addition, to select among the many labels, the analysis of more features will be required. For example, it will be difficult for annotators to mark up data with classes such as "Very Expensive", "Expensive", "Less Expensive".

3. Use multiple passes - data markup should be done by multiple annotators. This is necessary to improve the overall quality of the data. While this is more time consuming and resource intensive, this approach is used to build consensus within the team.

4. Establish a validation system - to reduce the chance of errors, ready-made data markup should be verified by another person or through self-improvement checks. This will allow any annotator to understand where he can improve, his level of accuracy, and the kind of training needed to improve his work.

3.5.6.3 Standard deviation - Standard Deviation in simple terms, this is a measure of how scattered a set of data is. By calculating it, you can find out whether the numbers are close to the average value or far from it. If the data points are far from the mean, then there is a large variance in the data set .When the data is more spread out, the standard deviation increases.

Standard Deviation Formula :

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \mu)}{n-1}} \quad (3.4)$$

Where:

σ standard deviation.

Xi are different sample values,

μ is the arithmetic mean of the sample, n is the sample size.

3.5 The k Nearest Neighbors (kNN)

The k Nearest Neighbors (kNN) method is a popular classification algorithm that is used in various types of machine learning problems. Along with the decision tree, this is one of the most understandable approaches to classification [68]. On an intuitive level, the essence of the method is simple: we look at the neighbors around, which of them prevail, that are what we are. Formally, the technique is based on the compactness hypothesis: if the space metric among examples is delivered successfully, then similar examples are much more likely to lie in the same class than in different ones.

• In the case of using the method for classification, the object is assigned to the class that is the most common among the k neighbors of this element, whose classes are already known.

• In the case of using the method for regression, the object is assigned the average value of the k objects closest to it, the values of which are already known.

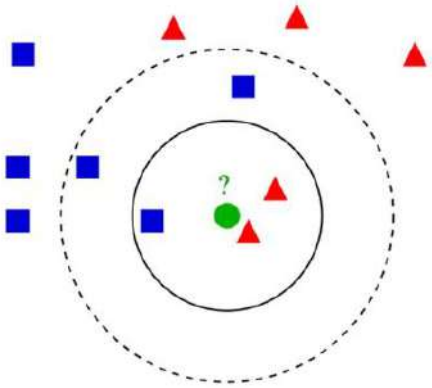


Figure 3.6: An example of k-nearest neighbor classification.

We have a test sample in the form of a green circle. The blue squares will be designated as class 1, the red triangles as class 2 .

• The green circle must be classified as class 1 or class 2 . If the area we are considering is a small circle, then the object is classified as 2nd class, because inside this circle there are 2 triangles and only 1 square.

• If we consider a large circle (with a dotted line), then the circle will be classified as 1st class, since there are 3 squares inside the circle as opposed to 2 triangles.

3.5.1 Theoretical component of the k-NN algorithm

3.5.1.1 Euclidean metric

Beyond the simple explanation, the expertise of the underlying mathematical components of the k-nearest neighbors algorithm is absolutely necessary Euclidean metric (Euclidean distance, or Euclidean distance) - a metric in Euclidean space, the distance between two points in Euclidean space, calculated by the Pythagorean theorem [69]. Simply expressed, this is the lowest feasible space across locations A and B. Although Euclidean distance is useful for small dimensions [70], it does not work for large dimensions and for categorical variables. The disadvantage of Euclidean distance is that it ignores the similarity between attributes. Each of them is seen as completely different from all the others.

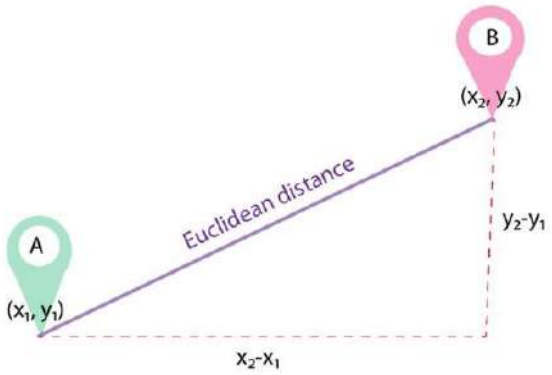


Figure 3.7: The Euclidean distance

The Euclidean distance is calculated using the following formula:

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \qquad (3.5)$$

Another important component of the method is normalization [71]. Different attributes usually have different ranges of values represented in the selection. For example, attribute A is represented in the range 0.01 to 0.05, and attribute B is represented in the range 500 to 1000). In this case, the distance values can be highly dependent on attributes with larger ranges. Accordingly, the data in most cases is going through normalization. There are two essential approaches to normalize data in cluster analysis: Min-Max normalization and Z-normalization.

When undertaking some type of study in which we have multiple variables measured on different scales and we want each of the variables to have the same range, we often normalize variables.

3.5.1.2 Feature selection

Experts in feature selection face a variety of challenges, including extracting and decreasing features. The features are what make every class in system have a variance or significance in term of it, sometimes required a sophisticated features generating [72] with specialists in order to increase features so raising dimensions of the feature space, that is in the event of a shortage of the required features, One of the most common techniques to

improve features is to work with statistical characteristics[73] such as spectral density, standard deviation, variance, and so o, using the available features. in the case of accumulation of the features, in order to reduce it , the features that ensure the best results and distinct differences in the distribution of situations must be identified. One of the most common ways to reduce the number of features and ensure the preservation of high accuracy is to use principle component analysis(PCA)[74], which allows you to reduce the number of features while also creating new features with different significance for each class.

3.5.1.3 Z-normalization

which normalize every value in data set so that the mean of every data set is 0 and the standard deviation is 1.

$$Nv = (x - M[x])/\sigma[x] \tag{3.6}$$

where:

σ is the standard deviation.

x is the original value.

M[x] is the mean of data.

In this case, most of the values fall within the range.

3.5.2 Min-Max normalization

Both values will be changed to the same scale/range when the data of the various scales is normalized. Both values, for example, will be in the range of 0 to 1.

The data will have a value of 0 for the lowest value and a value of 1 for the greatest value, with all other values falling between 0 and 1.

$$z = \frac{x_i - min(x)}{max(x) - min(x)} \tag{3.7}$$

where: xi

Is the original value. min(x) is the minimum value in data set . max(x) is the maximum value in data set .

What is the course of action?

• Download our data.

• Initialize k by choosing the optimal number of neighbors.

For each of the data samples:

. Using the data, calculate the distance between the querying example and the actual example.

. Add the index of the sample to the ordered collection, just like its distance.

• Sort an ordered list of distances and indices in ascending order from least to greatest.

• Choose the first k data points from the sorted collection.

• Take the labels of the chosen k entries.

• If we have a regression problem, we'll return to the average of the previously selected k labels.

• If we have a classification problem, we'll return to the most frequently occurring value of the previously selected labels k .

3.6 Spectral analysis

One of the most powerful techniques for processing experiments is spectral analysis. In particular, it is used for data analysis, detection of characteristic frequencies [75], noise suppression[76], etc.

The spectrum of the data set y(x) is some function of another coordinate (or coordinates, if we are talking about a multidimensional spectrum [77]) F(ω), Obtainable through the use of a specific algorithm . Fourier transform, power spectrum, and wavelet transform are examples of spectral analysis approaches.

3.6.1 Fourier Transform

The Fourier transform mathematically presents the signal y(x) as an infinite summation of sinusoids of the format F(ω) sin(ωx) .The function F(ω) is called the Fourier

transform, or the Fourier spectrum of the signal. Its argument represents the frequency of the appropriate component of the signal. The inverse Fourier transform transforms the spectrum F(ω) into the original y(x) signal. By definition,

$$F(\omega) = \int_{-\infty}^{\infty} \mathrm{y(x)} * e^{-i\omega x}\, \mathrm{d}x \tag{3.8}$$

Even if the signal is real, the Fourier transform is a complex value [78], as defined by definition in formula 3.8.The Fourier transform is particularly important in many mathematical applications, and a very efficient algorithm for it has been constructed. Called the FFT (Fast Fourier Transform) algorithm. it's so popular due to its super economical, which in almost all mathematical packages is organized as a subroutine.

The FFT algorithm has a rather strong limitation, which is not critical in practice. The idea is that the direct Fourier transform argument, i.e. the sample size y(xi) , must contain exactly 2n elements (n is any integer). Accordingly, the result of the FFT algorithm is a vector with 1 +2 (n−1) elements [79]. If the number of data does not match the power of 2, then to run the FFT algorithm, it is enough to pad the missing elements with zeros.

Considering the most typical situation for a physical experiment in calculating the Fourier spectrum of a real signal. we utilize the discretization of the following deterministic signal as model data (Figure3.11(a)) To further comprehend the Fourier transform:

y(x) = 1 sin(2π0.05x) + 0.5 sin(2π0.1x) + 0.1 sin(2π0.5x) On Figure3.11(b) shows the results of the FFT algorithm in the form of the Fourier spectrum modulus |F(ω)|, since, again, the spectrum itself is complex. It is very useful to compare the obtained amplitudes and the location of the spectral peaks in Figure 3.11 (b) with the definition of sinusoids [80] in the formula3.11. It is significant that if we subject the obtained absolute value of the Fourier spectrum Figure 3.11(b) to the inverse transformation

Fourier, which is also provided by the FFT algorithm, then the profile of the original signal will be reconstructed correctly, but will be shifted by a certain distance along the x-axis. This is because taking the absolute value of the complex spectrum destroys information about the relative phase of the data samples. Otherwise, the signal y(x) is restored with great accuracy, which is typical for a smooth signal change [81]. If, however, the complex Fourier

spectrum is used as the input data of the inverse Fourier transform, then the match will be complete [82].

3.7.Conclusion

The k-nearest neighbors approach is a basic supervised machine learning algorithm that may be used to handle classification and regression issues. It is simple to implement and understand, but has a significant drawback - a significant slowdown when the amount of data grows.

The k-Nearest Neighbor algorithm classifies based on the distance to a certain number (K) of training samples. This family of algorithms is called instance-based learning, since there are no parameters to study. the model assumes that the distance is sufficient for inference, otherwise it makes no assumptions about the underlying data or its distribution.

Chapter 4

Implementation of machine learning in photovoltaic system

Introduction

In this chapter, we will present the results of the operating of the multicellular inverter in two control modes sliding and exact linearization presented in the first chapter 1 in Both healthy mode case and with presence of faults in one , or all capacitors separately or instantaneous , and we use machine learning to classify this faults

4.1-- Simulation Results

In order to validate the behavior of the multicellular converter with the sliding mode and exact linearization controls in the case of flying capacitors faults, the system shown in Figure 1 is adopted.

Table 1 shows system parameters.

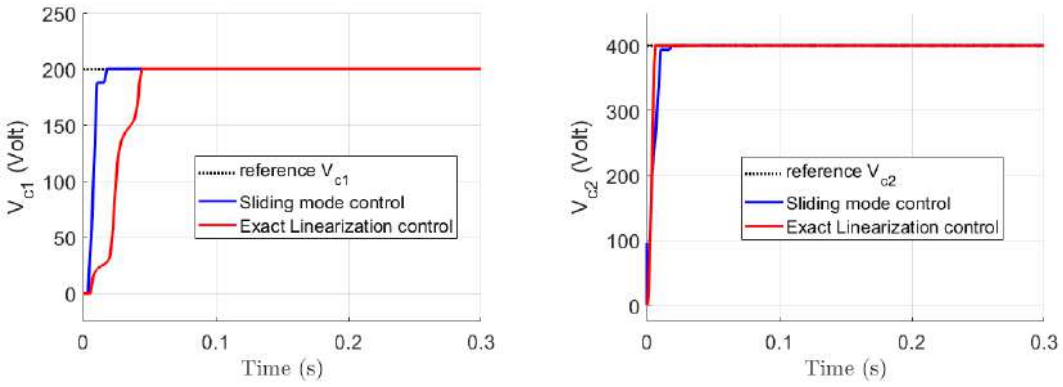
Table 4.1. System Simulation Parameters

Parameter	Value
Photovoltaic system voltage	$V_{dc} = 600V$
Flying capacitor	$C_1 = C_2 = 400\mu F$
Electric load inductor	$L_L = 2\ mH$
Electric load resistor	$R_L = 25\Omega$

A set of computer simulation runs is conducted using sliding mode control and exact linearization technique to demonstrate the effectiveness of proposed control strategies.

Therefore, a study of performance of these control methods applied to multicellular converter considering all capacitor faulty cases using MATLAB/Simulink software, starting from healthy mode to faulty in capacitor C_1, C_2 and C_1, C_2 simultaneously. A comparison is made to better illustrate the impact of failing in capacitors on system behavior and to show the better controller in this circumstance.

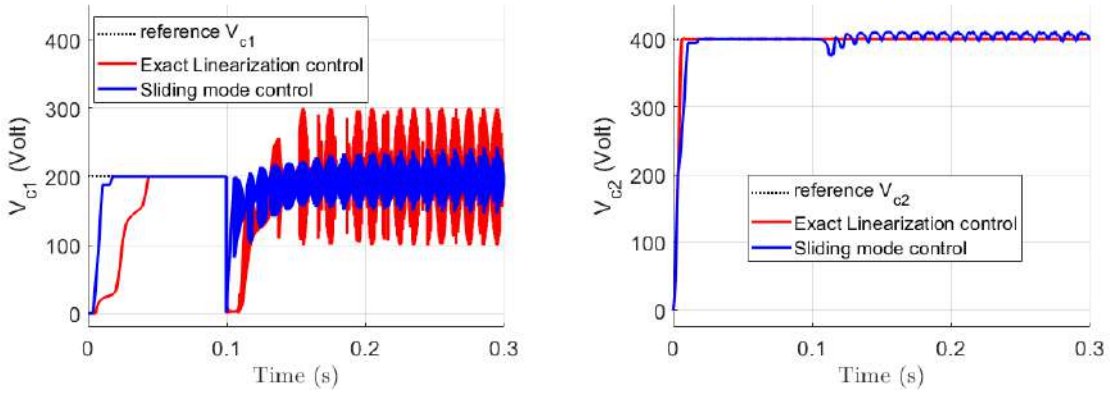
Figure 3.1 show the voltage and corresponding error compared to the reference of flying capacitors V_{c1} and V_{c2} for both controls in healthy operating mode. In sliding mode control, the two voltages track their references in steady-state with time response equal to 0.02s for both V_{c1} and V_{c2} and a ripple (ΔV) of 0.02V. However, in exact linearization time response for V_{c1} and V_{c2} are respectively 0.045s and 0.007s with a ripple of 0.02V. Hence, the sliding mode control performs better than the exact linearization control in terms of ripple and time response.



(a) **(b)**

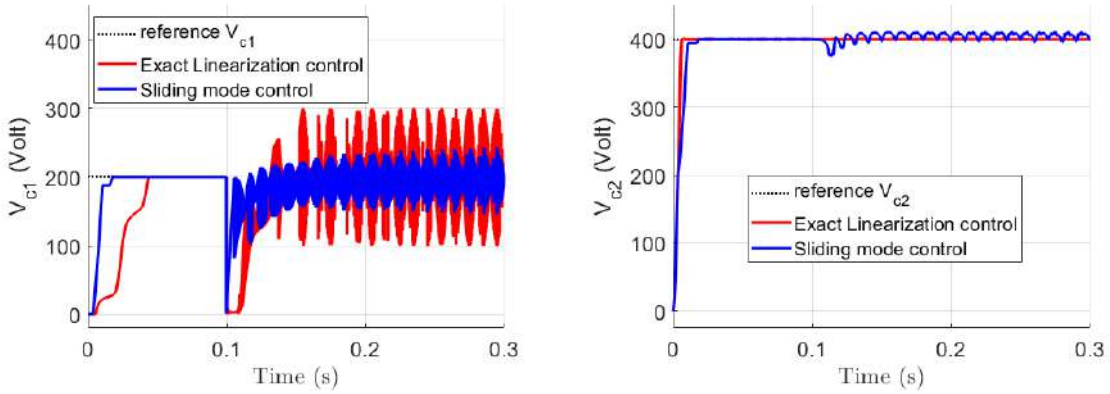
Figure 4.1. Error voltage C_1 and in C_2 healthy mode both control

As shown in Figures 4.2 and 4.3, in case of occurrence of a failure in C_1, the ripple magnitude of V_{c1} increases, in both controls, by 200V for exact linearization and by 95V for sliding mode. On the other hand, V_{c2} is less affected, and the ripple is less intense, i.e. 15v for the sliding mode and 1v for the exact linearization. This implies that the sliding mode control shows better robustness than the exact linearization control.



(**a**) (**b**)

Figure 4.2. The voltage of capacitor C_1 and C_2, failure C_1 in both controls.

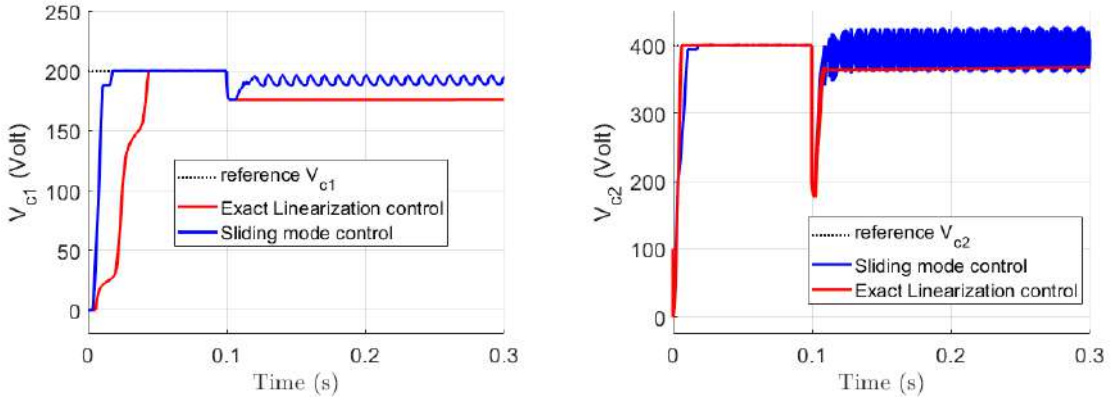


(**a**) (**b**)

Figure 4.3. Error voltage of capacitor C_1 and C_2, failure C_1 in both controls.

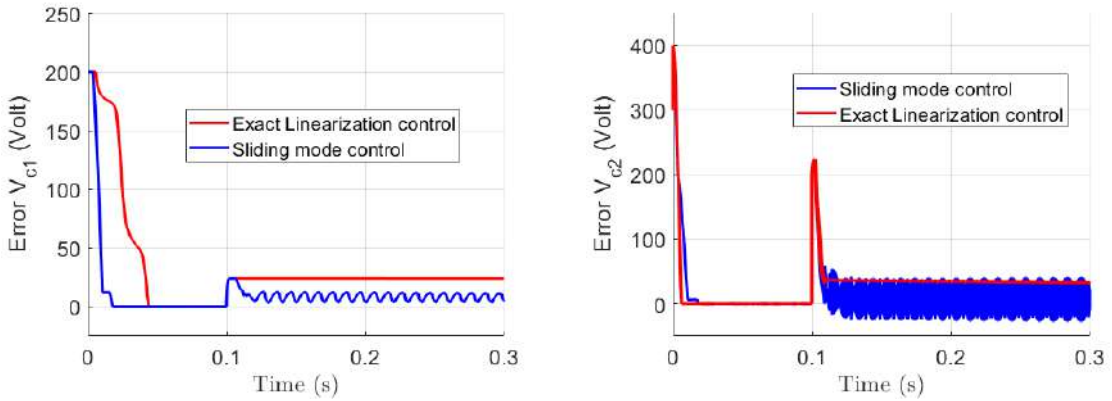
Figures 4.4 and 4.5 illustrate the results for the failure of capacitor C_2. This case shows that the failure in C_2 has less effect on the behavior of the system. This case shows fewer voltage ripples in V_{c2} for both controls, compared to ripples generated in V_{c1} due to the failure of capacitor C_1.

The associated ripple for V_{c2} is 1V in exact linearization and 60V in sliding mode, while for V_{c1} it is equal to 1V in the case of exact linearization and 10V in sliding mode. In addition, the sliding mode control has better accuracy than the exact linearization control. It is 6V for Vc1 and 10V for V_{c2} in the case of sliding mode control and 22V for V_{c1}; however, it is 33V for V_{c2} in the case of exact linearization.



(a) **(b)**

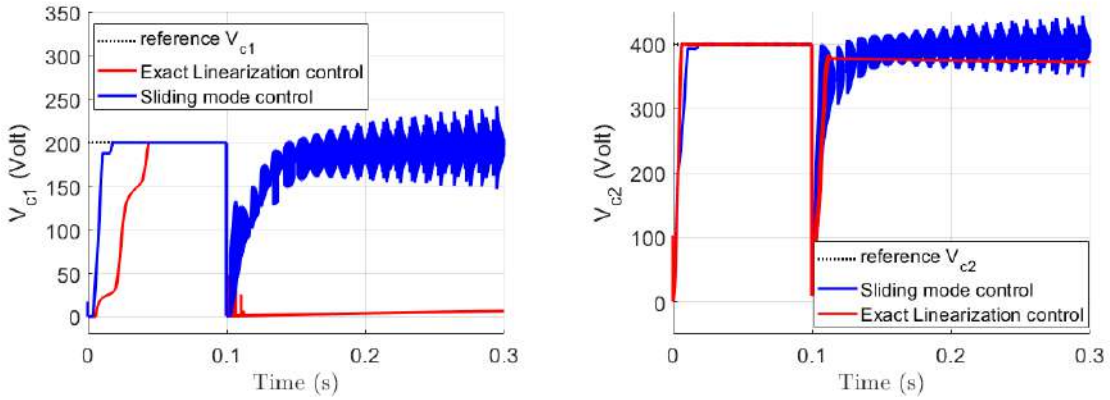
Figure 4.4 Voltage of capacitor C_1 and C_2, failure C_2 in both controls.



(a) **(b)**

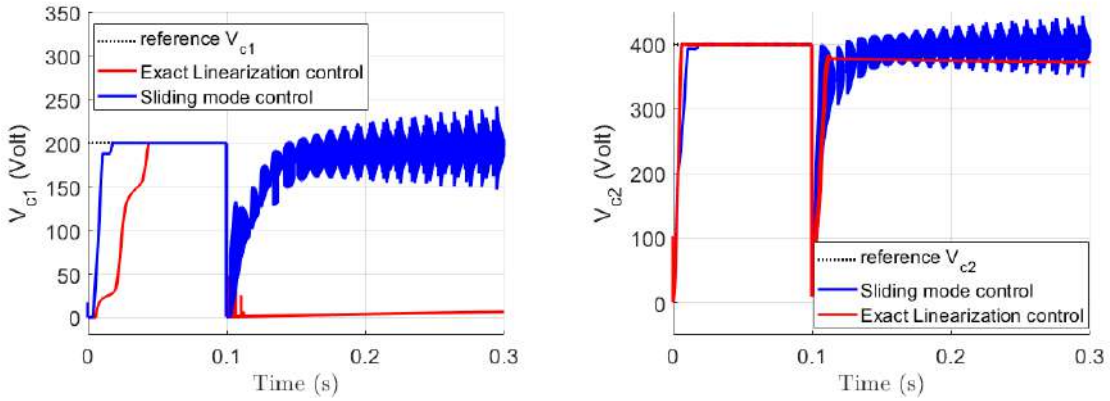
Figure 4.5. Error voltage of capacitor C_1 and C_2, failure C_2 in both controls.

The impact of the fault of the two capacitors C_1 and C_2 simultaneously is illustrated in Figures 4.6 and 4.7. This case shows that the failure of C_1 and C_2 has a significant impact on the system behavior. The flying capacitor voltages follow their references with strong ripples in sliding mode control of about 90 volts for V_{c1} and V_{c2}. On the other hand, for the control by exact linearization, the voltage V_{c1} drops to 0V without ripple, while V_{c2} follows the reference with 35 volts of precision. Form above discussion, the sliding mode control was less affected by this type of fault than the exact linearization control despite the high ripple generated.



(**a**) (**b**)

Figure 4.6. Voltage of capacitor C_1 and C_2, failure C_2 and C_2 in both controls



(**a**) (**b**)

Figure 4.7. Error voltage of capacitor C_1 and C_2, failure C_2 and C_2 in both controls.

Table 4.2 resume the comparison study results.

Table 4.2. Comparison between sliding mode and exact linearization controls in different operating modes

Control mode		Time response (s)		Precision (V)		ΔV(V)	
		V_{C1}	V_{C2}	V_{C1}	V_{C2}	V_{C1}	V_{C2}
Sliding mode control	Healthy mode	0.02	0.02	0	0	0.02	0.02
	C_1 Failure	0.12	0.02	5	5	95	15
	C_2 Failure	0.01	0.01	6	10	10	60
	C_1 and C_2 Failure	0.01	0.01	5	5	90	90
Exact linearization control	Healthy mode	0.045	0.007	0	0	0.02	1
	C_1 Failure	0.12	0	0	0	200	1
	C_2 Failure	0	-	22	33	1	1
	C_1 and C_2 Failure	0	-	200	30	1	1

From the above discussion about simulation results comparing sliding mode and exact linearization controls, applied to three cells multicellular converter used in photovoltaic system, for different modes (healthy, C_1 failure, C_2 failure, and both C_1 C_2 failure), obviously, the slider mode control shows more robustness against all types of faults considered and a shorter response time in a transient state, while the exact linearization control shows less voltage ripple. However, the major disadvantage of exact linearization control is zero load current when C_2 is faulty.

The results, also, indicate that when capacitor C_1 fails, it generates ripples in V_{c1}, and similarly for capacitor C_2 where it generates ripples in Vc_2. It is clearly noticed that the fault of the capacitor C_1 causes more brutal ripples than the fault relating to C_2. Moreover, the sliding mode control is more robust than exact linearization control, in terms of capacitor failure. Al though in the healthy case, the exact linearization shows slightly better performance than the sliding mode control.

In the next section, the changes that occur in feature spaces of the system resulting from faults in capacitors and the impact of capacitor fault on the feature space geometry and/or distribution are discussed.

4.2. Data Processing and Feature Space Analysis

Data processing and analysis is a critical step, for it creates a feature space that distinguishes between different operating conditions, i.e., the feature space differentiates between healthy and faulty operating modes. For this purpose, the chosen features must maximize the separation between the normal and each of the failure modes in the feature space. In order to define the best sensitive features to distinguish the normal class of each of the failure modes in the multicellular converter, knowledge of the physical dynamics of the latter is needed. The capacitor voltages Vc1 and Vc2 are used for this purpose. Thus, there are a number of features that correspond to the number of capacitors, to which can be added another feature such as the load current iL to ensure better separation of the different failure modes. As indicated by equation (4.3), the capacitor voltages are regulated to their reference values in the case of a healthy operating mode (normal condition), whereas if a fault occurs in one or two capacitors, the corresponding voltages cannot reach their references with the proposed control. This change over time between the real values of the capacitor voltages and their references makes it possible to detect faults through changes of normal class in the space of the characteristic curves.

Figures (4.8-4.10) show the characteristics for different operating modes with sliding mode and exact linearization control derived from simulation data. V_{c1}, V_{c2} and i_L was used as the axis for the representation of the space of 2D or 3D entities.

Figure 4.8 shows the two-axis feature space with sliding mode control for all healthy or faulty modes. İt is noticed that healthy feature spaces are very small regimes drowned in the global feature space. The spatial characteristics of the different failure modes are well separated but with some superposed or interlaced region. To overcome this situation, the current " i_L" is suggested as the third axis (see Figure 4.9). This last figure shows that all the modes are well separated with no ambiguity. These results show that through these features spaces, it is possible to detect or locate the different faults.

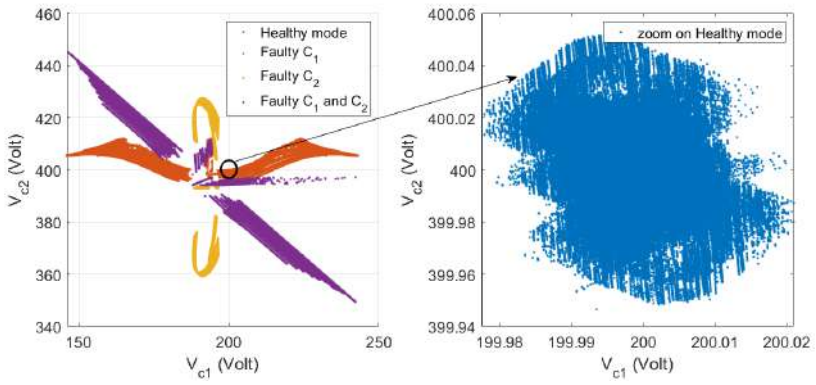


Figure 4.8. Feature space during failure of capacitor using sliding mode control

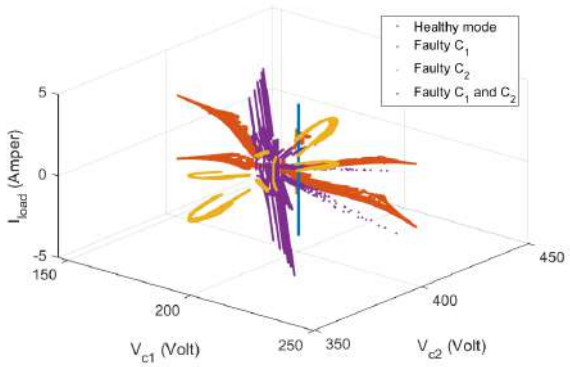


Figure 4.9. 3D feature space during failure of capacitor using sliding mode control

The feature space corresponding to the exact linearization control for all the modes is shown in Figure 4.10. It can be seen that the three fault modes in the capacitor C_1, C_2 and C_1/ C_2 simultaneously are well separated. On the other hand, the fault C_1 completely covers the healthy mode; the latter represents a restricted mode compared to the modes with defaults. As in the case of control by sliding mode and to overcome this problem of covering the feature spaces each other, the current "iL" is added as the third axis (see Figure 4.11). This latest figure shows that all the modes are better separated, but always with an ambiguity between the healthy mode and the faulty mode in C_1; however, this ambiguity is not very important. These results show that through these feature spaces, it is possible to detect or locate the various faults with a slight ambiguity in the case of a fault in capacitor C_1.

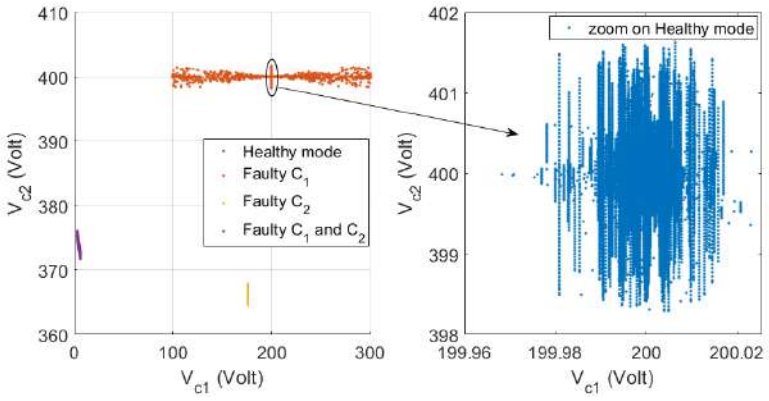


Figure 4.10. Feature space during failure of capacitor using exact linearization control

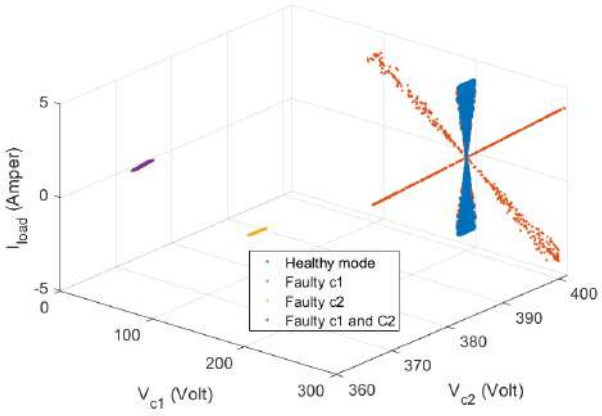


Figure 4.11. 3D feature space during failure of a capacitor using exact linearization control

According to obtained results from feature spaces, the sliding mode control presents a better separation between the different modes compared to the exact linearization control. In addition, the sliding mode control shows better robustness compared to the exact linearization control. This is observed by the zero-load current when C_2 is faulty, in the case of exact linearization control.

4.3-. K-Nearest Neighbor (KNN) implementation

After the data has been collected and features selected, the available data can now be used to apply the Machine Learning algorithm for classification.

Since the data will be classified into more than two classes, the k-nearest neighbor (KNN) algorithm was chosen because it is multiclass, based on determining the distance between all features. The classification is based on the property of closest distance, which is the feature with which it is similar or identical.

The process entails teaching the algorithm with the available data, i.e. training it to know the characteristics of each case in order to identify it in order to classify it, and this is done by teaching the data and dividing it randomly into 70% training and 30% testing and then submitting the training data to the machine learning algorithm to build a classified model based on it (Figure.4.12).

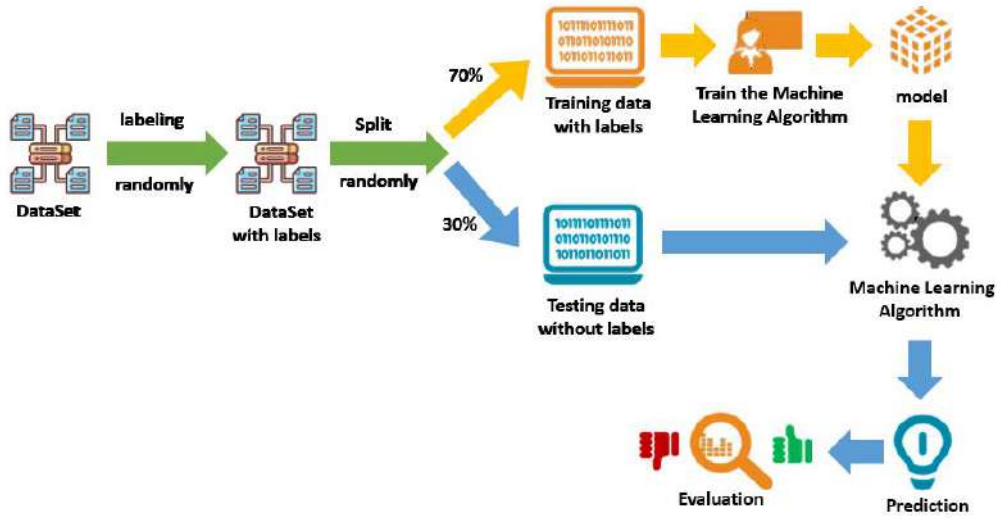


Figure 4.12. the KNN workflow.

The histogram in the figure 4.13 resume the behavior of the proposed fault detection approaches in term of accuracy. it was observed that the sliding mode classification model is more accurate than the exact linearization model.

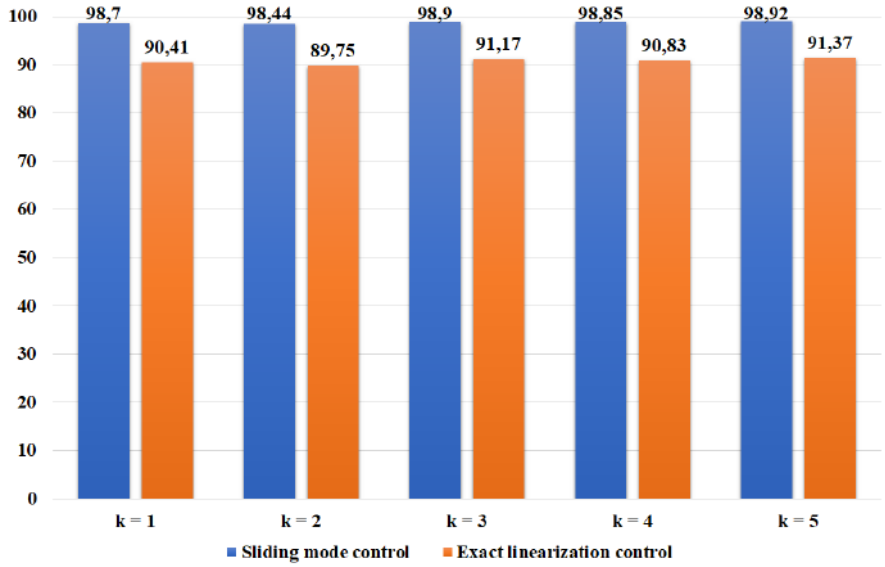


Figure 4.13 Histogram represent the comparison between the two modes of control

Conclusion

This chapter presented a comparative study between a sliding mode and exact linearization controls applied to three cells multicellular converter used in photovoltaic system under flying capacitors faults. The parameters used to differentiate between both controls are time response, precision and ripple of flying capacitors voltages in healthy and faulty modes (C_1 failure, C_2 failure and simultaneously failure of C_1 and C_2). In the field of time series, sliding mode control has better performance in terms of robustness and time response, compared to exact linearization control in healthy and faulty modes. It is found that the exact linearization control is smoother (less ripple) in both healthy and faulty modes.

In order to analyses the behavior of multicellular converters in four operating modes, two-dimension feature spaces are used with time domain axes to separate different operating modes. KNN algorithm for classification proved also that the sliding mode control presents a better class separation compared to the exact linearization control.

Conclusion

This work is an opportunity to address major issues related to renewable energy, especially in multi-cellular converters. Therefore, the study was carried out according to a professional project approach, which ensured the achievement of the objectives in terms of detection of capacitor faults in multi-cell power converters using KNN machine learning algorithms and determination of the optimal control method between exact linearization and sliding mode control. The realization of this project provided an excellent opportunity to study the main faults and their impact on the voltage and output current of a multi-cell battery converter.

This work could be of great benefit to the renewable energy sector. In fact, since multi-cell power converter systems run smoothly and produce a steady load current, capacitor faults can be located to handle high power equipment.

.As future works direction, the calculation of remaining useful life (RUL) in order to perform a fault prognosis of the considered system.

Biblioghraphy

[01] KN Nwaigwe, P Mutabilwa, and E Dintwa. “An overview of solar power (PV systems) integration into electricity grids”. In: Materials Science for Energy Technologies 2.3 (2019), pp. 629–633.

[02] Viktor Sebestyén. “Renewable and Sustainable Energy Reviews: Environmental impact networks of renewable energy power plants”. In: Renewable and Sustainable Energy Reviews 151 (2021), p. 111626.

[03] Boubakeur Rouabah. “Commande d’un convertisseur multicellulaire pour une applicationde filtrage actif”. PhD thesis. 2018.

[04] Boubakeur Rouabah and L Rahmani. “Direct LYAPUNOV control of multi-cells inverter applied to an active power filter with unbalanced canditions”. In: The International Conference on Electronics & Oil: From Theory to Applications. 2013.

[05] CH Basha and C Rani. “Different conventional and soft computing MPPT techniques for solar PV systems with high step-up boost converters: a comprehensive analysis”. In: Energies 13.2 (2020), p. 371.

[06] Houari Toubakh and Moamar Sayed-Mouchaweh. “Hybrid dynamic classifier for drift-like fault diagnosis in a class of hybrid dynamic systems: Application to wind turbine converters”. In: Neurocomputing 171 (2016), pp. 1496–1516.

[07] Boubakeur Rouabah, Houari Toubakh, and M Sayed. “Advanced fault-tolerant control strategy of wind turbine based on squirrel cage induction generator with rotor bar defects”. In: Annual Conference of the PHM Society. Vol. 11. 1. 2019.

[08]Boubakeur Rouabah et al. “Adaptive data-driven fault-tolerant control strategy for optimal power extraction in presence of broken rotor bars in wind turbine”. In: ISA transactions (2022).

[09]Boubakeur Rouabah et al. “Adaptive and exact linearization control of multicellular power converter based on shunt active power filter”. In: Journal of Control, Automation and Electrical Systems 30.6 (2019), pp. 1019–1029.

[10]Emmanuel Yofo and Philippe Djondiné. “Parallel Active Filtering based on a threephase multicell converter”. In: SCIREA Journal of Electrical Engineering 4 (2019), pp. 48–62.

[11]Boubakeur Rouabah. “Contribution à l’amélioration des performances d’un filtre actif parallèle de puissance par l’utilisation d’un convertisseur multicellulaire”.PhD thesis. 2021.

[12]Boubakeur Rouabah et al. “More Efficient Wind Energy Conversion System Using Shunt Active Power Filter”. In: Electric Power Components and Systems 49.4-5 (2021), pp. 321–332.

[13] Yifan Peng et al. “DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs”. In: Ophthalmology 126.4 (2019), pp. 565–575.

[14]Yifan Peng et al. “DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs”. In: Ophthalmology 126.4 (2019), pp. 565–575.

[15]Mohammad Zawad Ali et al. “Machine learning-based fault diagnosis for singleand multi-faults in induction motors using measured stator currents and vibration signals”. In: IEEE Transactions on Industry Applications 55.3 (2019), pp. 2378– 2391.

[16]Bahzad Charbuty and Adnan Abdulazeez. “Classification based on decision tree algorithm for machine learning”. In: Journal of Applied Science and Technology Trends 2.01 (2021), pp. 20–28.

[17]Michael J Bianco et al. “Machine learning in acoustics: Theory and applications”. In: The Journal of the Acoustical Society of America 146.5 (2019), pp. 3590–3628.

[18] Chakraborty, Arindam, “Advancements in power electronics and drives in interface with growing renewable energy resources,” Renewable & Sustainable Energy Reviews, Elsevier, vol. 15, no. 4, pp.1816-1827, May, 2011.

[19] Renewable Energy Today and Tomorrow, Proceedings of the IEEE, vol. 89, no. 8, Aug. 2001.

[20] C. Hua, Y. Fang, C. Wong, “Improved solar system with maximum power point tracking,” IET Renew. Power Gener. Vol. 12, no. 7, pp. 806-814, May 2018.

[21] https://www.energy.gov/eere/bioenergy/biomass-resources

[22]REN21, "Renewables 2013: Global Status Report (GSR)," [Online]. Available: http://www.ren21.net/, Jun. 2013.

[23]REN21, "Renewables 2019: Global Status Report (GSR)," [Online]. Available: https://www.ren21.net/wpcontent/uploads/2019/05/gsr_2019_full_report_en.pdf.

[24]https://cleanenergynews.ihsmarkit.com/research-analysis/global-renewable-capacity-growth-at-record-in-2021-and-will-ac.html

[25]REN21, "Renewables 2020: Global Status Report (GSR)," [Online]. Available: https://www.ren21.net/reports/global-status-report/ Jun. 2020

[26] https://www.iea.org/reports/renewables2021?utm_source=SendGrid&utm_medium=Email&utm_campaign=IEA+newsletters

[27]Y. Zheng, D. Hill, Z. Dong, "Multi-agent optimal allocation of energy storage systems in distribution systems", IEEE Trans. Sustainable Energy, vol. 8, no. 4, pp. 1715-1725, 2017.

[28] A. R. Sparacino, G. F. Reed, R. J. Kerestes, B. M. Grainger and Z. T. Smith, "Survey of battery energy storage systems and modeling techniques," 2012 IEEE Power and Energy Society General Meeting, 2012, pp. 1-8.

[29]Zhaoxia Yang "Power Decoupling Control and Optimization for a Photovoltaic Power Decoupling Control and Optimization for a Photovoltaic Inverter in D-Q Rotation Frame Inverter in D-Q Rotation Frame"Minnesota State University, Mankato March 2022

[30]A.Saliha, “Modélisation et commande d’un système de pompage photovoltaïque.” Magister thesis university < FERHAT ABBAS > ALGERIA, 2014.

[31] Y. hameed khattak, Modeling of High Power Conversion Efficiency Thin Film Solar Cells. PhD thesis, university polytechnic of valancia, 2019.

[32] S. petibon, Nouvelles architectures distribuées de gestion et de conversion de l’énergie pour les applications photovoltaïques. PhD thesis, University Toulouse 3 Paul Sabatier , France.., 2009.

[33] A.bELKAID, Conception et implémentation d’une commande MPPT de haute performance pour une chaine de conversion photovoltaïque autonome. PhD thesis, university of FERHAT ABBAS SETIF 1, Algeria, 2015.

[34] Elhor Abderrahmane "MPPT Technique Based on Neural Network for Photovoltaic System" Master Degree In Renewable Energy and Energy Efficiency Bragança 2021.

[35] Almas Hossain Mollah "Three Phase Grid Connected Photovoltaic System with Maximum ower Point Tracking" International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering Vol. 4, Issue 5, May 2015.

[36]Tat Luat Nguyen, Kay-Soon Low, "A Global Maximum Power Point Tracking Scheme Employing DIRECT Search Algorithm for Photovoltaic Systems," IEEETransactions on Industrial Electronics, vol. 57, no. 10, pp. 3456-3467, Oct. 2010.

[37]Makhlouf M, Messai F, Nabti K, Benalla " Modeling and simulation of grid connected photovoltaic distributed generation system". In: 1st international conference on renewable energies & vehicular technology, 26–28 Mar 2012, pp 187–193

[38] Saad M, El Hammoumi A, El Ghziza A"The Most Used MPPT Algorithms: Review and the Suitable Low-cost Embedded Board for Each Algorithm" Journal of Cleaner Production https://doi.org/10.1016/j.jclepro.2019.118983

[39]Anishma Singh "Comparative Analysis Of Mppt Algorithms For Photovoltaic Systems Delhi Technological University"Master Of Technology In Power Electronics Delhi Technological University JUNE 2022.

[40] Dorofte C, Borup U, Blaabjerg F. "A combined two-method MPPT control scheme for grid-connected photovoltaic systems". European Conference on Power Electronics and Applications, 2005: 10 pp

[41]Zhang H,Zhou H W, Ren J, et al. Three-phase grid-connected photovoltaic system with SVPWM current controller. IEEE 6th International Power Electronics and Motion Control Conference (IPEMC '09), 2009: 2161-2164.

[42] Milosevic M, Andersson G, Grabic S. Decoupling current control and maximum power point control in small power network with photovoltaic source. Power Systems Conference and Exposition (PSCE), 2006: 1005–1011

[43]ZHU YongLi “Comparative study of two stages and single stage topologies for Grid-Tie Photovoltaic Generation by PSCAD/EMTDC” 2011The International Conference on Advanced Power System Automation and Protection

[44]Jae Ho Lee, HyunSu Bae, Bo Hyung Cho. Advanced Incremental Conductance MPPT Algorithm with a Variable Step Size. 12th International Power Electronics and Motion Control Conference (EPE-PEMC), 2006: 603-607

[45]Xiao W D, Dunford W G. A Modified Adaptive Hill Climbing MPPT Method for Photovoltaic Power Systems. 35th Annual IEEE Power Electronics Specialists Conference, 2004: 1957-1963 4 Zhang H,Zhou H W, Ren J, et al. Three-phase grid-connected photo

[46] Yao Z Q, Zhang Q, Liu X M. Research on simulation of a three-phase grid-connected photovoltaic generation system based on PSCAD/EMTDC. Power System Protection and Control, 2010, 38(17): 76-81

[47] Gupta R, Gupta G, Kastwar D, et al. Modeling and Design of MPPT Controller for a PV Module using PSCAD/EMTDC. IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010: 1–6

[48] Dorofte C, Borup U, Blaabjerg F. A combined two-method MPPT control scheme for grid-connected photovoltaic systems. European Conference on Power Electronics and Applications, 2005: 10 pp

[49] A. Fernandez-Infantes, J. Contreras and J. L. Bernal-Agustin, “Design of grid connected PV systems considering electrical, economical and environmental aspects: A practical case”, Renewable Energy, 2005, pp. 2049.

[50] PETER R. SALAMEH “STUDY AND INTEGRATION OF PHOTOVOLTAIC SYSTEMS”,~ Faculty of Engineering at Notre Dame University-Louaize, AUGUST 2022.

[51] Osama Omari, “Conceptual Development of a General Supply Philosophy for Isolated Electrical Power Systems”, PhD Thesis, South Westphalia University of Applied Sciences, Campus Soest, Germany, February 2005.

[52] G. Seguier, F. Labrique., “Power Electronic Converters, DC-AC Conversion”, Springer-Verlag, Heidelberg, Germany, January 1993.

[53]Said El-Barbari, W. Hofmann., “Digital Control of a Four Leg Inverter for Standalone Photovoltaic Systems with Unbalanced Load”, TU Chemnitz, Chemnitz, Germany.

[54] J. S. Lai and F. Z. Peng, “Multilevel Converters—A new Breed of Power Converters, IEEE Transactions onIndustry Applications, Vol. 32, No. 3, 1996, pp. 509-517. doi:10.1109/28.502161.

[55] Senthil U, Fernandes BG (2003) Hybrid space vector pulse width modulation based direct torque controlled induction motor drive. Conf IEEE. PESC’03 3: 1112–1117.

[56]ABBASSI Smail et :(Modélisation et Commande non linéaire d`un Segway): Université Mohamed Khider Biskra 2018 / 2019

[57]Akram Ghorayeb:(Chapitre 6 COMMANDE DES SYSTÈMES NON LINÉAIRES) https://controlesautomatiques.files.wordpress.com/2020/01/6-system.-non-linc3a9aire-1.pdf.

[58]Amieur Toufik(Commande des Systèmes Non Linéaires par Mode Glissant Flou) 2009

[59]V. Utkin, J. Guldner, and J. Shi, (Sliding Mode Control in Electromechanical Systems). Taylor & Francis, London, 1999.

[60] R. B. Fernandez, J.K. Hedrick, (Control of multivariable nonlinear systems by sliding mode control), International Journal of Control, volume 46, n°3, pp. 1019- 1040, 1987

[61] Boubakeur Rouabah, Houari Toubakh, Moamar Sayed-Mouchaweh, Fault tolerant control of multicellular converter used in shunt active power filter, Electric Power Systems Research, Volume 188, 2020, 106533, doi: 10.1016/j.epsr.2020.106533.

[62] BENAZIZA Walid (Commande par modes glissants du suivi de trajectoires pour un robot mobile) Université Hadj Lakhdar Batna

[63] KASSOURI Dalila et SAIBI Hayat (Commande par mode glissant d’ordre fractionnaire d’un réseau électrique)21 juin 2018 Université Mouloud MAMMERI, Tizi-Ouzou .

[64] “IEEE Recommended Practices and Requirements for Harmonic Control in Electrical Power Systems”. IEEE std 519-1992.

[65] ATIR Nour el wiam etNEZLI Tayeb et BEN MEBAREK Latifa :(Etude et simulation de la commande d'un moteur à courant continu à excitation série par un redresseur

[66] Meynard TA, Foch H(1992) Multi-Level choppers for high voltage applications. EPE J 2(1):45–51

[67] M. Khodja “High-performance multicell series inverter-fed induction motor drive” journal of Electr Eng 2 November 2016.

[68]Erin L Allwein, Robert E Schapire, and Yoram Singer. “Reducing multiclass to binary: A unifying approach for margin classifiers”. In: Journal of machine learning research 1.Dec (2000), pp. 113–141.

[69]Pratap Chandra Sen, Mahimarnab Hajra, and Mitadru Ghosh. “Supervised classification algorithms in machine learning: A survey and review”. In: Emerging technology in modelling and graphics. Springer, 2020, pp. 99–111.

[70]Kilian Q Weinberger, John Blitzer, and Lawrence Saul. “Distance metric learning for large margin nearest neighbor classification”. In: Advances in neural information processing systems 18 (2005).

[71]Daniel Aloise et al. “NP-hardness of Euclidean sum-of-squares clustering”. In: Machine learning 75.2 (2009), pp. 245–248.

[72]Muhammad Ali Imron and Budi Prasetyo. “Improving algorithm accuracy knearest neighbor using z-score normalization and particle swarm optimization to predict customer churn”. In: Journal of Soft Computing Exploration 1.1 (2020), pp. 56–62.

[73]Yifeng Li, Chih-Yu Chen, and Wyeth W Wasserman. “Deep feature selection: theory and application to identify enhancers and promoters”. In: Journal of Computational Biology 23.5 (2016), pp. 322–336.

[74]B Chandra and Manish Gupta. “An efficient statistical feature selection approach for classification of gene expression data”. In: Journal of biomedical informatics44.4 (2011), pp. 529–535.

[75]Samina Khalid, Tehmina Khalil, and Shamila Nasreen. “A survey of feature selection and feature extraction techniques in machine learning”. In: 2014 science and information conference. IEEE. 2014, pp. 372–378.

[76]Aurélien Pichon et al. “Spectral analysis of heart rate variability: interchangeability between autoregressive analysis and fast Fourier transform”. In: Journal of electrocardiology 39.1 (2006), pp. 31–37.

[77]Toshio Yoshizawa, Shigeki Hirobayashi, and Tadanobu Misawa. “Noise reduction for periodic signals using high-resolution frequency analysis”. In: EURASIP Journal on Audio, Speech, and Music Processing 2011.1 (2011), pp. 1–19.

[78]Andre Rauh and Gonzalo R Arce. “Optimized spectrum permutation for the multidimensional sparse FFT”. In: IEEE Transactions on Signal Processing 65.1 (2016), pp. 162–172.

[79] Kui Fu Chen and Shu Li Mei. “Composite interpolated fast Fourier transform with the Hanning window”. In: IEEE Transactions on instrumentation and measurement 59.6 (2010), pp. 1571–1579.

[80] Glenn D Bergland. “A guided tour of the fast Fourier transform”. In: IEEE spectrum 6.7 (1969), pp. 41–52.

[81] JW Brault and OR White. “The analysis and restoration of astronomical data via the fast Fourier transform”. In: Astronomy and Astrophysics 13 (1971), p. 169.

[82] Steven M Kay and Stanley Lawrence Marple. “Spectrum analysis—a modern perspective”. In: Proceedings of the IEEE 69.11 (1981), pp. 1380–1419.

[72] Jason WH Wong, Caterina Durante, and Hugh M Cartwright. “Application of fast Fourier transform cross-correlation for the alignment of large chromatographic and spectral datasets”. In: Analytical chemistry 77.17 (2005), pp. 5655–5661.