NANYANG TECHNOLOGICAL UNIVERSITY

TRANSIENT EARTH VOLTAGE (TEV) BASED PARTIAL DISCHARGE DETECTION AND ANALYSIS

LUO GUOMIN

SCHOOL OF ELECTRICAL AND ELECTRONICS ENGINEERING

2013

Transient Earth Voltage (TEV) Based Partial Discharge Detection And Analysis

LUO GUOMIN

School of Electrical and Electronics Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2013

Acknowledgements

First of all, I would like to express my sincere appreciation to my supervisor, Dr. Zhang Daming, for his consistent help and encouragement throughout my research in Nanyang Technological University. His patience and kindness are greatly appreciated. His great knowledge and serious research attitude helped me to improve my research skills. I have learnt a lot from him.

I am also deeply grateful to Prof. Tseng King Jet for his knowledge, guidance, fruitful discussions throughout my study and research in Nanyang Technological University. I am greatly indebted to his broad interests and high enthusiasms for research.

I would like to thank all the technical staff in the Power Electronics Research Laboratory: Mr. Chua Tiam Lee, Ms. Lee-Loh Chin Khim, Ms. Tan-Goh Jie Jiuan, and Mr. Teo Tiong Seng; and the staff in the Electric Power Research Laboratory: Mr Lim Kim Peow and Ng-Tan Siew Hong, for their technical support during my research.

I also would like to express my thanks to my laboratory fellows and friends for their friendly assistance in my research and everyday life. I also appreciate the financial support for this research provided by Nanyang Technological University.

Special thanks are dedicated to my parents and my husband for their consistent encouragement.

Abstract

Partial discharge (PD) detection is an effective way to evaluate the insulation condition of electrical equipment in power systems. The non-intrusive TEV-detecting method which detects transient earth voltage (TEV) signals from the external surface of equipment does not require interruptions of electrical operations and is thus preferred by more and more researchers, engineers and users. However, as a new technique, TEV based PD measurement is not well developed in many aspects, for example, the measuring system and the de-noising methods. Consequently, the research and development of the TEV based PD measurement has become an interesting topic in recent decades.

This thesis presents an investigation on the sensing system and the de-noising methods of TEV based PD measurement system. First of all, the mechanism, popular measuring systems, noise types and existing de-noising methods of PDs are reviewed. Secondly, based on the characteristics of TEV signals, a TEV based PD measuring system was proposed and its effectiveness has been demonstrated by an experimental test. Next, the optimal settings of a popular de-noising method for non-impulsive noise, wavelet thresholding, are selected and its processing efficiency is enhanced by using parallelism algorithm in C environment. Furthermore, the wavelet entropy is proposed to classify PD pulses from impulsive noises. Finally, a noise reduction system using Fourier transform and time-frequency entropy is proposed to reject various kinds of noises.

The non-intrusive PD measuring techniques have been more and more popular in recent years. In this thesis, a TEV based PD measuring system is proposed. The major parts: non-intrusive sensor and high-pass filter are designed according to the characteristics of TEV signals. The performance of proposed system is demonstrated by an experimental test where the PD signals are collected by both TEV and HFCT sensors. By considering the features of proposed system, the detected TEV signals are well simulated.

Due to the external locations of TEV sensors, the performance of TEV based PD detection is limited by noises. To remove non-impulsive noises, wavelet thresholding is often used. As the de-noised results depend on the settings of algorithm, the optimal ones are selected according to the features of TEV signal and the proposed system. Further, the processing efficiency of wavelet thresholding with optimal settings is enhanced.

As wavelet transform is good at time-frequency analysis of PD signals, its capability in rejecting impulsive noises is also explored. Therefore, wavelet entropy is proposed. By comparing with the traditional energy spectrum, the wavelet entropy is more stable to represent a single pulse. With the help of a trained ANN whose parameters are selected carefully, the PD pulses can be extracted with good percentages.

The impulsive noise reduction based on features of single pulses is often ineffective when PD pulse and noise occur at the same time. Thus, a de-noising system is proposed to remove both impulsive and non-impulsive noises even if they occur at the same time. In this system, Fourier transform and time frequency entropy are employed. The de-noised results of two experimental signals and one field-collected signal show that the proposed noise-rejecting system is effective in extracting real PD pulses.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AE	Acoustic Emission
AI	artificial intelligence
AM	amplitude modulation
ANN	artificial neural networks
BP	back-propagation
BR	Bayesian regulation
CT	Current Transformer
DGA	Dissolved Gas Analysis
EL	Electroluminescence
EM	Electromagnetic
FFT	Fast Fourier Transform
FM	Frequency Modulation
GHz	Gigahertz
GIS	Gas Insulated Switchgear
HFCT	High Frequency Current Transformer
HHT	Hilbert-Huang transform
HV	High Voltage
Hz	Hertz
kHz	Kilohertz
LM	Levenberg-Marquardt
MHz	Megahertz
MSE	Mean Square Errors
MSI	Modulated Sinusoidal Interferences
NEMA	National Electrical Manufacturers Association
NN	Neural Network

PD	Partial Discharge
PFC	Power Factor Correction
QN	quasi-Newton
RIV	Radio Influence Voltage
RTD	Resistive Temperature Detector
SNR	Signal to Noise Ratio
STFT	Short-Time Fourier Transform
TE	Transverse Electric
TEM	Transverse Electromagnetic
TEV	Transient Earth Voltage
TF	Time Frequency
TM	Transverse Magnetic
UHF	Ultra High Frequency
WT	Wavelet Transform

List of Principal Symbols

$\|A\|$	Absolute value of A
$\|\|A\|\|$	Norm of A
A^T	Transpose of matrix A
$<A,B>$	Inner conduct of A and B
$A*B$	Convolution of A and B
$A \cup B$	Union of A and B
$A \cap B$	Intersection of A and B
$A \subset B$	A is a subset of B
$A \perp B$	A is orthogonal to B
$A \oplus B$	Direct sum of A and B
C_k	Capacitance of coupling capacitor
C_t	Capacitance of test object
f	Original signal
$\tilde{F}$	Recovered signal
Δf	Frequency bandwidth
f_m	Central frequency
f_{up}	Upper frequency limit
f_{low}	Lower frequency limit
$H(s)$	Transfer function
V_d	Voltage after detector in coupling capacitor method
$V_{\min}$	Breakdown voltage
V_{pd}	Voltage of PD pulse
Z_d	Impedance in coupling capacitor method
n	Pulse number
pC	Pico-coulomb

ppm	Parts per million
q	Charge magnitude of PD
τ	Decaying time of an exponential pulse
ϕ	Phase angle of pulse
ε	Permittivity
μ	Magnetic permeability
$\psi_{j,k}$	Wavelet function
σ	Estimation of white noise
$L^2(\mathbb{R})$	Finite energy function
$\mathbb{R}$	Real number sets

CHAPTER 1
INTRODUCTION

1.1 Motivation

Partial discharge (PD) is closely related to insulating conditions of electrical apparatus in power systems. When PDs occur in insulations, small currents arise. Without any treatment, the discharge currents bridge the electrodes completely which certainly results in large short-circuit current and breaks down the equipment. PD phenomenon is an indication of degradation of insulation materials. Thus, the detection of PD at early stages plays a crucial role in increasing the service life of power equipment.

The research on PD measurement can be traced back to the beginning of 19th Century [1, 2]. Previously, the detection of PD relied more on the judgments of experienced engineers or workers. For example, noisy sounds happen when PDs occur. Those sounds can be recognized by experienced workers. Over the years, the level of investigations related to the PD mechanism and its measurement has increased considerably and a number of measurement techniques were introduced. Electrical measurements which are more sensitive, suitable for all kinds of equipments, and do not affect the insulations are preferred in practical applications. Traditional methods, for example, coupling capacitor method and ultra-high frequency (UHF) method have been commonly applied and they are reliable in detecting PDs. However, those traditional methods have a very serious drawback. They require electrical shutdown: due to the large amount of interferences during on-line test, the coupling capacitor method is usually employed in off-line tests; although external UHF couplers were designed, their performance is much worse than internal ones whose installation needs electrical shutdown. Thus, a PD measurement that is reliable and has no interruptions of electrical operation is needed.

At the end of 19th Century, the transient earth voltage (TEV) caused by PD occurrence was discovered and used as a method to evaluate the insulation conditions [3]. According to previous research, an impulsive voltage signal is caused by the electromagnetic waves of PD and propagates from the PD source to the earth via the grounded parts. Therefore,

if there are any dielectric openings on the metal cladding of electrical equipments, it is possible to judge the PD existence by detecting TEV signals on the external surface of the cladding. Since the detection of TEV signals does not need installations of sensors or couplers inside the equipment, TEV based PD measurement becomes more and more popular in recent years. However, as a new concept, TEV based PD measurement technique has not been well developed and more investigations are needed.

First of all, the TEV based PD measuring system was rarely discussed in details. The measuring method is the most important factor of an effective PD detection. In TEV based PD measurement, the sensors are located on the external surface of HV equipment. Before captured by non-intrusive sensors, the TEV signals have to propagate a long way such that the high frequency energy of PD pulses distorts and attenuates greatly. Therefore, the sensitivity of TEV sensor should be high enough to collect any weak pulses. At the same time, the frequency response bandwidth of the sensors should be wide enough to collect as much TEV energy as possible and to distinguish individual pulses when PD pulse occurs in high density. Therefore, the design and frequency characteristics of TEV sensor and the structure of measuring system should be carefully discussed and considered.

In addition, the noise reduction methods for TEV signals need more investigations. For the non-intrusive PD measurement where sensors are mounted on the external surface of equipment, noise is always a major limitation of measurement accuracy. Because of the influence from the external disturbances, the oscillatory waves from amplifier, the random impulsive noises from unknown sources and so on, PD pulses are mostly immersed in noise. Therefore, to produce a reliable TEV detection, the de-noising ability of measuring system must be considered. Because of different principles, the frequency spectrums of TEV signals are different from others such as UHF signals. The noise reduction methods of other measurements cannot be used directly in TEV analysis. Therefore, the noise reduction methods for TEV signals should be developed specifically. For some popular methods which can be applied in any other conditions, their optimal settings also should be selected according to the characteristics of TEV signals.

On the whole, the TEV based PD measurement system is required to be capable of effectively capturing PD signals, rejecting noises of all types, and extracting PD pulses with high reliability. Consequently, the research in this thesis is motivated by the quest for designing the appropriate non-intrusive sensor and TEV measuring system, enhancing

the capability of TEV measurement by non-impulsive noises rejection, and developing algorithms to remove any PD-like noises.

1.2 Objectives

The research in my thesis aims to develop a TEV based PD detection system that can detect PD signals effectively and reject noises as much as possible. The objectives of my research are as follow:

Objective 1: design a non-intrusive PD measurement system based on TEV method.

The hardware of PD measurement system based on TEV method usually includes three parts: the non-intrusive TEV sensor, the high-pass filter and the signal display and storage equipment such as oscilloscope. The first two parts: sensor and filter, need special design to shield from external interferences and to capture signals with suitable frequency response. The main objective of this research is to design a TEV-based PD measurement system which consists of a non-intrusive sensor and a high-pass filter.

Objective 2: optimize the existing wavelet based noise reduction algorithms

Two types of noises are commonly encountered in PD measurement: non-impulsive and impulsive noises. Wavelet thresholding algorithms were found to be effective in PD signal processing and applied successfully in removing non-impulsive noises. However, some issues associated with practical application have to be addressed. The objective of this part of research is to optimize the existing wavelet thresholding algorithm in terms of selecting suitable threshold and wavelet, and of enhancing the processing efficiency.

Objective 3: develop impulsive noise reduction algorithms

The impulsive noises usually come from electronic equipment, switching operation, and random pulse sources. These pulses are unavoidable in practical tests. However, unlike non-impulsive noises, the impulsive noises are much more difficult to reject due to their similar features as those of PDs. For example, the noises have impulsive waveforms and some of them also have wide frequency spectrums. A number of methods were proposed to solve this problem. However, the impulsive noise rejections for TEV measurement have not been extensively reported. The main object of this research is to develop de-noising algorithms to reject impulsive disturbances and extract PD pulses.

1.3 Contributions

The contributions of this thesis are listed below:

1. Implementation of a PD measurement system with TEV method.

A PD measuring system with TEV method is proposed. This PD measurement system includes a non-intrusive TEV sensor and a Butterworth high-pass filter. The designs of non-intrusive sensors for different cases have been developed and the settings of high-pass filter are obtained. The effectiveness of this proposed system was illustrated with an experimental test by comparing the detected signals from both TEV sensor and commercial HFCT.

2. Selection of optimal settings of wavelet thresholding based non-impulsive noise reduction

Wavelet thresholding was proved effective in rejecting non-impulsive noises in PD measurement. The settings for best results vary a lot with different systems. Thus, the optimal settings of TEV based measurement are explored in terms of threshold, thresholding function and wavelets. Based on the characteristics of TEV measuring system, the optimal combination of threshold and thresholding function, and the optimal wavelets are selected. Simulation results under different scenarios have shown the excellent performance of wavelet thresholding algorithms with selected settings.

3. A speed-up algorithm for real time wavelet thresholding.

The capability of high-speed signal processing of wavelet thresholding algorithm by using parallelism is discussed. Based on the comparisons between the processing durations of the same function on both MATLAB and C platforms, C was chosen. Further, the problems in parallelism such as boundary distortion and threshold estimation are analyzed and solved accordingly. The processing durations of different data show the potential of applying high-speed wavelet thresholding in real time de-noising.

4. An impulsive noise rejection method by using wavelet entropy

Rejection of impulsive noises is one of the most difficult issues in PD measurement. When PD pulses do not overlap impulsive noises, the pulses can be classified one by one. Therefore, an ANN based PD classification algorithm is proposed. The wavelet entropy of each pulse is first calculated and then distinguished by a trained ANN. Entropy which

is stable to measure the disorder was proved to be effective in representing the distribution of the wavelet coefficients of a single pulse.

5. A de-noising system by using Fourier transform and time-frequency entropy

Concurring PD and noise pulses are often found in practical tests. A de-noising system is proposed to remove both non-impulsive and impulsive noises even if the PD and noise pulses occur at the same time. First, the repetitive noises which produce large singularities in frequency domain are removed by Fourier transform. With the help of TF entropy, the rejection of modulated sinusoidal noises is followed. Finally, the rejection of random noise is presented. The de-noising procedures of contaminated PD signals illustrate the effectiveness of proposed method.

1.4 Organizations of the thesis

A brief introduction of the contents of this thesis is given below:

Chapter 1 starts with the presentations on the motivation of our work. Then the objectives and contributions of this research are described. Finally, the organization of this thesis is described.

Chapter 2 gives a literature review of the PD phenomenon, the characteristics, the history and types of its measurement, and the existing noise rejection methods. The merits and drawbacks of different PD measuring methods are analyzed and discussed. The noises during PD measurement, existing de-noising methods, and challenges of de-noising TEV signals are introduced as a guide for the development of noise rejection.

Chapter 3 introduces a PD measurement system based on the fundamentals of TEV theory. The design of non-intrusive sensor and high-pass filter are followed. The effectiveness of this system is illustrated by an experimental test. The detected TEV signals have been simulated and the simulated pulses have been compared with the detected ones.

Chapter 4 describes the selections of optimal combinations of threshold and thresholding function as well as the wavelets. The simulations under different conditions verified the effectiveness of selection. Furthermore, the processing efficiency of wavelet thresholding algorithm with optimal settings is enhanced by using parallelism in C environment.

Chapter 5 proposes an impulsive noise reduction algorithm with wavelet entropy. The advantages of entropy are illustrated by comparing the wavelet entropy distribution and the energy distribution of a single pulse. The real PD pulses are extracted by a trained ANN. The PD de-noising results with combined signals shows the good performance of the proposed wavelet entropy based algorithm.

Chapter 6 presents a noise reduction system based on time-frequency entropy and Fourier analysis. Both the non-impulsive noise and impulsive noises can be rejected by this system. Several algorithms are developed to remove different kinds of noises. Two experimental and one field cases were carried out to demonstrate the effectiveness of TF entropy based method.

Chapter 7 provides the conclusions of this thesis as well as the recommendations for future research.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

The purpose of PD-detection-based diagnosis of insulation performance is to provide information to judge whether PD occurs or not, and then identifying the source location, defect types and condition of insulation performance, finally leading to risk assessment and prediction of the remaining life of an apparatus. In order to perform these actions using PD detection method with high reliability, one should consider the following concepts systematically: (i) the PD phenomena, (ii) the properties of PD signal, (iii) the properties of measurement system and sensor, and (iv) insulation diagnosis based on data acquisition, analysis, and assessment [4]. Fig.2-1 shows the procedure of insulation diagnosis of electric power apparatus based on PD detection. Although the methods proposed by different researchers might vary a lot, such procedure summarizes the essential steps of PD detection.

In this chapter, the four concepts in Fig.2-1 are reviewed systematically. First, the PD phenomenon is reviewed in terms of mechanism, breakdown and types. Next, the equivalent circuit and occurrence of PD are reviewed for a better understanding of PD properties. Then the review of PD measurements, both using non-electrical and using electrical methods, is presented. Further, the noise rejection which is the major problem in accurate PD measurement and insulation diagnosis is reviewed in aspects of the most common noise types and of de-noising methods. Finally, the applications of PD measurement are reviewed.

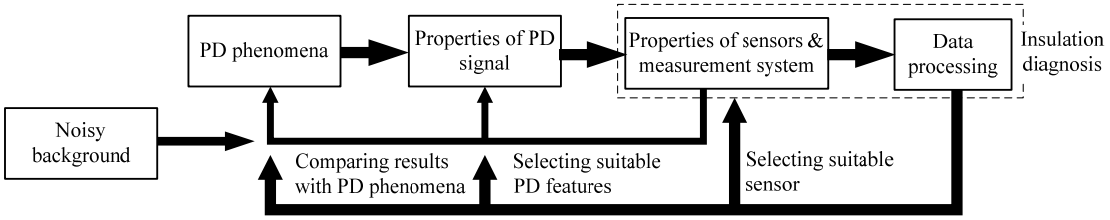

Fig. 2-1 Procedure of PD-detection based insulation diagnosis (after Hikita *et al* [4])

2.2 PD phenomena

Proper design and development of PD sensing and measurement circuit entail a certain degree of understanding of the PD phenomena [5]. According to IEC 60270 [6], "Partial discharge (PD) is localized electrical discharge that only partially bridges the insulation between conductors and which can or cannot occur adjacent to a conductor. They are in general a consequence of local electrical stress concentrations in the insulation or on the surface of the insulation. Generally, such discharges appear as pulses having a duration of much less than 1 μs. More continuous forms can however occur, such as the so-called pulse-less discharges in gaseous dielectrics." For a better understanding of PD phenomena, the fundamentals of PD measurement, the details of PD mechanism, electrical breakdown in different insulations, and PD types are reviewed.

2.2.1 General mechanisms

To start a PD, two conditions must be fulfilled. First, the electrical field must be large enough and have the suitable distributions to develop a self-sustaining discharge which will generally translate into a minimum breakdown voltage $V_{\min}$. Second, a free electron must be present at a suitable position to start the ionization process. This "starting" electron may be supplied by external sources, field emission or by capturing electrons from previous PD activities. Since the present of electron is a stochastic event, the time lag t_L varies in different cases. Here, the time lag t_L is the duration before PD occurrence and after voltage exceed $V_{\min}$ [7]. Fig.2-2 illustrates the voltage variations when PD occurs.

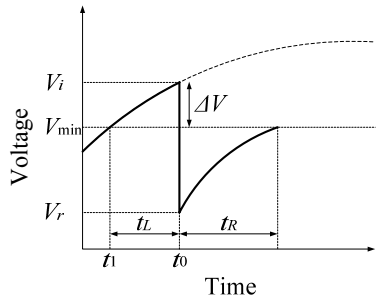

Fig. 2-2 Variation of voltage across the local insulations where a PD occurs (after Peter *et al* [7])

In Fig.2-2, the voltage across the local insulations such as a void for cavity discharge exceeds the minimum breakdown voltage at t_1, but PD occurs at t_0 only when a free electron is available. Therefore, it is possible that the voltage may already exceed the minimum breakdown voltage when PD occurs. In Fig.2-2, the PD ignites at voltage

$V_i = V_{\min} + \Delta V$. When PD happens, the voltage drops from V_i to Vr rapidly and thus result in an impulsive waveform of current.

2.2.2 Electrical breakdown

Generally, the PD detection mentioned in early research includes only pulse type discharges. In 1933, the oscillograph method have been reported to be employed in PD detections [5]. This technique can only detect the pulse type discharges. In fact, other types of discharges such as pulseless glow and pseudoglow discharges can also occur. Although in most cases, the occurrences of pulseless glow and pseudoglow discharges are accompanied by the pulse type PDs, it must be noted that the pulse type discharge cannot represent the full extent of intensity of PD present [5].

A) Ionization ways

After a long study on effect of overvoltage on the dielectrics and PD mechanism, Devins introduced two PD mechanisms named 'streamer-like' and 'Townsend-like' discharges [8].

According to reference [5], the streamer-like discharge theory is developed by Rather [9] and Meek [10], independently. The term 'streamer-like' was used to describe pulses with rapid rise-time short-durations. Occurrence of this type of discharges is related to presence of large gaps. The streamer-like discharges propagate rapidly across the gaps due to ionizing photon radiation at the streamer tips or leads. It is independent of cathode emission and dependent on photo ionization in the gas volume. Its occurrence can account for the relatively short breakdown time of the gaps.

On the other hand, the Townsend-like discharges are cathode emission-sustained discharges. They have essential features in contrast with streamer-like discharges. The typical Townsend-like pulses often happen in short gaps. They appear in different forms: slow rise time spark-type pulses, true pulseless glow and pseudoglow [11, 12]. The true pulseless glow discharge consists of weak ionization. However, pseudoglow discharge which is similar to true glow discharge in ionization may generate minute discharge pulses. These minute pulses can be detected electronically or optically by means of a photomultiplier [13-15]. It is agreed that presence of oxygen gas is an activator of sparks. It can be accounted from the predominance of spark or pulse type discharges in air and

the transition from a glow discharge process to pulse discharge type behavior in other gas (helium, neon) [5].

B) Breakdown in gas

Gas is the simplest and most commonly found dielectrics. In most cases, air is used as the insulating medium, and in some cases, other gas such as sulphur hexafluoride (SF_6) [16], nitrogen (N_2) [17] and carbon dioxide (CO_2) [18] are also used. There are usually two types of electrical discharges in gas: non-sustaining discharges and sustaining discharges. The spark breakdown is actually the transition of a non-sustaining discharge into a self-sustaining discharge [19]. Both 'Townsend-like' and 'streamer-like' discharges occur in gas under different conditions. In uniform electrical field, the breakdown voltage of a gap is a unique function of the product of pressure and the electrode separation for a particular gas and electrode material, which is known as Paschen's law [20]. On the other hand, in non-uniform electrical field, the discharges may appear at points with highest electric field intensity, namely at sharp points or where the electrodes are curved [19].

C) Breakdown in liquids

Liquid dielectrics are mainly used for filling transformers, circuit breakers, or as impregnants in high voltage cables and capacitors. Rather than being just a dielectric, liquid dielectrics have additional functions in which they act as heat transfer agent in transformers and as quenching media in circuit breakers [21]. The breakdown theories in pure liquid and impure liquid are often discussed separately. Various methods were employed to purify liquid such as filtration [22], centrifuging, degassing and distillation [23]. The breakdown in pure liquid is similar with 'Townsend-like' ionization in gases. The breakdown voltage depends on many factors, for instance, the field, gap separation and so on. Different from the limited knowledge of breakdown in pure liquid, the theories of breakdown in impure liquid has been well explained according to the causes: suspended particles, cavities and bubbles, and stressed oil volume. No matter which kind of causes, the actual mechanism of breakdown in liquid is a complex phenomenon and the breakdown voltage was reported to be determined by experimental investigation only [21].

D) Breakdown in solids

Solid dielectric is widely used in all kinds of electrical circuits and devices and insulates the parts operating at different voltages. Solid dielectrics have properties of good dielectric such as low electric loss, high mechanical strength, free from gaseous and moisture, and so on. They have better performance than gas and liquid [24]. However, the solid dielectrics cannot recover their electrical strength after discharge, while the dielectric strength of gas is fully recovered and of liquid is partially recovered.

The breakdown procedure in solid dielectrics is complex and depends on the duration of applying voltage. According to the duration of voltage application, the breakdown in solid can be divided into intrinsic breakdown, electromechanical breakdown, and thermal breakdown [24]. Intrinsic breakdown relies on the presence of free electrons. When the electric fields and temperatures exceed certain ranges, more and more electrons participate in the conduction procedure and lead to breakdown [25]. The electromechanical breakdown occurs when the electrostatic compression forces exceed the mechanical compressive strength. The compression forces arise from the electrostatic attraction between surface charges which appear when voltage is applied [20]. When an electrical field is applied to a dielectric, conduction current, no matter how small it may be, flows through the materials. Finally, the current heats up the specimen and results in the temperature rise. The thermal breakdown appears when the generated heat exceeds the heat dissipated [24].

2.2.3 PD types

The term ‘partial discharge’ includes a wide group of discharge phenomena: (i) continuous impact of discharges in solid dielectrics forming discharge channels (treeing); (ii) internal discharges occurring in voids or cavities within solid or liquid dielectrics; (iii) surface discharges appearing at the boundary of different insulation materials; (iv) corona discharges occurring in gaseous dielectrics in the presence of inhomogeneous fields [26]. Among these types, the discharges in trees can be considered as a typical kind of internal discharges [27].

A) Tree discharges

Among all the discharges, the PD caused by treeing is the dominant cause factor of insulation failure [28]. Treeing is the formation of branches, which are the channels of

discharged current, in insulation material under high electrical potential. For better understanding of PD, it is necessary to know more about treeing.

The phenomena which was first described as treeing in dielectrics was in 1912 [29]. Treeing is a type of pre-breakdown deterioration that has the general appearance of a tree-like path in the insulation wall. It is a damaging process due to partial discharges and progresses through the stressed dielectric insulation, whose path resembles the form of a tree. The tree often develops in line with electrical field direction.

Two kinds of treeing are usually discussed: electrical and water treeing. Water treeing is a very complex phenomenon involving electrical, physical and chemical mechanisms while electrical treeing seems to be basically a comparatively 'simple' phenomenon involving mainly purely electrical phenomena and gas discharges in a solid dielectric [30]. However, there are instances where a water tree can develop into an electrical tree, thus producing PD pulses. Although the two trees are related to each other, they are usually discussed separately since they have their own characteristics.

1) Electrical treeing

Electrical treeing is the first degradation mechanism experienced with dielectrics of HV apparatus. Its initiation under impulse voltage load occurs at higher voltage levels in the case of negative potential applied to the electrode [30]. Electrical treeing is observed to originate at points where impurities, gas voids, mechanical defects, or conducting projections cause excessive electrical field stress within small regions of the dielectric. An impurity or defect may even result in the partial breakdown of the solid dielectric itself [31]. Thus, one of the best improvements is to keep the apparatus clean and any surface boundary smooth. Practically, depending on different voltage magnitudes, different internal electrical stresses are generated, and different types of trees are thus formed: tree-like, brush-like or chestnut-like trees. The tree-like channel usually contributes most in PD formation. However, in contrast to the initiations of trees, the development of trees is strongly temperature dependent [30].

2) Water treeing

Water tree deterioration was first observed in 1967 [32]. It is a major cause of long-term deterioration in electrical HV equipment. Same as electrical trees, water trees grow from impurities either in the form of cavities or in the form of foreign material imbedded in the

insulating material in the presence of moisture and AC electric field. Both of the two trees are formed when bounds are cleaved. However, the water treeing can happen at much lower electric field because there are no hot electrons. Unlike electrical treeing, the precipitation of water is the main reason of water treeing. If there are foreign water soluble inclusions such as tiny salt particles in insulation materials or the relative humidity of the insulation materials exceeds the saturation pressure, a water droplet will be formed [30]. These water-filled cracks are favorable starting points for water trees [33]. They are diffused structures with a bush or fan like appearance [34]. Water trees are known to disappear on drying of the insulation and reappear on rewetting. They grow without detectable partial discharges. Once formed, water trees become partially or completely invisible upon drying of the material, but can be made visible again by exposure to (hot) water without applying an electric field [35].

B) Cavity discharge

Cavity discharges occur in inclusions of low dielectric strength. Usually, these are gas-filled cavities, but oil-filled cavities can also breakdown and cause gaseous discharges subsequently [27]. The inception voltage depends on the stress in the cavity and the breakdown strength of the cavity which is related to the cavity shape and location. Four kinds of cavities of internal discharges are portrayed in Fig.2-3.

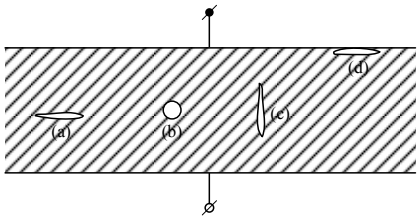

Fig. 2-3 Cavities of internal discharges, (a) flat cavity, (b) spherical cavity, (c) sharp cavity, (d) flat cavity near the surface

The field stress in the flat and spherical cavities is less than that of the other two cavities in Fig.2-3(c) and (d) which can be destructive as they cause considerable field concentrations after breakdown [27].

C) Surface discharge

Electrical discharges can occur in gas, liquid or vacuum in close vicinity to a solid dielectric surface. Such a discharge is known as a surface discharge. As a result of the proximity of the discharge plasma to the surface, considerable degradation of the solid dielectric surface can occur [36]. This kind of discharge applies to bushings, ends of

cables, the overhang of generator windings and where a discharge from outside touches the surface. Common forms of discharges are shown in Fig.2-4 [9].

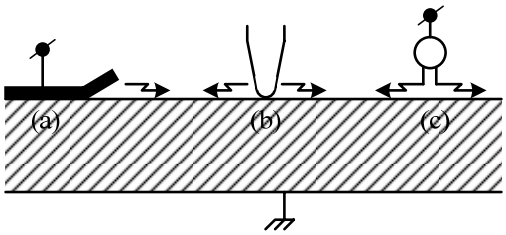

Fig. 2-4 Surface discharges, (a) plane-plane configuration, (b) sharp edges to plane, (c) spherical edge to plane (after Kreuger [27])

D) Corona discharge

In uniform field and quasi-uniform field gaps, the onset of measurable ionization usually leads to complete breakdown of the gap. In non-uniform fields various manifestations of luminous and audible discharges are observed long before the complete breakdown occurs [26]. This phenomenon is called 'corona'. They usually occur at the high-voltage side, but at sharp edges at earth potential, or even at half-way of electrodes also may cause corona discharge [27]. Fig.2-5 shows the corona fundamental schematically.

Fig. 2-5 Corona discharges

It must be noted that 'corona' is a form of partial discharge that occurs in gaseous media around conductors which are remote from solid or liquid insulation. 'Corona' should not be used as a general term for all forms of PD [26]. For the electrical equipment insulated with solid dielectrics, corona is often classified and recognized from cavity discharge which is the cause of insulation degradation [37, 38].

2.3 Properties of PD signal

PD signals occur in the form of individual or series of electrical pulses. Review on the equivalent circuit of single PD pulse and the reoccurrence of pulses in PD series is helpful in PD measurement which involves the capture, storage and processing of those PD pulses [39].

2.3.1 Equivalent circuit

It is convenient to analyze the properties of PD signal by using its equivalent circuit. Except for corona discharges whose breakdown path is limited by space charges, the behaviors of all other types of PD under AC voltage can be described using the well-known a-b-c circuit in Fig.2-6(b). This configuration was first suggested by A. Gemant and W. V. Phillippoff in 1932 [40] and then discussed by Whitehead [41] and Kreuger [27]. Although this representation is simplified and does not allow for complex factors, it is quite helpful in understanding the pulse voltage variations.

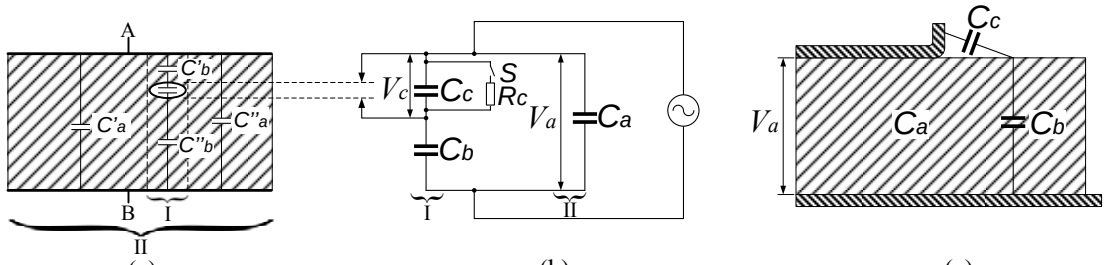

Fig. 2-6 Equivalent circuit, (a) dielectric circuit of internal discharge, (b) equivalent a-b-c circuit, (c) dielectric circuit of surface discharge (after Kreuger [27] and Kuffel [26])

Fig.2-6 takes the internal discharge as an example to explain this a-b-c circuit. In Fig.2-6(a), a simple discharge sample is sketched with the dielectric materials between the two electrodes A and B, and a gas-filled cavity. The defective part corresponds to I, and the non-defective part corresponds to II. In the defective part, the dielectric starting and ending at the cavity walls form two capacitances C'_b and C''_b which are combined into a capacitance $C_b = C'_b C''_b/(C'_b + C''_b)$. The capacitor of the cavity is represented by a capacitance C_c. The switch S is controlled by the voltage V_c across the cavity capacitance C_c. It is only closed for a short time, during which the PD current takes place. The resistance R_c simulates the magnitude of PD current. The PD current which is caused by the discharge of cavity capacitance C_c cannot be measured directly. It has a shape similar to Dirac function and a short duration in the nanosecond range. All the non-defective part is represented by a capacitance C_a which is connected in parallel with the faulty part. From Fig.2-6(a), the capacitance $C_a = C'_a + C''_a$. Due to the cavity dimensions, the magnitude of capacitances are always constrained by the inequality [42]

$$C_a >> C_c >> C_b \tag{2-1}$$

Fig.2-6(c) shows the dielectric circuit of surface discharge. The surface discharge can also be described by the a-b-c equivalent circuit. The discharge is represented with a

capacitance C_c on the surface. The rest of the surface which are made of insulation materials are represented by a capacitance C_b. All the other non-defective parts are represented by a capacitance C_a.

This a-b-c circuit is also applicable for treeing discharges whose dielectric circuit is not shown in Fig.2-6. All the branches correspond to capacitance C_c and the capacitance C_b represents the non-defective part of the dielectric in series with the tree. The rest of the non-defective part is again represented by the capacitance C_a.

2.3.2 Reoccurrence

As aforementioned in Section 2.2.1, the PD could only occur when the requirements on electrical field and free electron are both fulfilled. After the first ignition of PD, the discharge activity will deposit electrons which make the reoccurrence of PD more possible. The detail of PD reoccurrence is shown in Fig.2-7 with the assumption that free electrons are present and PD occurs at breakdown voltage.

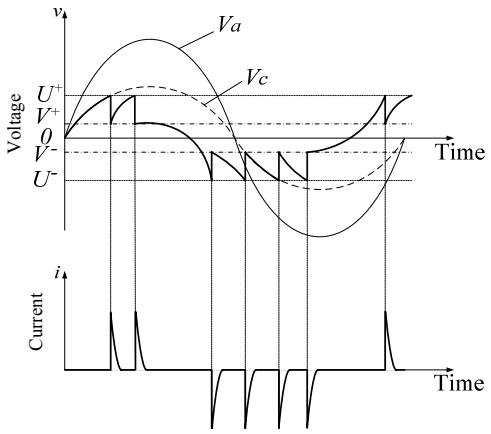

Fig. 2-7 Reoccurrence of PDs

According to the a-b-c equivalent circuit in Fig.2-6, V_a is the voltage across the dielectric and V_c is the voltage across the breakdown path. Also, U^+ and U^- are the breakdown voltage of positive and negative cycles. Their values are usually dependent on the kind and size of insulations such as gas, liquid and solid. V^+ and V^- are the voltages where PD extinguishes in positive and negative cycle, respectively. When the voltage V_c reaches the breakdown voltage U^+, a PD occurs. Immediately, the voltage V_c drops to V^+. This voltage drop takes much less than 1μs which is extremely short when compared with the

duration of a power frequency cycle that last 20 milliseconds[6]. After the PD distinguishes, the voltage V_c increases and reaches U^+ again. Then, a new PD occurs. This occurrence of PD will repeat until the voltage V_a decreases. However, when the voltage V_c decrease with the voltage V_a and reaches the breakdown voltage U^- in negative half-cycle, a new group of PDs occur.

Theoretically, the reoccurrence of PD is a regular sequence of pulses in both half cycles. However, the discharges may become intermittent if the breakdown voltage U^+ and U^- are not equal. Also, if there are more than one PD sources, irregular reoccurrence of PD will be observed [27].

2.4 PD Measurement

Instrumentation is one of the most important aspects of PD analysis. Without an accurate, fast and objective measurement of discharge pulse, no further analysis can be done. Although some detection methods have been employed in PD detection of electrical apparatus and cables, the pursuit of more accurate measuring method is still in progress [43, 44].

To be recognized as PD, the current induced in the external conductor must be large enough to be detected and occurs frequently enough to be recognizable as something other than noise. The discharge is a process of energy dissipation. The detection of PD is based on this energy exchange of discharges. These exchanges can be manifested as electrical current, dielectric loss, magnetic radiation, optical radiation, sound, increased gas pressure, and chemical reaction [35]. These different kinds of characteristics are detected to describe the PD occurrence. In the area of PD measurement, there are two main types of techniques: non-electrical measurement and electrical measurement.

2.4.1 Non-electrical measurement

Non-electrical methods are based on the non-electrical properties to detect the occurrence of PDs. The most common methods are the acoustic, the optical and the chemical methods.

A) Acoustic PD measurement

It is a very old but traditional way to detect PD by the human ear. There will be an elastic wave during the energy discharge in a solid structure. The discharge of a few hundred pico-coulombs can be heard by experienced workers in a quiet environment. However, this method is ineffective when there is a modest amount of background noise.

In recent times, acoustic emission (AE) technology is considered to be effective for non-destructive testing and material evaluation [35]. Usually, a narrow band of about 30~50Hz is chosen, in which noise is less and unimportant. Excellent results were obtained by detecting ultrasound [45-47]. The most important part of this method is integrated pre-amplified sensor [48]. Some ultrasound detection equipments are available in practical use.

Acoustic method is a useful tool in PD test. It can detect unexpected surface discharges. Large amount of PDs can also be detected by this method. By moving the ultrasound detector along the sample, a reasonable area of location can be covered [48]. However, the acoustic method cannot measure the magnitude of PD pulse, and its sensitivity is low since only pulses above 100pC can be detected.

B) Light detection

The most common optical methods are mainly based on two PD characteristics: electroluminescence (EL), and light emission caused by PD.

Since EL happens before the insulation degrades, it is a powerful technique to detect degradation before electrical treeing. It has been employed successfully in some practical cases [49, 50]. As EL is just an energy discharge before treeing and PD, it does not cause erosion in insulation. The degradation can be stopped before the insulation breakdowns. However, a high EL inception does not necessarily imply an eventual growth of tree. A long-time monitor and test are needed.

Different from EL, light emission implies the initiation processes which may lead to PD cavity and trees [51]. In the measurement of light emission, the parameters that often used are: luminous quantity, length [52, 53], intensity [54], area growth [55], radial extent [56], and fractal dimension [57]. Light detection is an excellent method for PD detection and location detection with a good sensitivity. It is unaffected by electrical disturbances in the circuit. Photographic records can also be captured even with consumer-grad cameras. At

least 1pC can be attained in light detection. However, this method has its own drawbacks. Besides the common shortcomings of all non-electrical methods in that PD magnitude is not easily measured, light emission detection can only detect discharges in transparent materials which are visually accessible from the outside. Opaque materials are not suitable for light detection methods [35].

C) Chemical detection

As PD occurrence could generate chemical by-products in insulating materials, detection of those chemical by-products is another way to check for PD existence. The predominant application of chemical detection is for oil and cellulosic materials, although new techniques have been developed to assess the state of composite insulators in gas insulated switchgear (GIS) or aged cable [58]. Currently, the most popular chemical detection methods are dissolved gas analysis (DGA) and Furan measurement.

Dissolved gas analysis (DGA) is probably the most widely used chemical detection for oil-insulated apparatuses such as power transformers. The in service apparatus undergoes various thermal, chemical, electrical stresses which causes the generation of some gases. The test which is conducted to analyze these dissolved gases is known as DGA Test [59]. The dissolved gases commonly detected are hydrogen (H_2), methane (CH_4), acetylene (C_2H_2), ethylene (C_2H_4), ethane (C_2H_6), carbon monoxide (CO) and carbon dioxide (CO_2). They are quantified in extremely low level in parts per million (ppm). Their values vary from a few ppm to thousands of ppm [60].

Utilizing liquid chromatography for furans in insulating oil is also a popular way of chemical detection. The level of concentrations of different forms of furans can indicate the deterioration of insulations at normal operating temperatures [40].

A major restriction to applying chemical detection is the lack of access to the insulating materials built into the electrical equipment. For example, it might be possible to obtain oil samples from the tank of oil-insulated transformers. However, the extraction of solid insulation is not possible because of its resulting damage to the insulation structure [40].

2.4.2 Electrical measurement

Due to the limited applications of non-electrical methods, electrical measurement methods which can be used in more kinds of insulations are more popular. The early

electrical methods used analogue instruments. Such instrument performed adequately well for cases with PD inception and extinction with voltage. Then crystal controller and A/D converters were employed in PD analysis. With the advent of personal computers in 1980s, the PD measurement system shifted from hardware instrumentation to software dominated techniques. The 1990s saw the introduction of rapid response digital circuit for PD measurement application [5]. Nowadays, the electrical measurement is used in three basic approaches by using coupling capacitor, UHF coupler, or TEV sensor, respectively.

A) Coupling capacitor method

Coupling capacitor method is a traditional way to detect PD. As described in IEC 60270, this method utilizes a coupling capacitor that is placed in parallel or in series with the test object, and the discharge signals are measured across external impedance, and then filtered by detectors which are usually band pass filters.

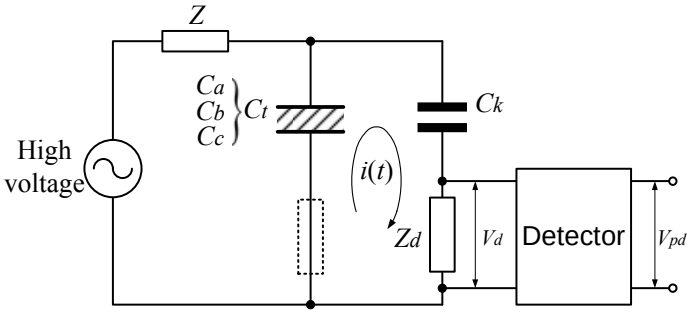

Fig. 2-8 Basic diagram for coupling capacitor method

A typical PD detection circuit with coupling capacitor is shown in Fig.2-8 where the PD sample is represented by C_t and the coupling capacitor is represented by C_k. The PD sample C_t can be modeled by the a-b-c circuit in Fig.2-6(b). When PD occurs in the cavity of C_t, the switch S is closed and the discharge current $i(t)$ releases a charge $\Delta q_c = C_c \Delta V_c$ from C_c. This charge will be absorbed by the whole system. By comparing the charges within the system before and after the discharge, the voltage drop ΔV_a across the sample C_t can be calculated as [42]

$$\Delta V_a = \frac{C_b}{C_a + C_b} \Delta V_c \tag{2-2}$$

On the other hand, the impedance Z_d disconnects the coupling capacitor C_k and test sample C_t from the HV source during PD phenomenon. Therefore, the C_k acts as a

storage capacitor and release a charging current or the actual "PD current pulse" $i(t)$ between C_k and C_t. If $C_k >> C_t$, the charge transfer is

$$q = (C_a + C_b)\Delta V_a \tag{2-3}$$

From equation (2-2), the charge becomes

$$q = C_b \Delta V_c \tag{2-4}$$

Here, q is the "apparent charge" of a PD.

If no stray capacitances are presented in parallel with the branch of C_k and C_t, and no PD leakage currents occur across the HV supply, the PD current $i(t)$ would be the same in both branches of C_k and C_t, but of opposite polarity. Therefore, the coupling or measuring impedance Z_d which is used for measuring PD current $i(t)$ can either be connected in series with coupling capacitor C_k or in series with the sample C_t which is indicated in Fig.2-8 by a dotted line [27]. In practice, the series connection of Z_d with PD sample C_t has higher sensitivities because the PD current $i(t)$ from C_t is best picked up by this connection. However, this series connection has disadvantages such as possible damage to the measuring circuit if test object C_t fails. Therefore, the detection circuit Z_d is usually placed in series with coupling capacitor C_k or in series with test sample C_t when coupling capacitor C_k is provided [42].

1) Impedance

When a PD occurs in C_t, an impulsive current is produced in the circuit in Fig.2-8. The impedance Z_d is thus needed to convert the impulsive current into an impulsive voltage which is easier to calibrate. As the PD measurement may be set up in well electromagnetically shielded laboratories as well as in industrial plants where interferences are unavoidable. The interferences such as harmonics can be reduced by using filtering circuits. Two types of filtering impedance Z_d are often employed: the RC circuit and RLC circuit. The connections of Z_d are shown in Fig.2-9.

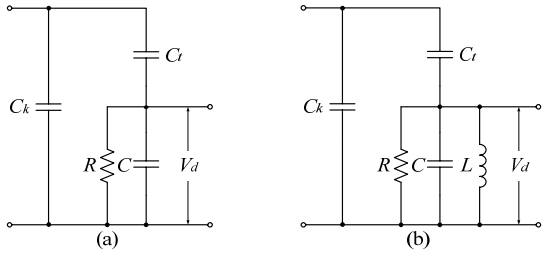

Fig. 2-9 Connections of Z_d , (a) the RC circuit, (b) the RLC circuit

When PD occurs in C_t , a voltage reduction across C_t is caused. As $C_k >> C_t$, the voltage drop can be reduced to $\Delta V_t = q/C_t$. Then the charge flows from C_k to C_t and C to equalize the voltage between them. As the PD current $i(t)$ is the same in the C_k - C_t loop, the voltage drop across impedance Z_d should be [61]

$$\Delta V_d = \Delta V_t * \frac{\frac{1}{C}}{\left(\frac{1}{C}+\frac{1}{C_k}+\frac{1}{C_t}\right)} = \Delta V_t * \frac{C_k C_t}{CC_k + CC_t + C_k C_t} = q/C_t * \frac{C_k C_t}{CC_k + CC_t + C_k C_t} = \frac{qC_k}{CC_k + CC_t + C_k C_t} \tag{2-5}$$

However, as the PD current $i(t)$ is an extremely short current pulse, the magnitude of ΔV_d may decrease with time. The waveform of V_d depends on the transfer functions of C_k - C_t loop with impedance Z_d. The transfer functions of RC-connected and RLC-connected circuits are expressed in the form of Laplace transform as

$$G_{RC}(s) = \frac{1}{m}\frac{1}{s+1/Rm} \tag{2-6}$$

$$G_{RLC}(s) = \frac{1}{m}\frac{s}{s^2 + s/Rm + 1/mL} = \frac{1}{m}\frac{s}{(s+1/2Rm)^2 + 1/\omega_0^2} \tag{2-7}$$

where $m = C + C_t C_k/(C_t + C_k)$ and $\omega_0 = \sqrt{(1/Lm - 1/4R^2m^2)}$. The value of Rm controls the decay time of the voltage V_d and ω_0 determines the oscillating frequency of voltage V_d [39]. Therefore, the waveforms of V_d due to different impedances are calculated to be: [27]

$$V_d = \frac{qC_k}{CC_k + CC_t + C_k C_t} \times \exp(-t/Rm) \tag{2-8}$$

$$V_d = \frac{qC_k}{CC_k + CC_t + C_k C_t} \times \exp(-t/2Rm) \times \cos\omega_0 t \tag{2-9}$$

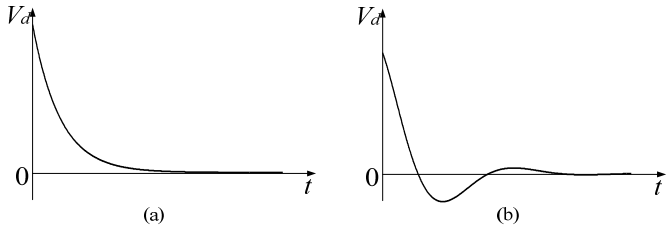

Fig. 2-10 Waveforms of voltage across impedance Z_d , (a) RC circuit, (b) RLC circuit

The waveforms of voltage V_d are shown in Fig.2-10. As the PD current is similar to a Daric function, the voltage V_d increases suddenly and then decreases gradually.

2) Detectors

A detector with a band pass filter is often employed to pick up the impulsive signal V_{pd}. The PD waveform V_{pd} is heavily influenced by the frequency response and bandwidth of the detector. Some parameters such as frequency bandwidth Δf and central frequency f_m were employed to characterize a detector [6].

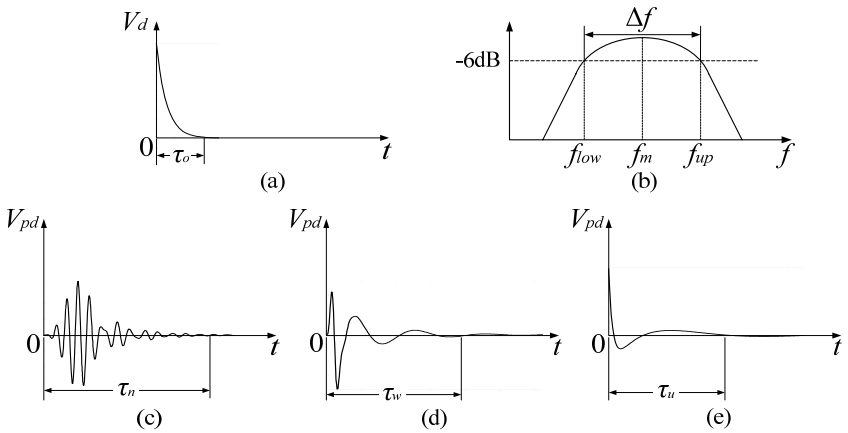

Fig. 2-11 Waveforms of V_{pd} , (a) original PD pulse waveform V_d , (b) frequency parameters of detector, (c) V_{pd} after narrow-band detector, (d) V_{pd} after wide-band detector, (e) V_{pd} after ultra-wide-band detector.

Fig.2-11(b) shows the frequency characteristics of a band pass filter. The lower and upper cut-off frequencies are represented by f_{low} and f_{up}, respectively. Bandwidth Δf is defined as the difference between f_{up} and f_{low} : $\Delta f = f_{up} - f_{low}$. Central frequency f_m equals $f_m = f_{low} + \Delta f / 2$. With different frequency parameters, the detectors are often classified into three fundamental types: narrow-band, wide-band and ultra-wide-band detectors. Different detectors have been applied under different conditions in the history of PD measurement. The damped exponential PD pulse V_d in Fig.2-10(a) is filtered by the various detectors to illustrate their different effects on V_{pd} .

a) Narrow-band detector

During the early days of PD measurements, some researchers tried to discriminate PDs from other interferences with some selective voltmeter which can be tuned within the frequency range of interest. This idea is similar to the radio frequency detection methods. Therefore, the measurement of radio influence voltage (RIV) which is the fundamental theory of narrow-band filter in PD measurement was first recommended by National Electrical Manufacturers Association (NEMA) in 1900's [62]. At that time, the RIV measurement was the only one available, which appeared to have the necessary characteristics for the measurement of low level, high frequency signals. This method has been subjected to many modifications in the past decades. Further, the large amount of data collected via RIV measurements on transformers has somehow added weight to the official recognition of RIV as one of the essential criteria during acceptance tests [63].

Nowadays, the term "narrow-band detector" is supposed to represent the band pass filter with a small bandwidth Δf and a central frequency f_m which can be varied over a wide frequency range, in which the frequency spectrum of the PD pulse is approximately constant. Recommended values for Δf and f_m are $9k\text{Hz} \le \Delta f \le 30k\text{Hz}$ and $50k\text{Hz} \le f_m \le 1\text{MHz}$, respectively [6]. A typical waveform of V_{pd} after narrow-band detector is portrayed in Fig.2-11(c). Here, frequency bandwidth Δf and central frequency f_m are set as 30kHz and 80kHz, respectively. It is noted that the impulsive PD waveform has changed to an oscillating waveform. The duration of a single PD pulse τ_o increases to τ_n which is usually larger than 100μs.

Based on the narrow-band detecting theory, some sensing devices already installed in electrical apparatus were adapted for PD measurement. The most famous method that applied resistive temperature detector (RTD) was proposed by Itoh *et al* [64] and Campbell *et al* [65]. The RTD sensor operates in the extended frequency range from 5MHz to 60MHz and a narrow-band data acquisition system is employed for phase resolved PD pulse height analysis [5].

However, with more and more research in PD detection, narrow-band detection reveals many problems. First of all, the reading of narrow-band detection is a complex function of the PD pulse amplitude and repetition rate. It is very difficult to find the relationship between reading and PD pulse features such as amplitude in pC [66]. Secondly, the resolution of narrow-band detection is extremely low. Take the RIV detector as an

example, the resolution is of the order of 30 to 40 impulses per quadrant. It is insufficient if the repetition of pulses is higher which may often indicates either a serious breakdown or existence of pulse-like interferences [27]. Moreover, the oscillatory discharge response generates errors in complex samples such as transformers and rotating machines. The PD pulses from transformers, rotating machines or cable have long travel length to the detector. Their high frequency content has been distorted and filtered out during the propagation. If those pulses are collected by narrow-band detectors, superposition errors will be induced which makes it very difficult to discriminate PDs from background noises.

Therefore, with such severe shortcomings, narrow-band detectors are no longer preferred by researchers and engineers although some such detectors are still in use. In some cases such as non-electrical PD measurement, the electrical narrow-band detectors are equipped to provide more information for PD measurement [42].

b) Wide-band detector

In last century, Franke *et al* pointed out that the shapes of PD pulses must be evaluated to determine the effect on both system data transmission and reliability objectives [67]. Further, the PD current was proven to have a rectangular shape with an extremely short duration and a resulting wide frequency spectrum extending to the region of 100 MHz which was the limit of measurement system at that time [27, 67, 68]. The narrow-band detector with bandwidth less than 30kHz and generating oscillating waveforms is apparently impossible to meet this requirement. To improve the accuracy in PD magnitude and resolution of nearly simultaneous PD pulses, wide-band detector was introduced.

According to IEC 60270 "High-voltage test techniques - partial discharge measurements", the term 'wide-band detector' is often used to describe the detectors with a bandwidth Δf between 100kHz and 400kHz. The lower and upper limit of wide-band detector are recommended to be $30 \text{ kHz} \le f_{low} \le 100 \text{ kHz}$ and $f_{up} \le 500 \text{ kHz}$, respectively [6]. Limited by technologies, the wide-band detector was a combination of several detectors at the early ages [67]. With the development of filter design, the wide-band detector was realized by a single current transformer (CT) with wide bandwidth [63].

The PD pulse detected by wide-band filter is in general a well-damped oscillating pulse. An example of that PD waveform is shown in Fig.2-11(d). To produce this waveform, the bandwidth Δf is chosen to be 400kHz. The lower and upper frequency limits are set as

50kHz and 450kHz respectively. From this waveform, the magnitude as well as polarity of PD pulse is easy to determine. The PD duration τ_w is smaller than τ_n of narrow-band detector and is commonly in the range of tens of microseconds. Although the rise time degradation and attenuation is still present, wide-band detector can response much better to the fast rise time of the PD pulse front when compared to narrow-band detectors. Accordingly, the resolution of wide-band detectors increases to at least 500 impulses per quadrant which is satisfactory for PD analysis.

With the improvement of detector bandwidth, the detected PD data became useful in more applications such as PD location and recognition. The most widely used PD location at that time was based on the comparison of the PD pulse magnitude at different sites [69]. The PD pulses with highest magnitude are assumed to be closest to the PD source, or to the nearest PD site in the case of a number of discharge sites. Further, with wide-band detector, more information about the number, magnitude, and phase angle of PD data was provided. With that information, the PD recognition technique had more development. The $\phi-q-n$ analysis was widely used in many international standards to identify PD types where ϕ, q, and n represent the phase, quantity and number of PD pulses, respectively[70].

However, as the bandwidth increases, some signals which are not PD pulses are also recorded. The most commonly encountered interferences in wide-band detection are the PD resonances with smaller magnitude. Also, some pulse-like interferences within the bandwidth of detector may introduce PD analysis errors.

c) Ultra-wide-band detector

After the first high frequency measurement done by Bailey who reported the rise time of PD pulse is around 1ns, the use of ultra-wide-band detection in PD measurement was adopted [71]. Since 1980's, widespread use of ultra-wide-band PD detection systems is observed for both laboratory experiments and field measurements on HV GIS, cables and generators [72].

Sometimes, the ultra-wide-band detectors collect PD pulses in nanoseconds region with a bandwidth up to 1GHz. Those studies were mainly designed to find a correlation between PD type on one hand and the characteristics in the nanosecond region on the other [27]. Generally, three main advantages are found in ultra-wide-band detection [73]: Firstly, the PD detection sensitivity increases when bandwidth extends to several hundreds of

megahertz. Secondly, it is much easier to reject external noises with ultra-wide-band PD signal. Furthermore, PD measurement with ultra-wide-band detectors has the ability to observe the shape of individual PD pulses directly. It helps to discriminate PD pulses from interferences, or from different sources.

Among all the ultra-wide-band detectors, Rogowski coil is usually regarded as the most widely used one. It has the advantage of measuring currents up to tens of microampere with no interference on the beam and without connections to high voltage conductors [74]. As it is not a closed loop, its connection is flexible. It can be wrapped around the live conductor. Further, it has an air core rather than an iron core which allows it respond to fast-changing currents. The PD pulses filtered by Rogowski coil was reported to have a fast rise time up to 1ns [75].

However, as it was not a long time after the ultra-wide-band detector was introduced into PD measurement, there is not any international standard that describes the range of its bandwidth. The detectors with a bandwidth larger than several megahertz are often regarded as ultra-wide-band detectors. An example of PD pulse filtered by an ultra-wide-band detector with bandwidth from 10kHz to 5MHz is shown in Fig.2-11(e). The pulse V_{pd} has a rise time as short as the original PD waveform V_d. When compared to the waveforms filtered by narrow-band and wide-band detectors, the waveform in Fig.2-11(e) is much less oscillating. The pulse duration τ_u follows the inequality $\tau_o < \tau_u < \tau_w < \tau_n$.

3) Advantages and drawbacks

The coupling capacitor method is a traditional PD detection circuit and recommended by many standards such as IEC 60270 [6]. It has been widely employed in the PD measurement for cable specimens [76], capacitors , transformers [77] and rotating machines [78]. With tens of years' of development, the coupling capacitor method is reliable and effective in PD detection. It can be applied in both on-line and off-line tests. Especially, a known pulse can be injected into the resonant circuit to calibrate the system. Furthermore, the optimum sensitivity can be obtained by maximizing the ratio of the coupling capacitance C_k to the capacitance of the test object C_a [79].

However, this method still has some drawbacks that need improvement. It can only identify discharges in a short isolated length. It has insufficient sensitivity for an electrical apparatus with large size due to the large capacitance of the entire equipment [79, 80].

Furthermore, its on-line application is limited by electromagnetic interference. Thus, this method is often used in off-line test which requires electrical shutdown.

B) UHF measurement method

The PD signal in metal-clad apparatus such as gas insulated switchgear (GIS) and oil insulated transformer has quite short durations which are usually less than 1ns. Therefore, the PD current radiates electromagnetic waves with energy spectrum extending to 2GHz [81]. Those electromagnetic waves propagate and reflect in the metal chamber. They indicate the occurrence of PD activities. Detection of those radiated electromagnetic waves which is normally called ultra-high frequency (UHF) method helps to monitor the insulation conditions of metal-clad equipments.

The UHF method was first proposed by Hampton in 1988 in UK [82]. The theory and fundamentals of UHF method were developed and improved by other researchers. The model of excited PD current on the inside surface of GIS chamber was set up by Judd *et al* [83] and Pearson *et al* [84]. Those models proved the effectiveness of UHF measurement and provided directional suggestions in further research. By comparing the frequency spectrums of PD signals and noises, the UHF couplers are often designed to have quite wide bandwidth and high central frequency [85]. The first practical application of UHF monitoring was on a gas-insulated substation (GIS) at the Torness nuclear power station in Scotland in 1986. Since then, UHF PD monitoring for GIS has evolved into a commercially viable on-line system that is providing real benefits to plant manufacturers and utilities around the world [85]. Besides the application in GIS, the UHF method is also introduced in monitoring transformers and good performance was reported [86]. To give a clear understanding, the theory of UHF signal, the couplers and the UHF data characteristics are reviewed.

1) Fundamental of UHF signal

It is well known that the PD pulses usually have quite short durations and those fast rising pulses have a wide frequency spectrum which typically extends 1GHz. Their occurrence can excite the metal chamber of GIS or oil-insulated transformer into various modes of electrical resonance. The resonances indicate the PD activity and are often detected by couplers which are located not far from the PD sources [84].

The propagation of UHF signals in metal-clad apparatus was studied by Farish and Judd [81]. Typically, without any barriers and discontinuities, the attenuation of signal up to 1GHz is only 3dB/km to 5dB/km. However, there are non-uniformities in the metal enclosure which cause partial reflections of UHF signal. Most discontinuities have a complicated reflection pattern because they do not reflect the signal at a plane but over a distributed volume. The partial reflection is the main reason of UHF signal attenuation.

Modeling the UHF signal via mathematical method is a helpful means of analyzing its propagation and extraction. In earlier years of UHF analysis, almost all the studies were focused on GIS. The structures of GIS are normally regarded as coaxial waveguide which includes inner conductor and outer chamber. The most familiar mode for coaxial waveguide is the transverse electromagnetic (TEM) mode which represents the electric and magnetic field in the plane transverse to the direction of propagation. At the same time, the higher order modes for sufficient high frequencies: transverse electric (TE) mode and transverse magnetic (TM) mode are also considered. The overall response to a PD current pulse is the superposition of all modes that make a significant contribution within the measurement frequency bandwidth [83]. Theoretically, the response is a function of PD current.

2) Frequency features of UHF measurement

As mentioned before, when electrical charges of PD move with anything other than a constant velocity, electromagnetic radiation occurs [87]. The electromagnetic wave radiates out from the PD site in all directions and propagates along the metal tank surface.

On the other hand, the arcing interferences associated with poor contacts, tap-changer mechanisms, floating objects at high potential will also radiate electromagnetic transient waves. In the confined space of metal enclosure, the electromagnetic waves reflect and resonant at the conducting surfaces and dielectric interfaces. The electromagnetic waves from interferences propagate along the metal chamber with similar pattern as that of PD signals. They could also be detected by the UHF couplers which are located at different sites of metal enclosure. How to filter out the interferences is a crucial problem in UHF measurement.

Different from interferences, the PD pulses, especially the internal PD pulses, have extremely short duration. Some researchers such as Judd [85] pointed out the due to the inverse relationship between time-domain and frequency-domain representations of a

pulse (Heisenberg uncertainty principle), shorter PD current pulses have more spectral energy at high frequencies. An example of two exponential damped PD current pulses with same charges but different durations is shown in Fig.2-12 to illustrate the frequency spectrums of different pulses.

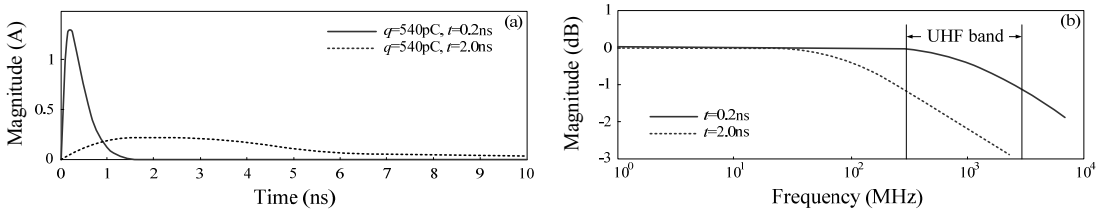

Fig. 2-12 Comparison of two current pulses, (a) waveforms, (b) frequency spectrums (after Judd [85]).

As portrayed in Fig.2-12, the two pulses have same amount of charge, $q = 450pC$, but different durations. The pulse with shorter duration has a wider frequency range than the other pulse that lasts longer. Therefore, in the UHF band, the PD pulses with shorter durations may contain larger energy than interferences and the signal-to-noise ratio (SNR) can be improved greatly. This can be demonstrated by the example shown in Fig.2-13, which displays the results of a PD signal filtered by high-pass filters with different cut-off frequency: 10MHz and 500MHz, respectively [88]. It is obvious to find the magnitude of impulsive noise is much greater than PD pulses when the cut-off frequency is only 10MHz. However, as in Fig.2-13(b), the noises are significantly attenuated after filtering and the SNR is also improved greatly.

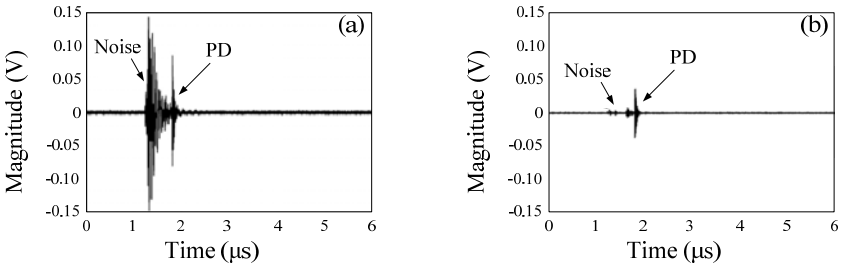

Fig. 2-13 PD signal obtained by different frequency bands, (a) PD with 10MHz high-pass filter (b) PD with 500MHz high-pass filter (after Fabiani *et al* [88])

The above discussion illustrates that UHF measurement is effective in PD detection. Usually, the UHF band is defined as the interval from 300MHz to 3GHz [81].

3) UHF couplers

In UHF method the PD is detected by couplers. According to the location of couplers in electric apparatus, two kinds of UHF couplers are often used: internal couplers, and external couplers that include windowed coupler and barrier coupler. Fig.2-14 portrays

how the couplers are fitted to the different locations of the clad of GIS [89]. The detailed locations of each kind of couplers are also portrayed.

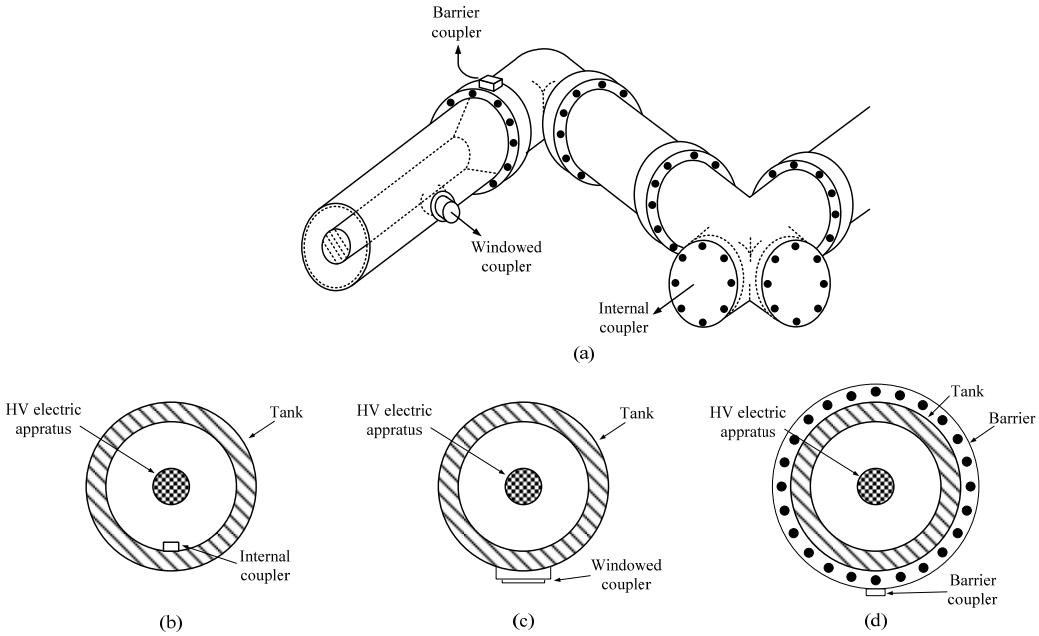

Fig. 2-14 Examples of mounting locations for UHF couplers on a GIS, (a) different locations of couplers, (b) internal coupler, (c) windowed coupler, (d) barrier coupler

a) Internal coupler

As the UHF signal resonates and propagates along the inside surface of the metal-clad, a UHF coupler located inside the clad could have high sensitivity. With the screen of metal-cladding, minimal interferences will be detected by the internal coupler. For the metal-clad apparatuses, internal coupler is a better choice than external couplers such as windowed coupler and barrier coupler. Because their locations are inside the metal enclosure, the design of internal couplers involves a compromise between the conflicting requirements of minimizing the field enhancement, while maximizing the UHF sensitivity and optimizing the match with measurement system [90]. The coupler should not create an additional risk of breakdown, and is normally mounted in a region of relatively weak HV field. However, this deployment often results in weak field which is hard to detect.

Therefore, various kinds of internal couplers are designed to fulfill the requirements of both high sensitivity and low field enhancement. The most mentioned designs of UHF couplers are circular plate [85], spiral [91], and monopole sensor [92, 93]. To avoid degrading the insulation system, internal couplers must be relatively large, with smooth faces and adequate radii on all edges that are exposed to the power frequency field [89]. Among those internal couplers, the monopole with a good sensitivity is unacceptable as it is more likely to induce breakdown [91]. Circular plate couplers have been proved useful,

and are more readily accepted in metal-clad apparatuses because they are similar to capacitive dividers and can be designed not to cause stress enhancement [90].

b) External coupler

Internal coupler which is mounted inside the metal-clad is the main approach for UHF based PD measurement. However, the internal coupler is mostly applied to new or refurbished equipment. For operational equipments, arranging an outage specifically to fit internal couplers rarely can be justified. However, older apparatus can particularly benefit from continuous UHF monitoring and this can sometimes be achieved by using external couplers to detect UHF signals in the metal cladding [89]. When external couplers are fitted to an in-service equipment, their placement is restricted by the availability of existing mounting locations, which may not provide adequate coverage of the electrical apparatus such that the coupler itself is unlikely to affect the insulation level [89]. However, this kind of couplers is remote from the UHF field inside the clad. It suggests that they may have a lower sensitivity than internal couplers. However, the external couplers were reported to cover a wider frequency band and have a good average sensitivity [89]. Due to the different locations of external couplers, they are classified as windowed coupler and barrier coupler.

Windowed coupler:

The theory of windowed UHF coupler was first proposed by Judd *et al* in 1997 [94]. The windowed coupler is usually designed to fit the window or openings on the metal enclosure of electrical equipment and detects the radiating electromagnetic waves. As an external coupler, the windowed coupler itself is unlikely to affect the insulation rating. Thus, it results in more freedom in designing sensor structures that are better suited to broadband reception and compensate for the less favorable coupler mounting position [91]. Furthermore, the windowed coupler can still be completely screened, making the sensing element electrically internal, as shown in Fig.2-15 that illustrates the diagram of a typical windowed coupler [89]. The UHF coupler is mounted in the window housing on the metal enclosure and is covered with metallic layer that ensures the coupler body be in good electrical contact with the window housing.

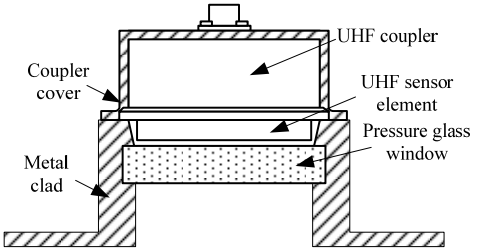

Fig. 2-15 Detailed diagram of windowed coupler

Barrier coupler:

Besides windowed coupler, barrier coupler is another commonly used external coupler. Since the PD pulse has an extremely short duration and rise time, its frequency spectrum usually extends to the range of gigahertz. As a consequence, a proportion of the PD energy radiates into the free-space through the joints in the metal-clad, giving rise to the fact that PD can be detected using radiometry [95]. Cast resin barrier, which is the jointing parts without sufficient metallic coverage, is an ideal electrical aperture in the cladding to radiate EM waves of PDs.

Because of the external location of barrier coupler, it is unlikely to affect the insulation ratings of HV equipment. On the other hand, its structure becomes an area of research to produce sensitivity as high as possible. Usually, the barrier coupler has the shapes of horn [96], biconical [97], log-periodic [98], loop [99], monopole or dipole antennas [100] and so on. Although a variety of barrier couplers are used and good results have been achieved, the optimal design of UHF antennas is still an immature technique and requires more investigations [101]. Besides the shape of barrier couplers, their size should also be considered carefully because it should neither be too small to ensure their resonant frequencies fall within the band between 300MHz to 3GHz, nor too large due to the space limitations of some electrical equipments [102].

4) Advantages and drawbacks

UHF measurement is very popular in PD detection of metal-cladding apparatuses, especially in GIS. The most prominent benefits of applying UHF are in two aspects:

First, PD location is possible to be determined if a large number of UHF sensors are mounted at different locations of the metal clad, no matter which kind of coupler is used. As the rise time of UHF signals is heavily influenced by the propagation environment, the physical location of PD sources can be found by analyzing the front part of PD waveform

that represents the direct path of the signal from the source before any reflections occur [95].

Second, continuous and remotely operated monitoring is possible. Continuous monitoring system can produce very large quantities of data for more accurate diagnosis, and it is less likely to overburden the engineer with its interpretation. Further, the PD data can be analyzed remotely and automatically, and the engineers are informed only when some condition arises which needs their attention [90].

However, the PD energy in UHF range is dependent on its propagation path. High frequency energy often attenuates greatly during propagation. The PD energy will be very small if there are any spacers or barriers. Thus, the waveform and energy peak of detected PD pulses would be totally different from its original status. No charge quantity information (electric quantity in pC) is delivered by UHF measurement [103]. Also, due to the small magnitude of UHF signals, noises cannot be ignored, even white noises from background. Furthermore, the external couplers are not as sensitive as internal ones. However, for the equipments without pre-installed internal couplers, it is impossible to shutdown the equipment for coupler installations.

C) TEV method

The investigations of coupling capacitor method and UHF method show that both of them have some drawbacks, for example, requirements of electrical shutdown, or absence of continuous online monitoring. To overcome those drawbacks, TEV method was proposed. The theory of transient earth voltage (TEV) was first proposed by Jennings and Collinson from EA Technology [3]. They pointed out that the radiating electromagnetic wave from PD source induces and forms a small pulse-like voltage on the metal tank surface. In TEV method, a non-intrusive sensor is mounted on the outside surface of metal-clad apparatus, normally next to the cable box or cable termination. This installation is much convenient than coupling capacitor method and UHF measurement and is preferred by increasing number of researchers and engineers. However, the metal tank acts as a receiver and amplifier of external interferences, and the influence from noises cannot be ignored. The PD extracting ability can be enhanced by using de-noising techniques. As a newly developed technique, many issues of TEV measurement are not well addressed yet and more investigations are needed. Since in this thesis, the research focus is on detection by TEV methods, details of TEV-based PD measurement are separately introduced in Chapter 3.

2.5 Noise rejection

Because of the external locations of many PD sensors, noise is always a major problem. The review of types and features of noises, and existing popular noise rejection ways via hardware and software are essential for developing reliable PD measurement technique.

2.5.1 Major noises in PD measurement

Noise can be due to several kinds of sources and can couple with the systems in different ways and with different features. Therefore, noise rejection has no omnipotent solution and is best approached by devising several techniques, each of them tailored for a specific kind of noise [104]. To develop suitable tools for each kind of noise, the noise types and features should be analyzed. Much previous work and field tests suggests that the noises that most likely need to be rejected during on-site PD measurements are: white noise, sinusoidal and harmonics, repetitive pulses and random pulses [105]. Those noises have different patterns and can be classified into two groups: non-impulsive interferences and impulsive interferences. Features of these two groups of interferences are introduced in details in the following paragraphs.

A) Non-Impulsive Interferences

Non-impulsive interferences include white noise and sinusoidal noises.

White noises are the most common background noise. They are usually generated by amplifier, oscilloscope or any electrical equipment. White noises are equal-power signals that follow Gaussian distribution. A white noise signal is shown in Fig.2-16(a). Also, the frequency spectrum and TF spectrum of white noise in Fig.2-16(d) and (g) demonstrate the white noise has equal power density throughout the whole frequency range.

The sinusoidal interference is a regular signal whose magnitude decreases greatly when frequency is not equal to their oscillating frequency. It includes amplitude modulation (AM) radio, frequency modulation (FM) radio, and mobile communication signals in air [106]. Two typical kinds of sinusoidal interference: harmonics and modulated signals are usually encountered in practical measurement. The harmonic signals usually come from communication systems or electronic equipment. They contain the same frequency components all the time. Their energy decreases greatly in the frequency range that does not equal to their oscillating frequencies. Therefore, they appear to be sharp singularities

in frequency domain and strips that are parallel with time-axis in time-frequency domain. Fig.2-16(b) shows a sample of harmonic signal. The singularities in frequency domain appear at the oscillating frequencies. The time-axis-paralleling strips are also obvious in its TF spectrum in Fig.2-16(h). However, the modulating signal, for instance, signal from mobile telephone, looks like pulses. The pulses are actually a segment of sinusoidal signal. They also have sharp edges in frequency domain and time-frequency domain. This is illustrated by a modulated signal from mobile phone collected in laboratory in Fig.2-16(c). Similar to harmonics, large singularities are also seen in the frequency distribution of modulated sinusoidal in Fig.2-16(f). However, in its TF spectrum shown in Fig2-16(i), only some dots with sharp edges are found rather than time-axis-paralleling strips.

Commonly, the white noise and harmonics can be rejected by frequency-dependent thresholding. Both of them are very easy to remove when comparing with impulsive interferences. However, the modulated signal whose magnitude changes with time and frequency is not easy to be filtered by thresholding.

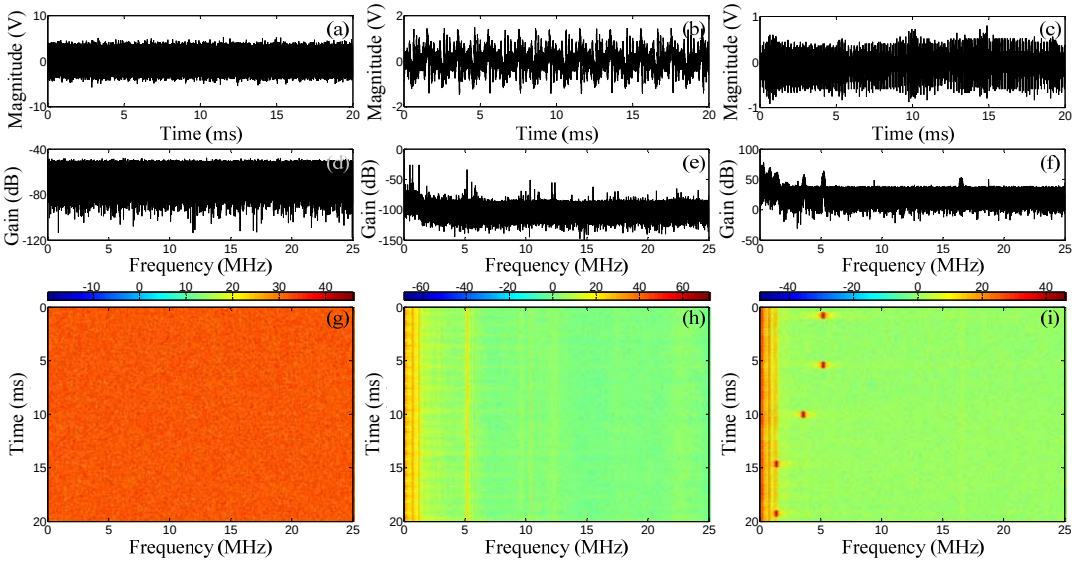

Fig. 2-16 The original data, frequency spectrums and time-frequency spectrums of non-impulsive noises, (a) white noise, (b) harmonics, (c) modulated sinusoidal, (d) to (f) frequency spectrums of signals in Fig.2-16(a) to (c) accordingly, (g) to (i) time-frequency spectrums of signals in Fig.2-16 (a) to (c), accordingly, the coefficients in TF spectrums are denoted in dB.

B) Impulsive Interferences

Impulsive interferences usually include repetitive pulses and random pulses. Impulsive disturbance is difficult to distinguish by using one technique alone because of its similarity to PD pulses in some aspects. The methods such as thresholding which is effective to remove white noise and harmonics are often ineffective to remove pulse-like

disturbances whose time-frequency spectrums appear to be strips that are parallel with frequency-axis.

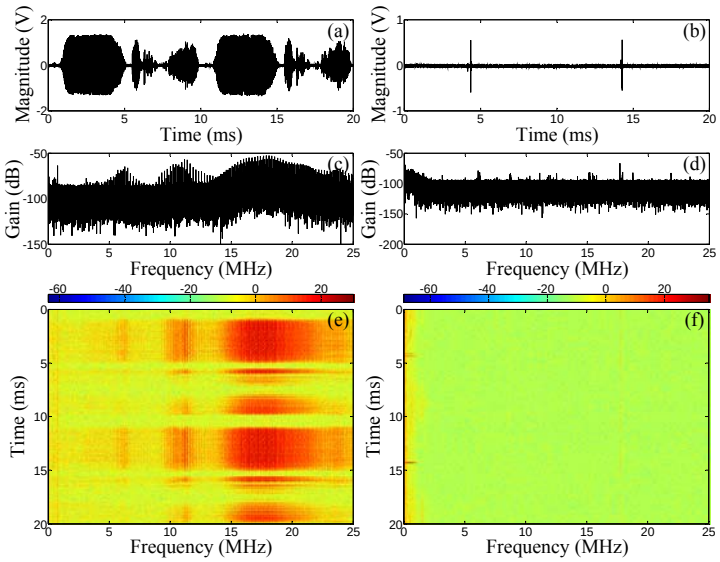

Fig. 2-17 The original data, frequency spectrums and time-frequency spectrums of impulsive noises, (a) repetitive pulses, (b) random pulses, (c) frequency spectrum of signals in Fig.2-16(a), (d) frequency spectrums of signal in Fig.2-16(b), (e) TF spectrums of signal in Fig.2-16(a), (f) TF spectrum of signal in Fig.2-16(b), the coefficients in TF spectrums are denoted in dB.

Repetitive pulses usually come from electronic apparatus such as AC/DC converter and rectifier. The repetitive pulses from the same source must have the same features (i.e. frequency distribution). Meanwhile, because of the regular switching behaviors of electronic equipment, the repetitive pulses tend to group at equally-spaced phase values. Highly-repetitive occurrence of these exactly the same and equally-spaced pulses can produce large-amplitude singularities in frequency domain, as shown in Fig.2-17(c). This characteristic suggests a possible solution of removing repetitive impulsive noise in frequency domain. Furthermore, the frequency-domain method is possible to separate pulses occurring concurrently.

Besides the repetitive pulses from electronic equipment, random pulse is another type of impulsive interference that is often encountered in field test. Random pulses come from breaker operations, lightning and so on. In general, there is no correlation between supply voltage wave and random pulses, and the random pulses from the same source are not identical at different time moments. Thus, unlike repetitive pulses, the large-amplitude singularities in frequency-domain which are caused by repetitive occurrence of same pulses are seldom found in the frequency domain of random pulses. This can be demonstrated by Fig.2-17(d). It is very hard to discriminate PDs from random pulses via frequency-domain analysis. However, the frequency distributions of pulses from the same

source must be highly similar and different from those of pulses from other sources. For example, the PD pulses from the same source that travel along the same path should have identical distortion during propagation. Their frequency distributions must be different from those of pulses that happen in the immediate vicinity of PD sensor which means less distortion. This difference in frequency distribution of each pulse suggests that the PDs and impulsive noises can be classified and recognized pulse-by-pulse according to their frequency distributions.

2.5.2 Methods for noise rejection

After the widely accepted application of PD measurement in insulation evaluations, extensive research has been carried out to improve the measurement accuracy and remove noises [40]. The noise rejection methods mainly fall into two categories: the hardware and software based methods. Usually, the software based methods are also called signal processing based methods.

A) Hardware based noise rejection

Hardware based noise rejection in PD measurement has been developed for over half a century. Most of those methods are effective and can be used to establish a reliable measurement system for PD detection. The noises are removed by different approaches which employ their different characteristics of interferences. The most popular techniques in hardware based methods include: sensors improvement, noise gating, and differential circuit.

1) Sensor improvement

As an important part of PD measurement system, PD sensor detects PDs and converts them into electric signals. Its performance was often investigated with the goal to increase the SNR of detected data. The sensors are different for different PD measurement systems. For instance, detector is used in coupling capacitor method while antenna is employed in UHF measurement. Normally, in coupling capacitor method, the frequency band was widen to improve the SNR [27]. In UHF measurement, the design such as size and shape, location and frequency range were usually discussed and studied for sensor improvement [101, 103]. The details of those sensor characteristics are introduced in Section 2.4.2 and shall not be repeated in this section.

Besides the UHF sensor and coupling capacitor mentioned above, different sensor types and methods to distinguish noise and PD have been proposed and partially put into service, for example, directional sensor, resistive temperature detector (RTD) [107], fiber-optic sensor [108], and so on. Most of them were reported to have effective performances, especially the directional sensor. Previously, directional sensors have been applied to stator bars [109], HV transformers [110] and power cables [111]. For HV equipment with earthed metal sheaths or enclosures, electrical noise sources are mainly external, for example, corona on the conductor connected to the HV terminal. On such a conductor, PD pulses travel from the equipment terminal outwards and electrical noise travels inwards, as shown in Fig.2-18 [40].

Directional sensors are often characterized by their coupling factor and their directivity. The coupling factor describes how much energy is coupled from the traveling PD waves into the output ports. Directivity quantifies the ability to distinguish between forward and backward propagating signals [111]. Much previous research has proved their effectiveness in rejecting external noises. However, this method could only distinguish the noisy pulses from external environment. The pulses that occur inside the HV equipment are still found in the measured data.

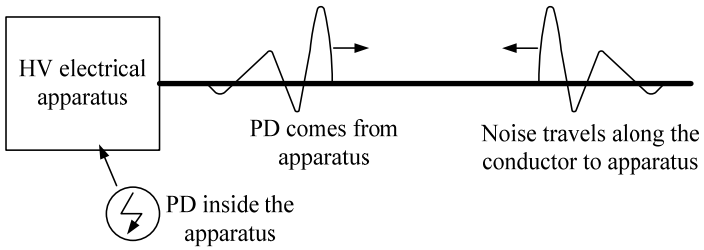

Fig. 2-18 The PDs from HV apparatus and external noise travel in opposite directions

2) Noise gating

Gating is an effective noise rejection method if the noise sources are known. It has been successfully used in PD measurements with pulse-type interference. A diagram of noise gating process is shown in Fig. 2-19.

Normally, the noise gating system includes fast analogue or digital switching circuits or gates. The gate is controlled by a triggering circuit that is activated whenever noisy pulse is detected [40]. If the noise pulse exceeds a pre-set trigger level, the gate opens, the input is disabled and the noise pulse is prevented from being logged into the output pattern [112]. The gate keeps open for a period which depends on the behaviors and oscillating natures of noisy pulses. In order to improve the measurement accuracy and sensitivity, the

PD signal after gating is amplified before being sent to the A/D converter. This noise removal method is optimum if the gate is only active when noise is present. Reported by some papers, with proper design, the gate can not only handle phase stable pulses but also external corona or brush noise from DC machines [113].

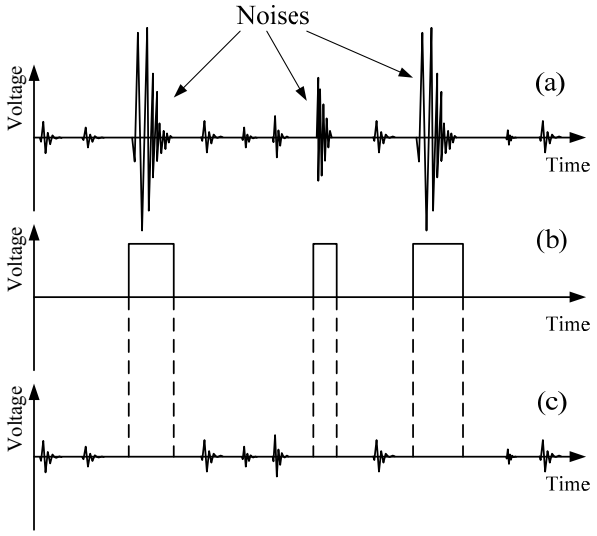

Fig. 2-19 Noise gating procedure, (a) PD signal with noisy pulses, (b) gating windows triggered by noise, (c) PD signal after noise gating (after James *et al* [40])

However, the application of noise gating method is limited to distinguishing noise pulses from known sources. The trigger level is difficult to set when the noisy pulses are unknown and unpredictable. Further, as this method is based on the pure time-domain features, it is likely to lose PD pulses if the PD and noise pulse occur at the same time.

3) Differential circuit

The noise rejection technique by the balanced circuit or the use of two sensors in differential mode has been well developed and widely used in PD measurement [114]. A typical model of this circuit, as portrayed in Fig.2-20 [5], was first introduced by Kurtz *et al* [115], and subsequently by Stone *et al* [73, 116]. In this balanced circuit, the outputs of two parallel circuits are compared. If certain criteria are met, the external noises at the two outputs are identified and can be canceled by the differential amplifier. Otherwise, the outputs are considered as PDs.

As shown in Fig.2-20, the basic PD measuring circuit in Fig.2-8 is duplicated and used in each branch of the balanced circuit. Here the capacitor coupler, terminated in a resistor, is installed differentially with one coupler per circuit and two parallel circuits per phase. The coupler pairs, C_1 and C_2, with same bus bar length and coaxial lines are matched in

their equivalent electrical length at the input of the differential amplifier. That means the electrical length of each circuit should equal such that $x + y = s + r$. Therefore, the incident interference pulses arriving from the apparatus with equal traveling times are thus canceled in the differential mode [5].

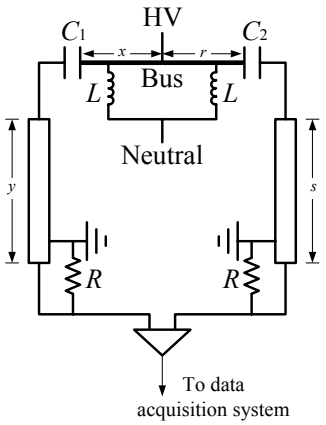

Fig. 2-20 Balanced permanent coupler connections

Although differential circuit is a well-known technique, this measurement system has some disadvantages [114]: (i) balance over a wide frequency range is very difficult, if not impossible; (ii) identical apparatuses or dielectric characteristics requirements are difficult to satisfy; and (iii) slight time shift which causes interference at the input of amplifier is difficult to avoid in two branches, especially for medium length power cables. The most important one is its inapplicability to on-line use.

B) Signal processing based noise rejection

As discussed above, it is quite difficult to reject noise by using existing hardware methods directly in on-line TEV measurement. However, if the TEV signals are collected with a wide frequency spectrum, the software based technique could be a powerful tool to remove noises. Signal processing has been used in insulation condition monitoring for many years. With the help of computational techniques, it costs much less than hardware method if same performance is required. Previously, the time-domain features such as magnitude, durations, and waveforms were used to distinguish PDs from noises [117]. Then, the frequency-domain features, like frequency distribution and the Fourier coefficients were employed [118]. With much development, the time-frequency analysis which could provide more information becomes increasingly popular in PD analysis. In following content, the noise rejection methods via signal processing are reviewed in the aspects of domains.

1) Time domain methods

At the beginning of PD measurement, the time-domain analysis was used to remove the noises in detected data. The features in time domain are closely related with the waveform of PD pulses. According to those features, the de-noising methods are developed into three main branches: statistical evaluation, probability analysis and wave shape analysis.

a) Statistical evaluation

In statistical evaluation of PD activities, several quantities were chosen to describe the aging and deterioration level of insulations: (i) pulse magnitude (denoted by q): this parameter is correlated to the voltage stress across the discharging gap; (ii) pulse phase angle (denoted by ϕ): this parameter is correlated to deterioration energy, and indicates the nature of physical changes in the discharge area; (iii) pulse number (denoted by n): this parameter indicates the frequency of PD reoccurrence. Pulse magnitude, phase angle and number are very suitable quantities for digitization and automatic investigation, which are the basic requirements for the statistical PD evaluation [119].

Usually, different patterns of these quantities are used in PD discriminations: $q-\phi$ pattern that is called pulse-height distribution evaluates the distribution of PD magnitude, either maximum or average values, and the phase angle of pulses [120]. $q-n$ or $\phi-n$ pattern is known as pulse-count distribution [121]. The $\phi-q-n$ distribution is the most classical distribution and adopted by some researchers [122] and international standards [6].

With the help of mesh technique, visual recognition of PDs can be realized by displaying the statistical distributions of different patterns [70, 123]. However, discriminating PDs from background noises with statistical evaluation requires experiences. The pulse-height-count distributions provide a visual model which is difficult to make a judgment in quantitative way. Therefore, the application of statistical evaluation is often accompanied by artificial intelligence such as neural network.

b) Probability analysis

The probability analysis was first proposed for better analysis of statistical evaluation of PD measurement. As in $\phi-q-n$ distribution patterns, the PD pulse occur in positive and negative half cycles. Therefore, the symmetry of distributions of pulse height and numbers in two half cycles are evaluated by several probability operators: (i) skewness

which describes the asymmetry of distribution with respect to a normal distribution; (ii) kurtosis which represents the sharpness of the distribution with respect to the normal distribution; (iii) cross-correlation factor which describes the differences in shapes of two half cycles; (iv) asymmetry which describes the difference in the mean discharge level of two half cycles [124]. Thus, with the probability operators mentioned above, the PD pattern can be described by more values such that more information could be provided for artificial intelligence and increase the possibility of correct judgments.

Besides the application in statistical evaluation, the application of those probability operators was extended into wider areas. Recently, the probability operators such as skewness and kurtosis were employed in PD shape analysis and good performance was reported [37]. The skewness and kurtosis of each pulse are calculated and used as the input of classifier.

c) Wave shape analysis

An important step forward in time-domain based pulse separation has been obtained by using digital instrumentation which allows the acquisition of the whole shape of pulse signals at a sampling rate high enough to avoid the frequency aliasing. This PD pulse extraction problem has been approached under the assumption that signals from the same source have similar shapes and those from different sources are characterized by different waveforms [125].

In some wave shape analysis, the pulses are characterized by some basic parameters such as rise time, fall time, duration and amplitude [126]. In some other research work, correlation factor between any two pulses is calculated and used to describe the differences between pulses. The features of pulse wave shape are clustered and thus the PD pulses and noise pulses are separated.

2) Frequency domain methods

The frequency spectrum in pure frequency domain reveals the frequency features of PD. A major advantage of frequency-domain methods is to distinguish concurring PDs and interferences. Normally, the frequency domain de-noising is realized in two main ways: Fourier analysis and filtering.

a) Fourier analysis

The frequency spectrums of PD signals can be obtained by transferring the time-domain signal into frequency domain. This procedure is usually done by fast Fourier transform (FFT). Then the interference threshold was set and a series of frequency bands of noise were determined by ways of threshold searching method [127]. Peak values of those frequency bands were set to zero [128], or smooth compensations were made to replace those parts with interference frequency bands [127].

b) Filtering

Filter is an equipment to pass some signal and block others. For PD analysis, the filters can be divided into two kinds: band rejection filter for filtering known frequency band and adaptive filter for filtering unknown frequency bands [129].

According to the bandwidth and cutoff frequencies, band rejection filters are known as low, high and band pass filters. As mentioned in previous sections, high pass filters can remove low frequency noises in UHF measurement. Further, a notch filter is often used to remove the sinusoidal noise present in the signal [130].

On the other hand, the adaptive filter or predictor filter calculates the next sample from a certain amount of previous sampling values. It often has a closed loop design, in which filter coefficients are varied to reduce the noises. The filter accomplishes the objective by automatically updating the filter coefficients with the availability of each new sample of the data. After processing a number of samples in this way, the algorithm evolves the optimum filter coefficients after which the adaptation can be stopped [130]. Finally, the noises in following PD signal can be canceled by this adaptive filter with optimum coefficients.

3) Time-frequency domain methods

It is quite clear that time-domain methods cannot distinguish pulses occurring at the same time, and frequency-domain methods cannot reject interferences that fall in the same frequency band with PDs. The properties of time-frequency analysis were explored for a more effective approach.

In early research in signal processing of PDs, some researchers tried to combine the time and frequency domain features to provide more information for de-noising. The basic

idea is similar with the statistical evaluation in time domain [131]. Rather than the magnitude, phase angle, pulse number, and probability operators, some frequency parameters such as frequency spectrum are added. Subsequently, to obtain a better characterization of pulses and shorten the processing time, different time and frequency operators were proposed. The most famous algorithm was proposed by Contin *et al* [132]. The separation of PD pulses from different sources and of PD pulses from noise is based on a clustering technique that relies on a time-frequency characterization applied to each recorded pulse. In the time-frequency characterization, two quantities (the equivalent time length and the equivalent bandwidth) were extracted from each detected pulse. Then the pulses with similar quantities were clustered into several groups.

However, in this approach neither the time-domain and frequency-domain features alone nor their combination could reveal the energy variation with time and frequency. Thus, the time-frequency analysis was proposed thereafter. With many years' development, much fruitful work has been obtained in the area of PD noise reduction with TF analysis.

Wavelet transform is one of the most popular time-frequency transform used in PD analysis. Wavelet based thresholding often produces a good estimation of PD pulses. In this method, the signal is decomposed into several spaces: one approximate space and many detail spaces. All the coefficients in detail spaces are scanned with threshold. The coefficients that are smaller than threshold are regarded as noise related and then removed. To generate better estimation, the most important factor is the selection of thresholds. The universal threshold [39], SURE threshold [133], minimax threshold [134] and empirical threshold [105] have been proved effective in PD recovery. However, all the wavelet-based thresholding methods aim to remove white noises and harmonics. If the reduction of pulse-like interferences is required, other tools or information must be added, for example, artificial neural networks [37] or noise references [135]. Furthermore, the resolution of wavelet transform is unequal in time-frequency domain. It has lower frequency-resolution and higher time-resolution in high frequency range and higher frequency-resolution and lower time-resolution in low frequency range. This may not be appropriate for all kinds of PDs, for instances, the PDs that occur in SF_6 have wide frequency range. However, the frequency range of PDs detected on cable surface is much narrower since the high frequency energy has degraded greatly through propagation.

Besides wavelet transform, the de-noising ability of other time-frequency transforms were explored by some researchers. Wong [136] pointed out the possibility of employing

Wigner-Ville distribution in PD analysis. However, the Wigner-Ville distribution could not be easily applied in PD de-noising because of the interferences created by the transform's quadratic properties. Although the interference can be removed by averaging the Wigner-Ville distribution with appropriate kernels, the time-frequency resolution will be reduced.

X. Wang *et al* [137] introduced Hilbert-Huang transform (HHT) in PD analysis. They compared the frequency spectrums of PD by using both Fourier transform and HHT. The HHT displayed more energy of PD than Fourier transform. However, in another comparison [138], the HHT is observed to amplify some noises. This reveals the drawback of HHT that its performance cannot be evaluated easily.

Short-time Fourier transform (STFT) is a basic time-frequency transform. It has equal resolution in time-frequency domain when the sliding window is chosen. The spectrum generated by STFT displays the power distribution of signal. Furthermore, no interferences will be induced by transform. Some research has proved the potential of applying STFT in PD analysis. Caironi *et al* [139] employed STFT in PD noise reduction. The noises were removed manually according to the TF spectrum of noisy data. Then neural network was induced in STFT-based PD de-noising by Thayoob *et al* [140] and good performance was reported. However, these methods only discussed the potential of employing STFT in PD extraction. A systematic analysis has not been proposed.

2.6 Applications of PD measurement

As a reliable and convenient tool for evaluating insulation conditions, TEV based PD measurement can be widely applied in monitoring electrical apparatus such as cables, rotating machines, transformers and switchgears.

The insulating materials of cables are usually oil, paper and polyethylene. The PD detections of oil-paper insulated cables were mainly applied to their terminations and joints because the PD pulses far away from measuring equipments are difficult to measure because of the large attenuations [146]. Due to the introduction of polyethylene extruded cables in power system, the PD measurement of cable specimens developed rapidly [5]. Currently, PD measurements were commonly used to evaluate new cable specimens. When PD pulses propagate along the cables, the high frequencies present in original PDs attenuate greatly and rapidly such that the pulse durations increase and the

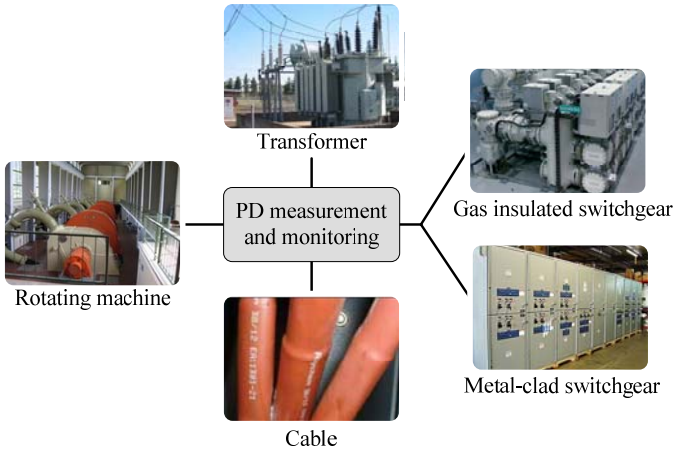

Fig. 2-21 The applications of PD measurement on different electrical apparatus (the figures of electrical apparatus are from [141-145])

pulse heights decrease. Thus, the TEV detection for cables usually does not require ultra-wide range of frequency response [147].

Nowadays, many transformers still use oil as insulation which can be evaluated by DGA as mentioned in Section 2.4.1. However, for some transformers whose insulations are mainly epoxy or polyester, the PD measurement is far more complex than for cables. Since the transformer is an inductive device, the PD occurring inside the windings have to travel a long way before detection. The high frequency components are filtered and distorted greatly. Further, the occurrence of resonance introduces errors in measurement [5]. Therefore, noise separating techniques should be developed to improve the detection sensitivities [148, 149]. Those methods have achieved fruitful results in both experimental and practical applications.

The insulation condition of switchgears, which are important equipment in transmission level, is also monitored by PD measurements. On-line tests are highly recommended and preferred by switchgear owners because a suitable interruption is difficult to arrange and the phase-to-phase insulation cannot be tested [146]. The most difficult part of TEV based PD measurement on switchgear is how to reject repetitive noises from switching operations.

Finally, PD measurement can also be applied in rotating machines. However, different from other HV equipment where PD phenomenon is not allowed, discharges in the insulation systems of stator cannot be eliminated entirely and past experiences show that rotating machines could operate a long time after the presence of PD [5]. Therefore, the PD measurements on rotating machines emphasize more on recognizing PD pulse

distributions and patterns instead of ascertaining the occurrence and density. Furthermore, to improve the rate of correct judgments, PD noise reduction is very important.

2.7 Conclusion

The application of PD technology to diagnose the condition of the electrical insulation in HV equipment has grown dramatically over the past decades. The technology has been widely accepted for many kinds of electrical apparatus. With appropriate detection of PD activities, some insulation problems are found before catastrophic failure and maintenance costs are reduced. For proper design and application of PD detection system, it is very important to explore the mechanism and properties of different forms of PD. Based on a comprehensive understanding of PD features, the PD measurements and de-noising methods are reviewed. It can be concluded that there are many ways to achieve an effective PD detection.

The popular measuring methods: coupling capacitor method and UHF method have drawbacks in on-line PD measurement. To decrease the interruptions to operations, inspections without electrical shutdown are preferred. Further, the installation of instrument should be convenient and the detection needs to be effective in a wide range of equipment. TEV-based measurement which only requires the placement of non-intrusive sensor on the outside surface of metal-clad is a good way for on-line insulation diagnosis and has the potential to be widely employed in practical applications. However, as a new technique, its measurement theory and design of detection system have not been well developed.

Due to the external location of TEV sensors, noise is always a barrier to reliable detection. To improve the measuring accuracy, noise rejection method should be effective and with low cost. When compared to hardware based methods, software or signal processing based methods could achieve same de-noising performance with less cost. Among all signal processing algorithms, the time-frequency analysis provides more information of signals and was proved to be more effective than time or frequency analysis alone. Many software based de-noising methods have been proposed. However, not all of them are suitable for TEV signals. For those methods which can be widely used for all existing kinds of PD signals, their parameters should be adjusted when applied in TEV analysis.

Further, the impulsive noise is the most difficult part of PD measurement. Although many methods have been developed, the ones for TEV signals were seldom mentioned. Thus, there is a need for impulsive noise rejection especially for TEV signals. Among all signal processing algorithms, the time-frequency analysis provides more information of signals and was proved to be more effective than time or frequency analysis alone. Hence, one of the potential research areas is to develop de-noising algorithms with time-frequency analysis, for instance, wavelet transform and STFT.

Therefore, all the above directions of insulation diagnosis using TEV-based measurement and its de-noising algorithms will be described with original research results in the following chapters.

CHAPTER 3
TEV THEORY AND ITS MEASUREMENT SYSTEM

3.1 Introduction

Partial discharge (PD) instrumentation plays a paramount role in determining system reliability. Its measurement and interpretation on metal-clad apparatus represent a far more complex and intricate task than that on other equipment such as cables [5]. Electric shutdown of HV equipment is often required in traditional methods such as coupling capacitor method to improve the detection sensitivity. For the UHF method, the PD sensors are often mounted inside the metallic enclosure to achieve high signal to noise ratios (SNR). However, for operational switchgears and transformers without such UHF sensors, arranging a shutdown specifically to fit internal couplers rarely can be justified [89]. So the transient earth voltage (TEV) method with non-intrusive sensors which are easy to install and have no interruptions of operation are becoming more and more popular for devices which are not suitable to shutdown.

In this chapter, the TEV theory, measurement system, and simulation technique are introduced as the basis of further analysis in following chapters. This chapter starts with the review of the fundamentals and characteristics of TEV signals. Based on the knowledge of TEV, a measurement system with non-intrusive sensor and three-order Butterworth filter is proposed and introduced in details. To verify the effectiveness of such measurement system, experimental tests are carried out. Finally, the measured TEV signals are simulated for further theoretical investigations in Chapter 4.

3.2 TEV measurement

Although the non-intrusive PD detection using transient earth voltage (TEV) method has been developed for more than 30 years, it is not a well-established technique when compared with other PD measurement methods. Many aspects of its theory and application are still under development. However, much previous research work has been

done to explore its principles and characteristics, and therefore is introduced in this section.

3.2.1 Principles of TEV method

The TEV phenomenon was first proposed and named in 1974 by Dr John Reeves in EA Technology. He pointed out a large part of the electromagnetic (EM) waves produced by PDs conducts away by the surrounding metal work and a small proportion radiates onto the inner surface of the casing [150]. However, there are some discontinuities in the metal-claddings of GIS or other equipments, for example gasket, insulator, insulated pillar, and flange junction, which would let those EM waves leak out from the shielding enclosure and then pass across the external surface of the metal cladding to the earth [151]. Usually, the detected TEV has a rise time less than one microsecond and typically lasts for a few microseconds. Its amplitude varies widely from milli-volts to volts [152].

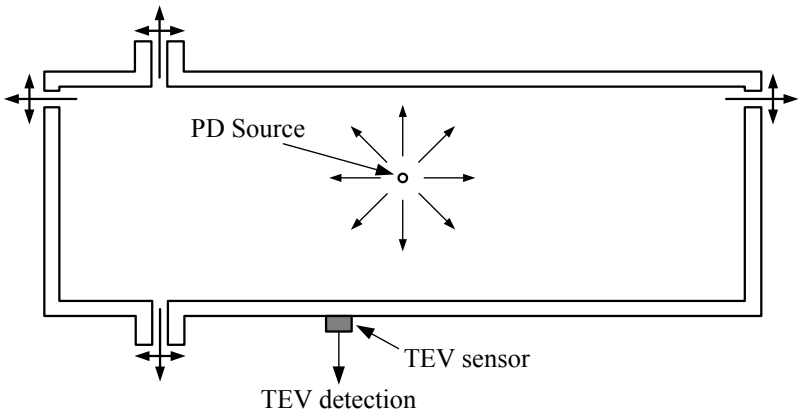

Fig. 3-1 Mechanism of transient earth voltage (TEV)

Fig.3-1 gives an example to illustrate the mechanism of TEV. The EM waves escape from the dielectric discontinuities in metal tank and then propagate on the external surface of cladding to the ground. Such propagating EM waves can be obtained by capacitive sensor and recorded as TEV signals [151]. Currently, the research on TEV based PD signals attracts increasing number of researchers. However, the characteristics of TEV-based PDs and pulse recognition methods are still rarely discussed.

3.2.2 Characteristics of TEV signal

To study the characteristics of TEV signal, modeling is a simple and basic method and is employed here. A metal enclosure of gas or oil insulated apparatus is modeled as in Fig.3-

2. It is a rectangular box with the origin of coordinate at the center of the external surface of its front wall. On the front wall, there is a small opening which is filled with dielectric materials. An area on the front wall which is bounded with dashed line is magnified and shown in Fig.3-2(b). The location of PD source is assumed to be in the center of the box with a coordinate $(0,0,a)$. The front metal wall has a thickness of b. Accordingly, the center of dielectric opening is (x,y,b) on the inner side, and $(x,y,0)$ on the external side. The distance between the point (x,y,z) and z-axis is ρ which equals $\sqrt{x^2+y^2}$. The distance between the point (x,y,z) and the PD charge is r, and $r=\sqrt{x^2+y^2+(z-a)^2}$. The electric potential V_d on the outside surface of the front wall can be calculated if the current of PD and the dimension of box are given.

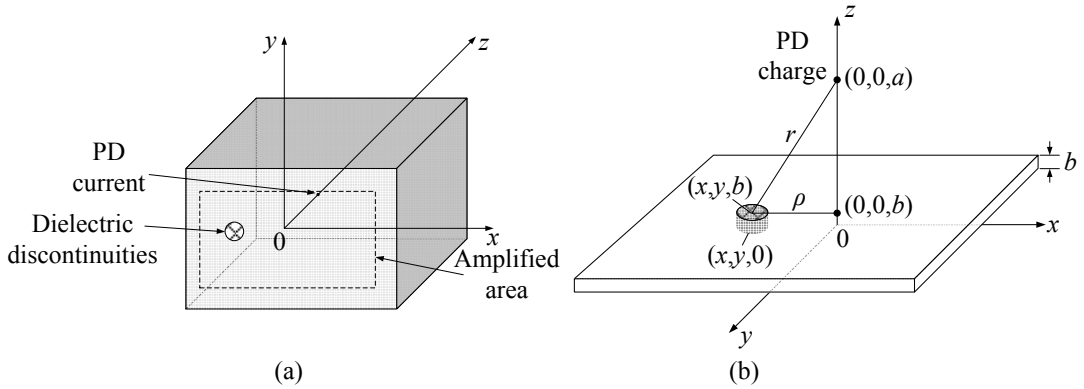

Fig. 3-2 Illustration of PD model inside a metallic enclosure, (a) PD occurs inside a metal box, (b) magnified area on the frontal surface

However, exact modeling of the original TEV signal V_d is a complex and tough procedure. Many factors will affect its value, for example, the amount of PD charges, the induced charges on all the other walls, the thickness of dielectric opening, the distance between detecting point and dielectric opening, and so on. Thus, this section gave a brief analysis of TEV signal under some assumptions to help modeling the potential V_d at the detecting point.

Assumption 1: the induced charges on the other walls are so small that they can be ignored in calculation. When a changing magnetic field is present near a conductor, an electric current is produced [153]. In time-varying electromagnetic field, a displacement current is added to keep the Ampere's Law valid. At higher frequencies, the two terms are comparable and $\mathbf{J}_d$ cannot be ignored. At low frequencies, $\mathbf{J}_d$ is usually neglected compared with $\mathbf{J}$. In TEV-based measurement, the frequency response range is much lower than that of UHF sensors. Further, the PD energy decreases with the increase of frequencies in the popular damped exponential PD model. Since only an approximate

model of TEV signal is calculated, the influences from displacement current are ignored here. Therefore, at the point (x',y',z') on the metal walls except front wall, its magnetic field $\mathbf{H}$ caused by PD current I_{PD} is

$$\mathbf{H} = \int_L \frac{I_{PD} d\boldsymbol{l} \times \mathbf{a}_r}{4\pi \mathbf{r}'^2} \tag{3-1}$$

where $r' = \sqrt{x'^2 + y'^2 + (z'-a)^2}$, $\mathbf{a}_r$ is the unit vector along $\mathbf{r}'$. Hence, the induced current density K on the metal wall which is a good conductor should be

$$\mathbf{K} = \mathbf{a}_n \times \mathbf{H}, \tag{3-2}$$

where $\mathbf{a}_n$ is the unit vector along normal direction of metal walls. Such induced current produce a magnetic field $\mathbf{H}_{ind}$ at the center of inner side of dielectric opening (x,y,b). It equals

$$\mathbf{H}_{ind} = \int_S \frac{\mathbf{K} d\mathbf{S} \times \mathbf{a}_{r-r'}}{4\pi(\mathbf{r}-\mathbf{r}')^2} \tag{3-3}$$

However, the metal enclosures in practical use often have large sizes which suggest large $\mathbf{r}'$ and $\mathbf{r}-\mathbf{r}'$. If the energy loss during radiation is also considered, the influence from induced charges is quite small and could be neglected. Therefore, for an easier calculation, only the effect from PD charge is considered. Consequently, the potential V_{in} induced on the inner side of the dielectric opening should be

$$V_{in} = \frac{Q_{PD}}{4\pi r^2} \tag{3-4}$$

where $r^2 = \rho^2 + (a-b)^2 = x^2 + y^2 + (a-b)^2$.

Assumption 2: the thickness of dielectric opening is extremely small such that the differences between the inner and external sides of dielectrics can be ignored.

According to the boundary condition of dynamic electromagnetic theory, the electric field $\mathbf{E}_{in}$ on the insulation-side surface of dielectric opening and $\mathbf{E}_{die}$ on the other side are different:

$$(\mathbf{E}_{in} - \mathbf{E}_{die}) \times \mathbf{a}_n = 0 \text{ and } (\varepsilon_{in}\mathbf{E}_{in} - \varepsilon_{die}\mathbf{E}_{die}) \cdot \mathbf{a}_n = 0, \tag{3-5}$$

where ε_{in} and ε_{die} are the permittivities of insulation and dielectrics, respectively. It is clear that the electric field $\mathbf{E}_{in}$ at (x,y,b) is not equal to $\mathbf{E}_{die}$ at $(x,y,0)$. Accordingly, the potential V_{out} at the external surface of dielectric opening is not equal to V_{in} if the thickness b of metal wall is non-zero.

However, in practical test, the thickness of dielectric plate is only several millimeters, which is quite small when compared with the distance r which is in the order of meters. Therefore, if we assume the thickness is extremely small $b \to 0$ and the vertical distance from PD source to dielectric opening is far larger than the thickness of dielectrics $a \gg b$, the differences between V_{in} and V_{out} will be negligible such that $V_{in} \approx V_{out}$. Thus,

$$V_{out} \approx \frac{Q_{PD}}{4\pi(\rho^2 + a^2)} \tag{3-6}$$

Furthermore, according to the Maxwell's equations, the following equation will hold when the PD charge is contained in the closed surface of enclosure.

$$Q = \int_S \mathbf{D} \cdot d\mathbf{S} = \int_S \varepsilon \mathbf{E} \cdot d\mathbf{S} = \int_S \varepsilon (\frac{\mathbf{J}}{\sigma}) \cdot d\mathbf{S} = \frac{\varepsilon}{\sigma} \int_S \mathbf{J} \cdot d\mathbf{S} = \frac{\varepsilon}{\sigma} I \tag{3-7}$$

Therefore, the potential can be written as

$$V_{out} \approx \frac{\varepsilon I_{PD}}{4\pi\sigma(\rho^2 + a^2)} = \frac{\varepsilon I_{PD}}{4\pi\sigma r^2} \tag{3-8}$$

As mentioned in Chapter 2, the PD current is generally modeled by an exponential signal $I_{PD} = I_0(e^{-t/\beta} - e^{-t/\alpha})$ where α and β are the rise and decrease time respectively, the potential of TEV in (3-8) can rewritten as

$$V_{out} \approx \frac{\varepsilon I_{PD}}{4\pi\sigma r^2} = \frac{\varepsilon I_0(e^{-t/\beta} - e^{-t/\alpha})}{4\pi\sigma r^2} \tag{3-9}$$

Assumption 3: the distance between detecting point and the external surface of dielectric opening is quite close such that the attenuation during propagation is very small.

When EM waves propagate along the metal wall, their field and associated currents are confined to a very thin layer of the conductor surface [154]. This is the so called skin effect. The potential V_d at detecting point $(x_d, y_d, 0)$ is usually lower than the potential V_{out}

at the external surface of dielectric opening $(x, y, 0)$ because of the loss on surface resistance which is defined as

$$R_s = \frac{l}{w} R_{unit} = \frac{l}{w} \sqrt{\frac{\pi f \mu}{\sigma}} \tag{3-10}$$

In (3-10), l and w are the length and width of propagating area, f is the frequency of propagating current, and R_{unit} is the resistance of a unit width and length of conductor. Accordingly, the TEV signal V_d at the detecting point can be modeled as

$$V_d = V_{out} - R_s I_{TEV} \tag{3-11}$$

Hence, it is obvious the TEV potential V_d depends on the value of R_s if V_{out} and I_{TEV} are given. According to the definition of surface resistance, the unit surface resistances of common kinds of metal materials are listed in Table 3.1.

Table 3.1 Unit surface resistances of some typical materials of cladding

Material	Relative permittivity μ_r	Conductivity σ	Unit surface resistance R_{unit}		
			f=100kHz	f=10MHz	f=100MHz
Aluminum	1	3.8×10^{7}	1.02×10^{-4}	1.02×10^{-3}	3.22×10^{-3}
Copper	1	5.8×10^{7}	8.25×10^{-5}	8.25×10^{-4}	2.61×10^{-3}
Silver	1	6.1×10^{7}	8.04×10^{-5}	8.04×10^{-4}	2.54×10^{-3}
Stainless steel	1	1.1×10^{7}	1.89×10^{-4}	1.89×10^{-3}	5.99×10^{-3}

Note: the permittivity of vacuum μ_0 equals $4\pi\times10^{-7}$.

Table 3.1 shows the unit surface resistances of those materials are very small when frequencies of TEV current are not too high. In practical application, the upper limits of TEV sensors are lower than 100MHz. Therefore, if the length l of propagation is very small and the width w is much larger than l ($l \ll w$), the differences due to energy loss can be ignored. Thus, the TEV potential at detecting point which is quite close to the dielectric opening can be expressed as

$$V_d = V_{out} - RI_{TEV} \approx V_{out} = \frac{\varepsilon I_0 (e^{-t/\beta} - e^{-t/\alpha})}{4\pi\sigma r^2} \tag{3-12}$$

With the approximate TEV potential in (3-12), it is clear that the magnitude of TEV is proportional to the amplitude of PD current, and inversely proportional to the square of

distance between measuring point and PD source. This was demonstrated by Li *et al* in [155]. They modeled a metal-clad GIS tank with 1:1 ratio in electromagnetic simulation software, XFDTD. The dimension of that box is 8.5m×5m×16m with a PD source located at the center of it. The detecting point is at the external surface of front wall. The simulation results in Fig.3-3 shows the magnitudes of TEV signal increase with the amplitude of PD current and decrease with the distances from PD source. They verified that the approximate TEV potential in (3-12) can represent the major characteristics of actual TEV signals.

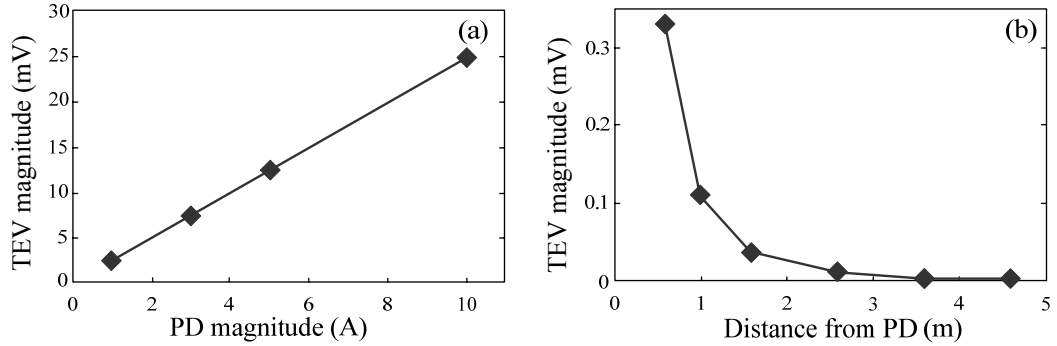

Fig. 3-3 Characteristics of TEV signals, (a) variations with PD magnitudes, (b) variations with the distances from PD sources (after Li *et al* [155])

3.3 TEV measurement system

According to the discussions in Section 3.2, the TEV measurement has some features such as good sensitivity, non-intrusive detection, and attenuations of high-frequency components during propagation. Consequently, the TEV measurement system should have suitable frequency response range and be sensitive enough to catch the useful signal. The measurement system is composed of three parts: the non-intrusive sensor that collects the TEV signals on the external surface of metallic tank, high-pass filter which removes the low-frequency disturbances and ensures the impulsive waveform of detected signal, and the data storage and display equipment which is often a computer or an oscilloscope. The first two parts: non-intrusive sensor and high-pass filter need special design and are introduced in this section.

3.3.1 Non-intrusive sensor

The non-intrusive sensor for TEV detection should be as sensitive as possible due to its external location which often results in small magnitude of PD pulses. The monopole antenna which is capable of equally receiving signals from all directions in one plane is

adopted in our design. In order to shield the interferences from surroundings, this monopole sensor is integrated with a metal cover. A basic-designed sensor was made and used in experimental environment. However, because of the defects in cover design, it collects interferences as well in field test where the SNR is much lower than in laboratory. The sensor was subsequently improved. Both the basic and improved sensors are introduced.

A) Basic-designed sensor

The non-intrusive sensor with basic design was originally made for the experimental use in laboratory tests. It is a coaxial sensor which includes four parts: part 1 is the female BNC interface, part 2 is a metal cover that shields environment interferences, part 3 is the outer conductor and part 4 is inner conductor that is made of copper. The BNC interface (part 1) is integrated with the cover (part 2). The structure and size of this coaxial sensor is shown in Fig.3-4(a). It is easy to find the protruding inner conductor (part 4) extends beyond the bottom of outer conductor (part 3). This design ensures close contact between the inner conductor and external surface of metallic enclosure. However, due to the small difference between the lengths of inner and outer conductor, the outer conductor cannot shield all interferences from surroundings which are more likely to be collected by inner conductor.

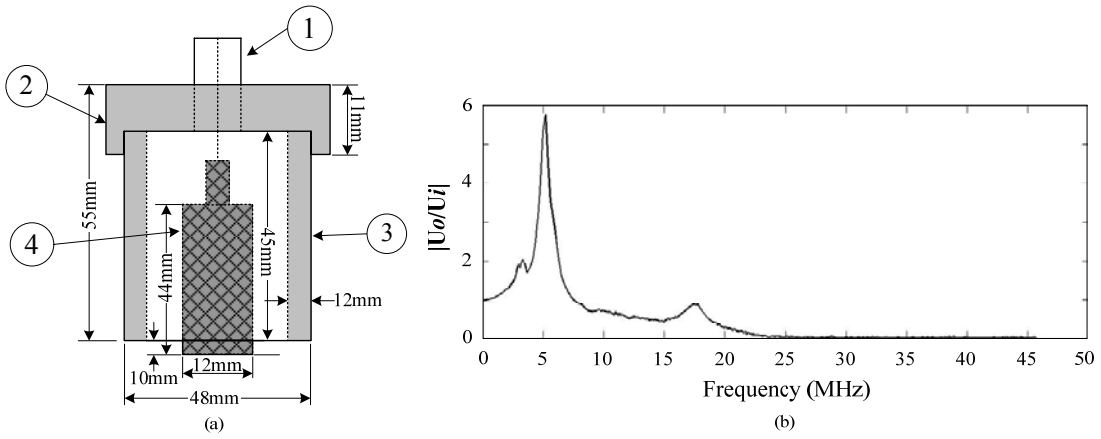

Fig. 3-4 Design of basic coaxial PD sensor developed for non-intrusive PD measurement, (a) construction of the sensor, (b) frequency response.

Normally, the pulses generated at PD source have wide frequency range, up to tens or hundreds of megahertz, or even several gigahertzes. However, in TEV measurement, the power in high frequency range attenuates greatly in propagation. The sensors for non-intrusive measurement can have a bandwidth of tens of megahertz which is easy to realize by hardware design and small lower-cut-off frequency to pick up less attenuated low

frequency components. The coaxial sensor design in Fig.3-4(a) meets such requirements. Its amplitude response is portrayed in Fig.3-4(b). Here, $|U_o/U_i|$ equals the absolute value of output voltage divided by input voltage at different frequencies. The -3dB points of the amplitude response are around 80Hz (low frequency) and 9MHz (high frequency). The -3dB point is defined as the frequency where $20\log_{10}|U_o/U_i|$ equals -3. Although such frequency response range is narrower than some industrial TEV sensors, for example, the -3dB response of HVPD TEV sensor is from 1MHz to 50MHz, a large number of tests were carried out to prove its effectiveness [156].

B) Improvement on sensor design

As aforementioned, the non-intrusive sensor with basic design has an inner conductor which is longer than outer conductor to ensure a close contact between inner conductor and cladding surface. Such design is effective in laboratory tests where the signal-to-noise ratio (SNR) is high, but it is also susceptible to noises when employed in field tests where external noises are unavoidable. In order to shield the interferences as much as possible, the design of TEV sensor is improved.

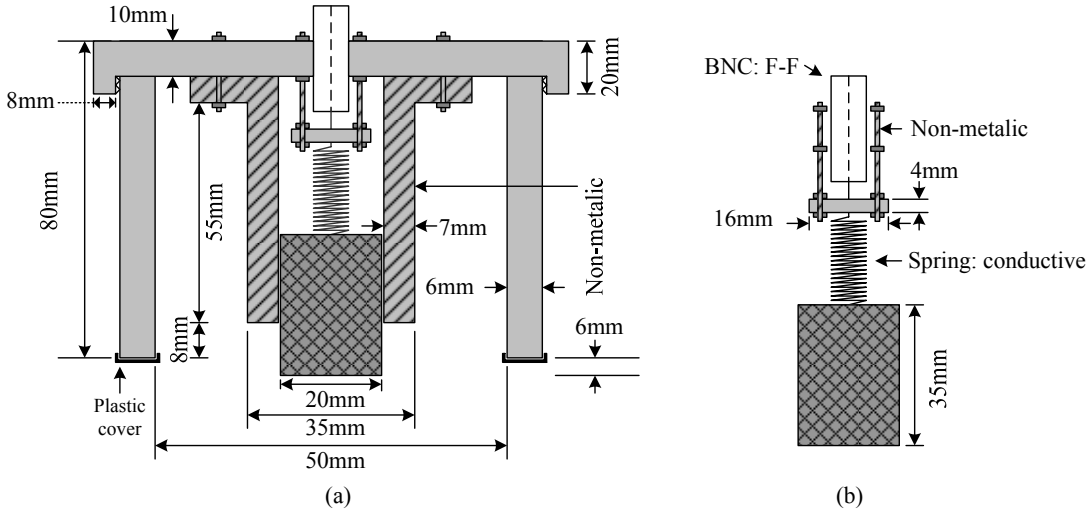

Fig. 3-5 Design of improved coaxial sensor for non-intrusive measurement

The structure and size of improved sensor is shown in Fig.3-5. The improved sensor still comprises four parts: BNC interface, cover, outer conductor, and inner conductor. Similar with basic design, the cover and outer conductor are connected together to shield the interfaces. The improvements are done in the aspects of inner conductor and its connection between BNC interface.

Rather than direct connection with BNC interface in basic design, the inner conductor and BNC interface are connected via a conductive spring, as portrayed in Fig.3-5(b). When the sensor is not in use, the inner conductor extends beyond the bottom of outer conductor. However, when the sensor is mounted on the surface of metal enclosure, the spring in compressed, and the outer conductor and inner conductor are aligned on the same level. This design ensures that both the inner and outer conductors are fully in contact with enclosure surface. Therefore, it could collect the TEV signal with high sensitivity and shield the external interferences as much as possible.

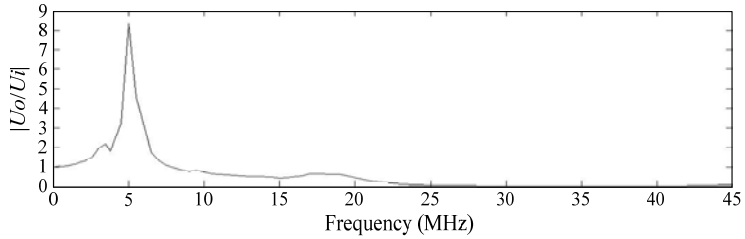

Fig. 3-6 Amplitude response of improved coaxial sensor

Although the size and connection between inner conductor and BNC interface are different, the frequency responses of the two sensors are similar. The amplitude response of improved coaxial sensor is shown in Fig.3-6. The -3dB response is from around 10Hz to 10.5MHz.

3.3.2 Software based high-pass filter

According to the descriptions in Section 3.3.1, the lower cut-off frequencies of non-intrusive sensors are very small: 80Hz for basic-designed sensor and 10Hz for improved one. Thus, the PD pulses detected by such sensors are more likely to present a longer duration. Therefore, a high-pass filter is needed to generate a sharp impulsive waveform with rapid decrease time and remove the low-frequency interferences.

A) Selection of filter type

Two kinds of filters: hardware and software filters are often used in practical applications. In our research work, the detected TEV signals are analyzed in computer, a software-based filter is thus adopted to reduce the design cost.

The most frequently used types of filters are Butterworth, Chebyshev, inverse Chebyshev and elliptic filters. The optimal selection depends on the gain of filters. The examples of them are shown in Fig.3-7. All of these filters have same order $n = 5$. As illustrated in

Fig.3-7(a), the frequency response of Butterworth filter is designed as flat as possible in the passband. However, Butterworth filter rolls off slowly around the cut-off frequency when compared with other filters. Comparison of Fig.3-7(a), (b) and (c) demonstrates that the sharpness of out-of-band attenuation of the Chebyshev filter and inverse Chebyshev filter is higher than that of Butterworth filter, if one assumes equal filter order, passband ripple and stopband attenuation. Among all of the four filters, the elliptic filter has the sharpest response at cut-off frequency. However, it has equalized ripple behavior in both the passband and the stopband [157].

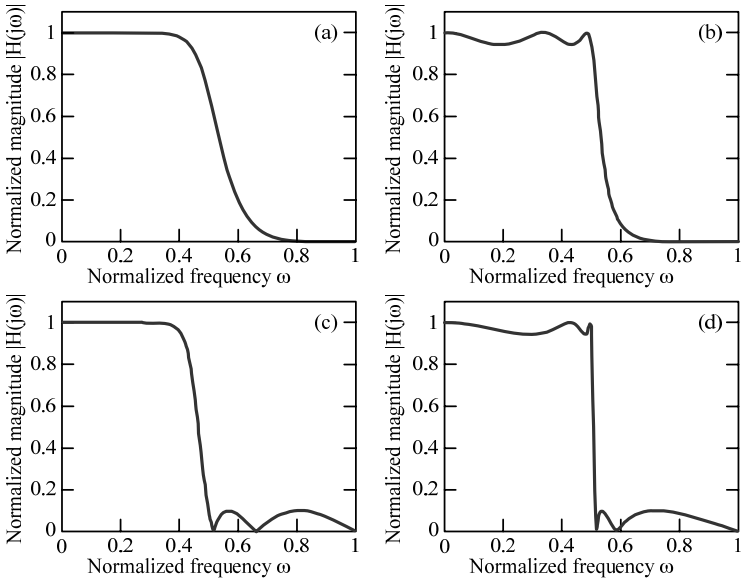

Fig. 3-7 Magnitude characteristic of different filters, (a) Butterworth filter, (b) Chebyshev filter, (c) inverse Chebyshev filter, (d) elliptic filter. All filters are fifth-order

When selecting a suitable filter, tradeoff is needed between the smoothness of passband and sharpness of attenuation. Due to the characteristics of original PD pulses and frequency response of TEV sensors, the detected PD pulses have wide and smooth frequency spectrum which suggests a lower requirement on the sharpness at cut-off frequency. Further, a flat response at pass band is preferred in further TF analysis. Therefore, the Butterworth filter with flattest response and simplest design is chosen.

B) Selection of filter order

The filter with Butterworth approximation is the simplest one whose basic concept is to approximate the desired frequency response by a Taylor's series. When designing a high-pass filter, the low-pass filter is often designed first as the prototype. The low-pass approximation of Butterworth filter is derived by assuming that $L(\omega^2)$ is a polynomial of the form [158]

$$L(\omega^2) = b_0 + b_1\omega^2 + b_2\omega^4 \cdots + b_n\omega^{2n} \tag{3-13}$$

where $L(\omega^2)$ is closely related with transfer function $H(j\omega)$, and $L(\omega^2) = \frac{1}{H(j\omega)H(-j\omega)}$. In order to approximate the filter response in a maximally flat sense, the coefficients in (3-7) should satisfy:

$$b_0 = L(0), \quad b_k = \left.\frac{d^k L(\omega^2)}{d(\omega^2)^k}\right|_{\omega^2=0} = 0 \quad \text{for } k \le n, \tag{3-14}$$

where n is the order of filter. Therefore, combining the two equations (3-13) and (3-14), we could get the general form of $L(\omega^2)$ as

$$L(\omega^2) = 1 + b_n\omega^{2n}, \tag{3-15}$$

and its form for a normalized approximation as follows:

$$L(\omega^2) = 1 + \left(\frac{\omega}{\omega_c}\right)^{2n}. \tag{3-16}$$

Here, ω_c is the cut-off frequency. Hence, the loss in a normalized Butterworth approximation is

$$A(\omega) = 10\log(L(\omega^2)) = 10\log\left(1 + \left(\frac{\omega}{\omega_c}\right)^{2n}\right) \tag{3-17}$$

With the designed prototype low-pass filter, the high-pass filter is converted via lowpass-to-highpass transformations. The difference between a lowpass and highpass filter is essentially an inverse: the frequencies below ω_c are mapped into frequencies above ω_c and vice versa [158]:

$$\omega = -\frac{\omega'_c}{\omega} \tag{3-18}$$

where ω'_c is the cut-off frequency of high-pass filter. If the cut-off frequencies of lowpass filter $\omega_c = 1$, then

$$A(\omega) = 10\log(L(\omega^2)) = 10\log\left(1+\left(\frac{-\frac{\omega'_c}{\omega}}{1}\right)^{2n}\right) = 10\log\left(1+\left(\frac{\omega'_c}{\omega}\right)^{2n}\right) \quad (3\text{-}19)$$

Obviously, the order of filter can be calculated if the parameters: ω, ω'_c, and $A(\omega)$ are given. According to the practical experiences, the cut-off frequency of high-pass filter ω'_c is set 100kHz [6]. The attenuation should reach -60dB at 1MHz, which means $A(1\text{MHz}) = -60\text{dB}$. Thus, the order of the filter is $n = 3$.

C) Transfer function of high-pass filter

Actually, the TEV signal that is saved in computer is the output of high-pass Butterworth filter, as in (3-17).

$$V_{TEV}(s) = H(s)V_d(s) \quad (3\text{-}20)$$

Here, $H(s)$ is the transfer function of high-pass Butterworth filter. To obtain the waveform of V_{TEV}, the parameters of $H(s)$ need to be studied first.

As the order $n = 3$ is used, the normalized transfer function of the prototype filter (low-pass third-order Butterworth filter) is

$$H_{LP}(s) = \frac{1}{(s+1)(s^2+s+1)} \quad (3\text{-}21)$$

To transform the low-pass filter into a high-pass filter, the equation in (3-12) should be rewritten as

$$s = -\frac{j\omega'_c}{s} \quad (3\text{-}22)$$

Substitute (3-19) into (3-18), we can have

$$\begin{aligned} H_{HP}(s) &= \frac{1}{(-\frac{j\omega'_c}{s}+1)((-\frac{j\omega'_c}{s})^2 - \frac{j\omega'_c}{s}+1)} \\ &= \frac{s^3}{(s-j\omega'_c)(s^2 - j\omega'_c s - \omega'^2_c)} \end{aligned} \quad (3\text{-}23)$$

The normalized transfer function of high-pass filter is

$$H_{HP}(P)=\frac{P^3}{(P-1)(P^2-P+1)} \text{ with } P=s/j\omega'_c. \quad (3\text{-}24)$$

Here, ω'_c is the lower cut-off frequency which is 100kHz. It is obvious that the exact parameters of the transfer function depends on the sampling frequency of detected TEV V_d.

3.4 Experimental test

As aforementioned, PDs inside metal enclosure generate EM waves and a part of them radiate out from gaps at insulated parts, gasket joints and cable insulation terminals and then produce TEV signal on the outside surface of the metal cabinet. A TEV measurement system was proposed to detect such signals. To illustrate this idea and test the proposed system, a laboratory test was set up and the PD signals collected by using different sensors at different locations are compared.

3.4.1 Measurement setup

The laboratory experimental setup is shown in Fig.3-8, where a PD generator was placed inside a metallic enclosure. This experiment system includes four parts: metallic enclosure, PD generation, sensors, and signal display and processing equipment.

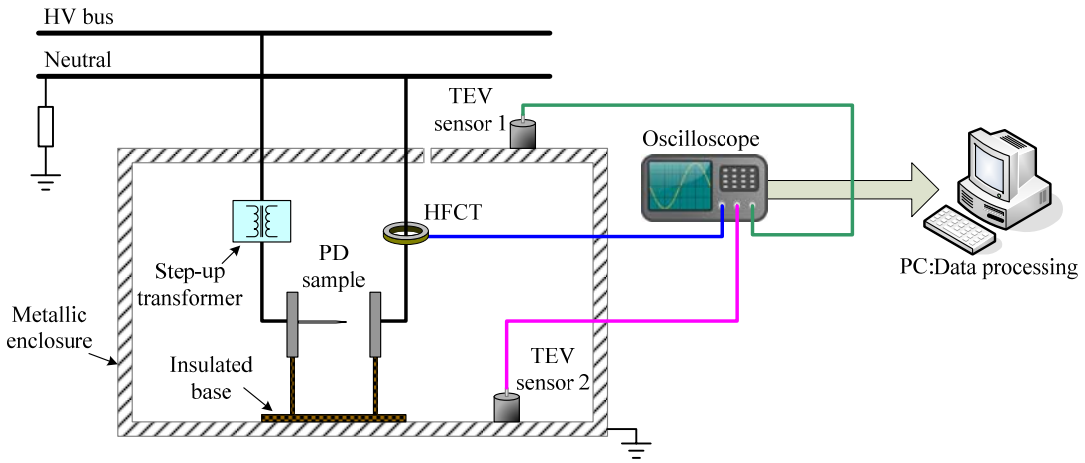

Fig. 3-8 Measurement setup of laboratory test

A) Metallic enclosure

In our experiment, an aluminum box is used to simulate the enclosure of metal-clad apparatus. The box has a dimension of 1.5m×0.8m×1m and a thickness around 8mm. Fig.3-9(a) and (b) shows the enclosure with its cover open and close, respectively.

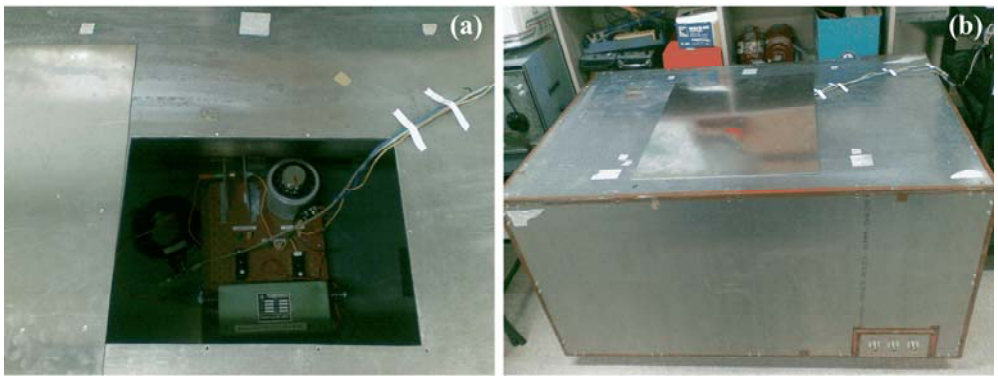

Fig. 3-9 PD generator placed inside a metallic enclosure, (a) enclosure with its cover open, (b) enclosure with its cover close

B) PD generation

The PD generating part contains PD defect samples and electric system which is composed by two transformers for high potential generation. The PD generation system is isolated from the metallic enclosure by an insulated base.

1) Discharge samples

Needle-to-plane discharges and cavity discharges are the most common causes of insulation failures in field tests of metal-clad apparatus. To study the features and effectively discriminate them from other interferences, these two typical defect samples with the structures of a needle-to-plane, and void in dielectric were fabricated. The structures of those two samples are shown in Fig.3-10.

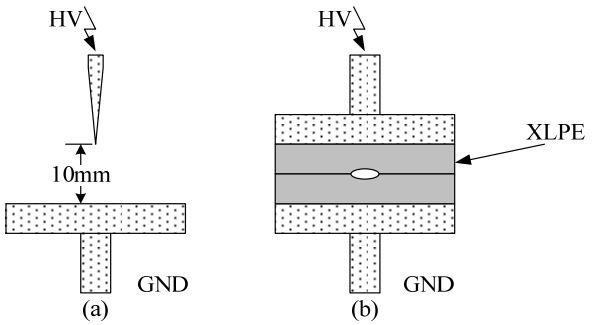

Fig. 3-10 Two types of discharge samples, (a) needle-to-plane discharge sample, (b) cavity discharge sample

a) Needle-to-plane sample

The experimental setup of a needle-to-plane defect sample is shown in Fig.3-10(a). The needle electrode is a stainless-steel needle, with a tip radius of about 5μm and a tip angle of about 30 degree. The bottom electrode is a cylindrical plane with a diameter of 25mm and a thickness of 12mm. The needle electrode is energized, and the bottom electrode is grounded. The distance between the needle tip and bottom electrode is 10mm.

b) Cavity sample

A cavity defect sample is shown in Fig.3-10(b) and its detailed parameters are shown in Fig.3-11. The basic idea of this sample is to make a void in the XLPE plates. XLPE is the abbreviation of Cross-linked polyethylene which is a form of polyethylene with cross-links. It is used predominantly in insulation for HV electrical cables. Since it is quite difficult to dig small cavities in XLPE materials, a cavity was made on the surface of one plate and two plates were fixed together to simulate the hollow void.

As shown in Fig.3-11(a), two smoothed XLPE plates with small cavities on the inner surface are fastened together by four plastic bolts. The inner surfaces of the two plates are close enough to avoid air existence between them. Two copper plates with a height of 1mm and a diameter of 9mm act as the electrode. When a high potential is applied on these two plates, a uniform electrical field will apply in the XLPE materials.

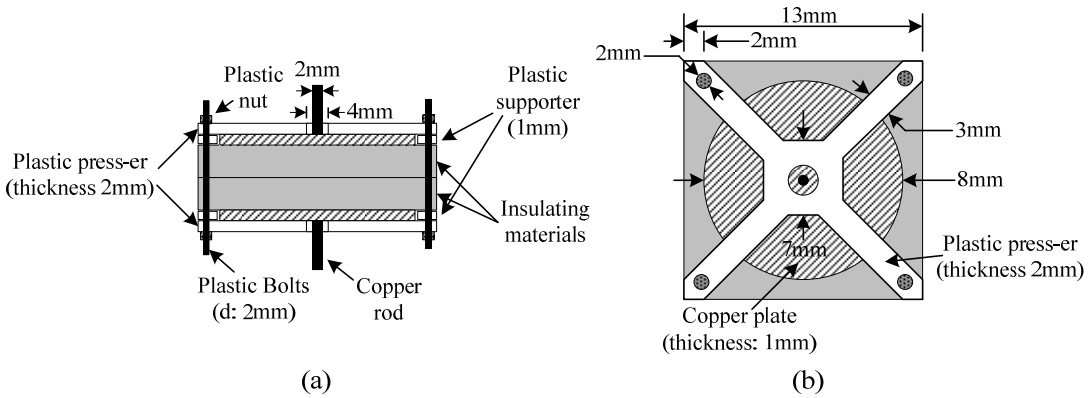

Fig. 3-11 Design of XLPE sample, (a) diagram of XLPE sample, (b) side view, (c) front view

2) Electric system

A schematic circuit diagram of the electric system of the PD test setup is shown in Fig.3-12. A variable transformer T_1 together with a step-up transformer T_2 are used as the voltage source. The output of variable transformer is connected to the primary side, or

low voltage side of the step-up transformer. The output voltage of T_1 ranges from 0V to 260V. The turns ratio of T_2 is 240/15k, and its output is connected to the top electrode of defect sample and the bottom electrode is connected to the ground. Hence, a high potential with maximum 15kV can then be applied on the defect sample.

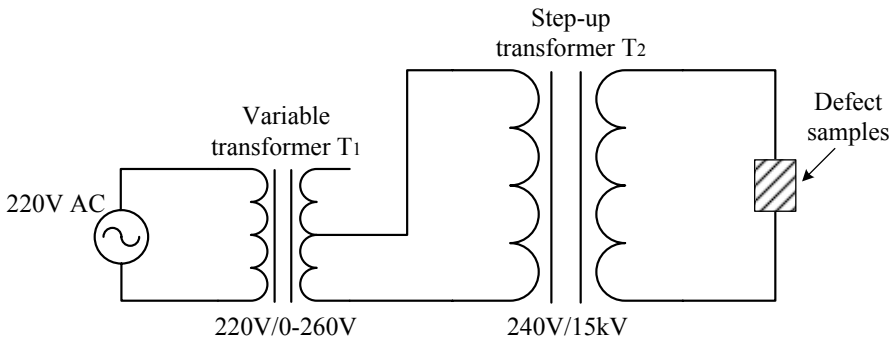

Fig. 3-12 Schematic circuit diagram of electric system

C) Sensors

To demonstrate the possibility of non-intrusive TEV measurement, the PD signals detected by both the TEV sensor mentioned in Section 3.3.1 and the industrial HFCT are compared.

Two similar TEV sensors are placed inside and outside the enclosure. The sensor placed outside the metallic enclosure has its inner electrode electric contact with the external surface of the top of the metallic enclosure. Similarly the sensor placed inside the metallic enclosure has its electric contact with the interior surface of the bottom of the enclosure.

In Fig.3-8, besides the PD generator, there is a HFCT placed inside the metallic box. The industrial HFCT sensor (type: IPEC OSM HFCT 140/100) is a large-size, split-core, high frequency current transformer with a frequency response from 100kHz to 12MHz. A probe with attenuation of one tenth connects the industrial HFCT to oscilloscope.

D) Signal display and processing

After collecting the PD signals via different sensors, the PD data is displayed on an oscilloscope (Tektronix TDS7104, band width: up to 1GHz and sampling rate: up to 10GHz/s) and then saved into a computer for further analysis.

3.4.2 Comparisons with direct detection

Using the two non-intrusive sensors and the HFCT, PD measurement was carried out in our laboratory. The detected results on oscilloscope are shown in Fig.3-13 for two different durations, where the top wave (blue) is measured PD pulses using HFCT; the middle wave (magenta) is the measured PD pulse by the sensor placed inside the enclosure; the bottom wave (green) is output from the sensor placed outside the enclosure. From the figures, one can see that the measured PD pulses are almost the same as from the two non-intrusive sensors placed inside and outside the metallic enclosure. However, due to the small lower cut-off frequency of non-intrusive sensors, the pulse waveform is heavily affected by low-frequency energy.

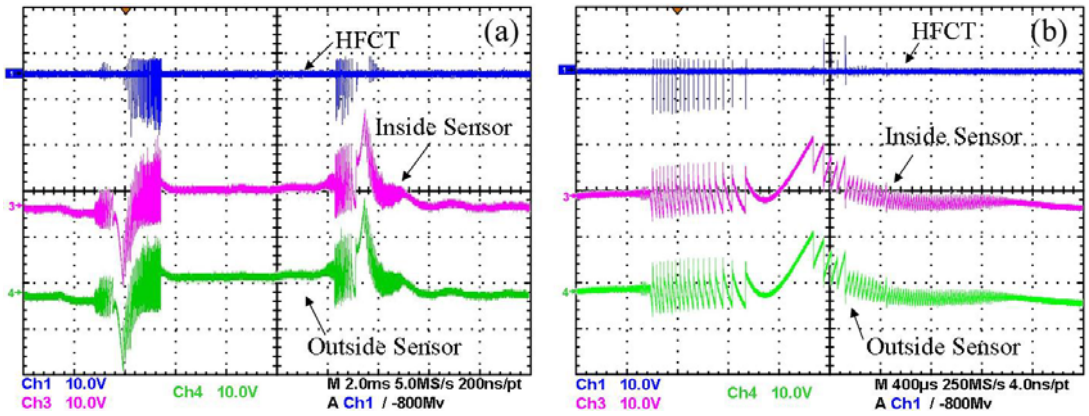

Fig. 3-13 Measured PD pulses at different locations, (a) PDs with duration of one cycle (20 milliseconds), (b) PDs with duration of 4 milliseconds

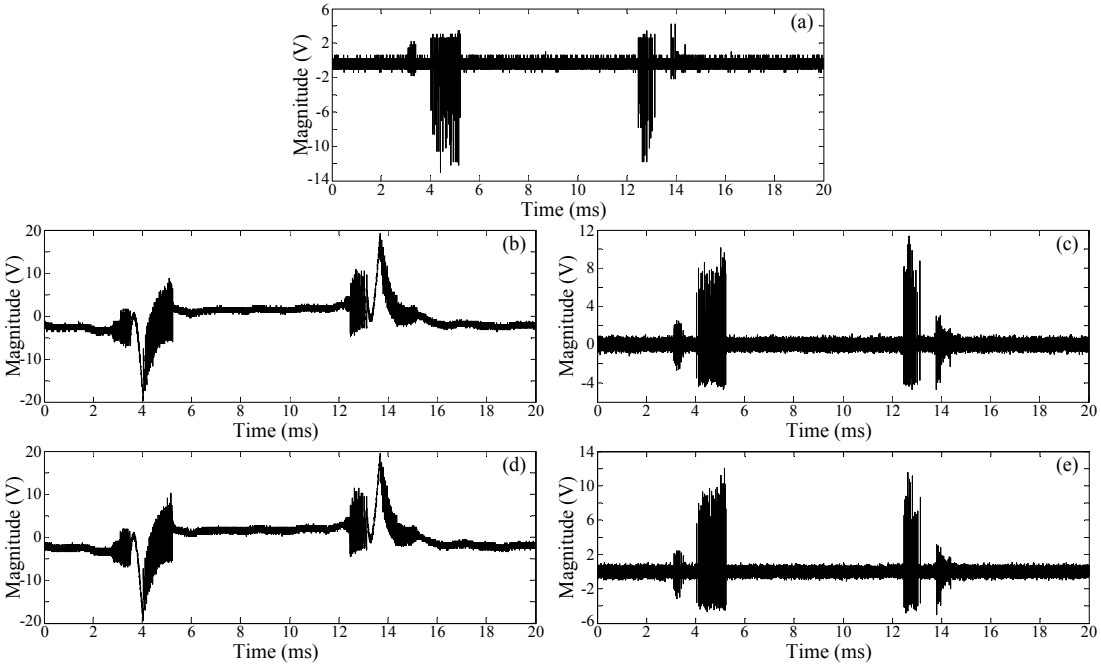

Fig. 3-14 TEV signals before and after filtering, (a) HFCT-detected signal, (b) signal collected by non-intrusive sensor inside the enclosure, (c) the signal in Fig.3-14(b) after filtering, (d) signal collected by non-intrusive sensor outside the enclosure, (e) the signal in Fig.3-14(d) after filtering

Fig.3-14 gives an example of PD signal before and after filtering. The magnified single pulses before and after filtering are also portrayed in Fig.3-15. Generally, some parameters such as rise time, decrease time and pulse height are employed to evaluate the waveform of a PD pulse [159]. The rise time refers to the duration that the magnitude increases from 10% to 90% of pulse height, and decrease time is defined as the duration that magnitude decreases from 90% to 10% of pulse height. As illustrated in the figures, the original PD pulses from non-intrusive sensors have an extremely short rise time of less than 1μs, and a decrease time of much more than 35μs. After filtering out the energy components in the frequency bands below 100kHz, the signals detected by non-intrusive sensors present an impulsive waveform. The rise time increases a little to 1μs. The decrease time changes greatly, and reduces to about 3μs. The pulse height also decreases from about 12V to about 9V. After filtering, the PD signals from HFCT and non-intrusive sensors at different locations are similar. This result verifies the theory of TEV and provides basis for field test of gas or oil insulated equipments using non-intrusive PD sensing technique.

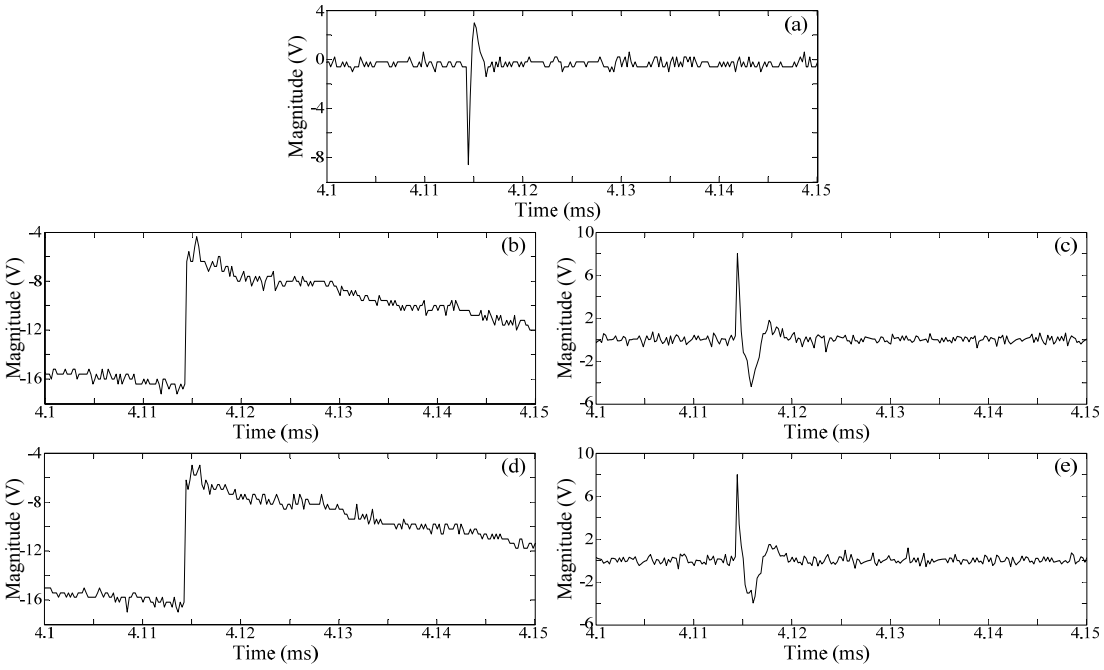

Fig. 3-15 Magnified single pulse before and after filtering, (a) magnified HFCT-detected signal, (b) magnified pulse collected by non-intrusive sensor inside the enclosure, (c) the signal in Fig.3-15(b) after filtering, (d) magnified pulse collected by non-intrusive sensor outside the enclosure, (e) the signal inFig.3-15(d) after filtering

3.5 Simulation of TEV signals

Simulation is considered as an effective method in signal analysis. The PD pulses detected by coupling capacitor method and UHF measurement were modeled and their

models were often used in theoretical analysis [39, 83]. Thus, the TEV signals collected by proposed system are also simulated for further analysis.

According to the characteristics of TEV signal mentioned in Section 3.1 and the features of proposed measurement system introduced in Section 3.2, the PD pulse in Fig.3-15(d) and (e) are simulated by combining the equations in (3-12) and (3-20). Both the simulated and measured pulses are shown in Fig.3-16. The parameters are set properly: $\alpha = 0.8\mu s$, $\beta = 450\mu s$, and the maximum amplitude is set 11V. In order to compare the two original TEV signals in a same figure, the starting point of simulated impulsive signal is shifted to -16.5V.

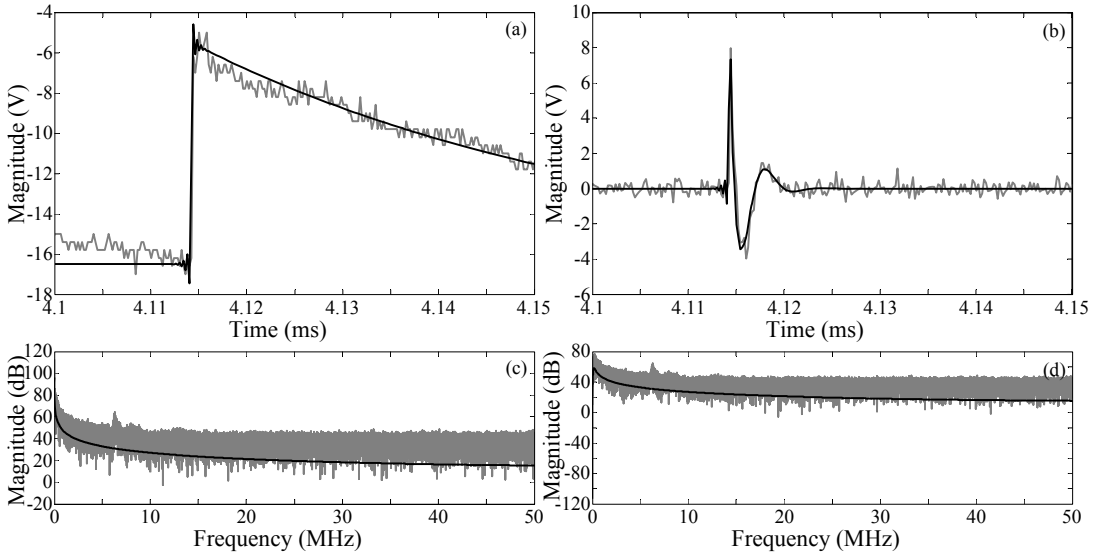

Fig. 3-16 Comparison between simulated and measured TEV signals, (a) the original TEV signals, (b) the filtered TEV signals, (c) frequency spectrums of signals in Fig.3-16 (a), (d) frequency spectrums of signals in Fig.3-16(d), the measured signal is marked in grey and simulated signal is marked in black.

As shown in Fig.3-16(a) and (b), only one pulse is simulated. The simulated signals are highly similar with the measured ones, especially the two pulses in Fig.3-16(b). Besides the time-domain waveforms, the frequency spectrums of measured and simulated signals are also similar. The only differences are the peaks at around 6MHz. Such difference is due to the transfer function of non-intrusive sensor which has a peak around 6MHz and is not considered in simulation. However, this difference is not so great that the measured signal can be well approximated and simulated with the models in section 3.5.1.

3.6 Conclusion

Partial discharge measurement with TEV technique is preferred in on-site detections of metal-clad apparatus. This chapter presents a TEV measurement system with non-

intrusive sensors and software based filters. The structures and characteristics of this measuring system are introduced. On the basis of analysis of TEV fundamentals and introduction of measuring system, experimental test was carried out in our laboratory to illustrate the effectiveness of non-intrusive measurement. The test results verified that the TEV measurement is effective in detecting the existence of PDs. Finally, the measured TEV signals are simulated. Comparisons between measured and simulated signals show the simulated signals can well present the features of TEV signal and be used in further theoretical analysis.

CHAPTER 4
OPTIMAL WAVELET THRESHOLDING FOR NON-IMPULSIVE NOISE REDUCTION

4.1 Introduction

Due to the external location of non-intrusive sensor, noise is always a major barrier for precise pulse detection. Wavelet thresholding was regarded as the most effective method for non-impulsive noise rejections and commonly employed in PD signal analysis [106, 114, 160, 161]. Although the theoretical and implemental aspects of its application have been explored and are well understood now, optimal thresholds and wavelets for TEV-detected PD signals are still not addressed fully. Further, due to the large amount of data and the requirement of high-speed processing for on-line measurement, the efficiency of de-noising procedure needs improvements.

This chapter investigates the selection of optimal thresholds and wavelets for wavelet thresholding of TEV-detected PDs as well as the efficiency improvements of the thresholding algorithm. First, the wavelet thresholding method is introduced in the aspects of wavelet transform, thresholding algorithm, and evaluation of de-noising. Next, based on the brief introductions of thresholds and wavelets, the optimal thresholding function, thresholds and wavelets are studied and selected with simulated PD signals under different conditions. Finally, in order to speed up the processing of measured data, a fast realization of wavelet thresholding based on paralleling computing is proposed.

4.2 Wavelet thresholding

PD pulse is usually a large-amplitude impulsive signal. Its energy is larger than that of noises in most frequency bands. Therefore, wavelet thresholding was proposed to remove non-impulsive noises, especially, the noises that follow Gaussian distribution. It was proved to be a more effective tool than others [162]. To have a clear understanding of wavelet thresholding and its performance, its fundamentals and evaluation are introduced in following contents.

4.2.1 Wavelet transform

PD signals carry a large amount of useful information which is difficult to find by using ordinary time or frequency domain analysis. The discovery of orthogonal bases and local time-frequency analysis opens the door to the world of sparse representation of signals. The orthogonal wavelet analysis uses a smaller number of coefficients to reveal the information of signal we are looking for [163]. The generation of these coefficients is an approximation of the original signal by linear combination of wavelets. For all f in $L^2(R)$,

$$P_j f = P_{j+1} f + < f, \psi_{j,k} > \psi_{j,k} \tag{4-1}$$

where $< f, \psi_{j,k} >$ stands for the inner product of f and $\psi_{j,k}$, P_j is the orthogonal projection onto a multi-resolution approximation space V_j which satisfies $\cdots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots$, closure $\{\bigcup_{j\in\mathbb{Z}} V_j\} = L^2(R)$ and $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$ [164]. Commonly, series of conjugate mirror filter pairs are used to decompose the approximation space V_j into a lower resolution space V_{j+1} and a detail space W_{j+1}, and project signal f onto different spaces. The two spaces, approximate and detail spaces, of a same scale j satisfy $V_{j+1} \perp W_{j+1}$ and $V_{j+1} \oplus W_{j+1} = V_j$ [164].

Since the approximate space denotes the lower frequency band and detail space represents higher frequency band in signal processing, the conjugate mirror filters are also called low pass filter and high pass filter, respectively. At decomposition, the wavelet coefficients of approximate space V_{j+1} (lower frequency band) are calculated with low pass filter $\bar{h}[k]$ and the coefficients of detail space W_{j+1} (higher frequency band) are calculated with high pass filter $\bar{g}[k]$ where $\bar{h}[k] = h[-k]$ and $\bar{g}[k] = g[-k]$. The approximate coefficients a_{j+1} and detail coefficients d_{j+1} are calculated as follow:

$$a_{j+1}[p] = \sum_{k=-\infty}^{+\infty} h[k-2p] a_j[n] = a_j * \bar{h}[2p],\ d_{j+1}[p] = \sum_{k=-\infty}^{+\infty} g[k-2p] a_j[n] = a_j * \bar{g}[2p]. \tag{4-2}$$

Accordingly, the filter banks at reconstruction are $h[k]$ and $g[k] = (-1)^{1-k} h[1-k]$, respectively [165]. The approximate coefficients in V_j are

$$a_j[p] = \sum_{k=-\infty}^{+\infty} h[p-2n] a_{j+1}[n] + \sum_{k=-\infty}^{+\infty} g[p-2n] a_j[n] = \breve{a}_{j+1} * h[p] + \breve{d}_{j+1} * g[p]. \tag{4-3}$$

Fig.4-1 shows the decomposition and reconstruction procedure with multi-resolution analysis. Here, only two decomposition scale is employed. a_0 denotes the space of original signal f.

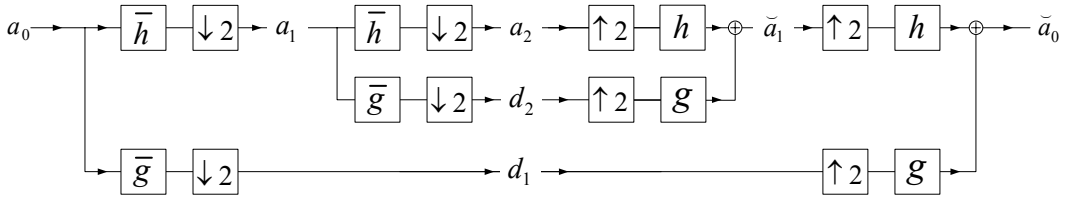

Fig. 4-1 Decomposition and reconstruction procedure with multi-resolution analysis

4.2.2 Thresholding algorithm

The main aim of wavelet thresholding in PD signal recovery is to recover the signal to be as similar with the original one as possible. The wavelet thresholding procedure for PD analysis includes four steps [105, 166]:

A) Decomposition

A filter bank of conjugate mirror filters decomposes the discrete signal in a discrete orthogonal basis. The wavelet function $\psi_{j,k}[n]$ and scale function $\phi_{j,k}[n]$ both belong to the orthogonal basis $B=[\{\psi_{j,k}[n]\}_{L<j\le J,0\le k<2^{-j}},\ \{\phi_{J,k}[n]\}_{0\le k<2^{-j}}]$. The scale parameter 2^j varies from $2^L=N^{-1}$ up to $2^J<1$, where N is the sampling rate of signal X.

In this step, appropriate wavelet should be chosen carefully. With different wavelets, the coefficients in the detail space W_j are different. An optimal wavelet only uses a few large-amplitude coefficients to represent an impulsive PD signal such that the small-amplitude noises can be removed as much as possible after thresholding.

B) Threshold estimation

Actually, de-noising with wavelet thresholding is a kind of noise estimation. Threshold is the estimated noise level in wavelet basis. The values larger than threshold are regarded as signal, and the smaller ones are regarded as noises. However, the estimation of noise level is possible only if some prior information is available. As most non-impulsive noises in PD measurement follow Gaussian distribution, for example, white noise, this distribution is used as the prior distribution of noises in threshold estimation. It was proved by Mallat [162] that the distribution of noise is not influenced by the

decomposition procedure. The non-impulsive noise remains white noise in orthogonal bases. Therefore, the estimators of white noise can also be used to estimate noise level in orthogonal basis.

So far, many kinds of thresholds are proposed to estimate the noise level after decomposition. They are calculated according to different estimation methods and effective for different applications. Therefore, thresholds need selection for particular applications.

C) Thresholding

After decomposition and threshold estimation, a recovered PD signal in the basis is written as

$$\tilde{F} = \sum_{j=L+1}^{J} \sum_{k=0}^{2^{-j}} \rho_T(< X, \psi_{j,k} >)\psi_{j,k} + \sum_{k=0}^{2^{-J}} < X, \phi_{J,k} > \phi_{J,k} \tag{4-4}$$

where $\rho_T(x) = a_m(x)x$ are the wavelet coefficients after thresholding. The function $a_m(x)$ is called thresholding function which usually includes soft and hard thresholding. Each thresholding function has their own advantages and generates different de-noised results. The selection of thresholding function is also needed with the considerations of thresholds.

D) Reconstruction

Finally, after thresholding, all the coefficients are used to reconstruct the de-noised signal.

4.2.3 Evaluation of de-noising

The output X acquired by non-intrusive sensor and PD measurement system can be regarded as a measurement of original PD signal f and noise signal W. After thresholding, the recovered PD signal can be denoted by $\tilde{F}$. The optimal wavelet thresholding is designed to minimize the error between original PD signal f and recovered one $\tilde{F}$. The mean-square distance is often employed to measure such differences. The mean-square distance is not a perfect model but it is simple and sufficiently accurate for one-dimension signal such as PDs [162]. The difference between f and $\tilde{F}$ is defined as the risk of estimation and calculated by (4-5):

$$r = E\{\| f - \tilde{F} \|^2\} \tag{4-5}$$

Here, the values of recovered signal $\tilde{F}$ heavily depend on the prior information available on the signal and the estimation methods it uses [162]. In other words, the methods of thresholding and the bases that are used in projection will result in different recovered signals and thus different estimation risks. Therefore, the estimation risk can be used to evaluate the performance of thresholding.

However, in practical applications, the estimation risk is rarely employed. The signal-to-noise ratio (SNR) which is measured in decibels and much more straightforward for understanding is commonly adopted to reveal the differences between original and recovered signal.

$$SNR_{dB} = 10 * \log_{10}(\frac{E\{\| f \|^2\}}{E\{\| f - \tilde{F} \|^2\}}) \tag{4-6}$$

where f is the original data without noise and $\tilde{F}$ is the recovered signal. It is easy to find for a certain original signal, larger estimation risk leads to smaller SNR, and vice versa.

4.3 Optimal threshold selection

The biggest challenge in wavelet thresholding is to find an appropriate threshold and suitable thresholding function. The appropriate threshold and thresholding function should be the combination that leads to smallest estimation risk and highest SNR. Some automatic thresholds such as universal threshold, minimax threshold and SURE threshold are regarded to have better performances in PD signal de-noising. Their de-noising capability with different thresholding functions for TEV signals are discussed and demonstrated by using simulated signals.

4.3.1 Thresholding functions

The thresholding function suggests the way that the thresholding algorithm revises the wavelet coefficients. Usually, the wavelet coefficients which are greater than threshold are kept or revised and the smaller ones are removed. For the noisy signal in orthogonal basis, the coefficients after thresholding can be written as $a_m(X_B[m])X_B[m]$ where a_m is the thresholding function. Commonly, according to the processing ways, the thresholding functions can be classified as hard and soft thresholding.

A) Hard thresholding

When employing hard thresholding, the wavelet coefficients whose amplitude exceeds threshold T is unchanged and the smaller ones are removed directly. A hard thresholding function is shown as follows [162]:

$$a_m(x) = \begin{cases} 1 & \text{if } |x| \geq T \\ 0 & \text{if } |x| < T \end{cases} \tag{4-7}$$

B) Soft Thresholding

Different from hard thresholding which keeps the larger-amplitude coefficients untouched, the soft thresholding method revises all wavelet coefficients. If the coefficients are greater than threshold T, their amplitude decreases. A soft thresholding function is implemented as follow [162]

$$0 \leq a_m(x) = \max(1 - \frac{T}{|x|}, 0) \leq 1 \tag{4-8}$$

The $\rho_T(x) = a_m(x)x$ that denotes the wavelet coefficients after hard thresholding and soft thresholding is portrayed in Fig.4-2.

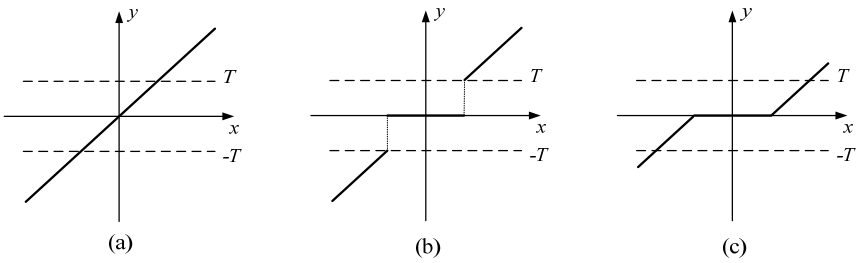

Fig. 4-2 Thresholding functions, (a) original signal, (b) hard thresholding, (c) soft thresholding

The large-amplitude coefficients which are untouched by hard thresholding ensures the magnitude of recovered signal $\tilde{F}$ is the same as the original one f. However, ripples and oscillating errors are often induced due to the reconstruction of unaffected noise coefficients whose magnitudes are only a little higher than that of threshold. On the other hand, in soft thresholding, the magnitude of coefficients that are greater than threshold is also reduced such that the amplitude of recovered signal $\tilde{F}$ is smaller than that of original one F. However, the errors which often cause low SNRs are less obvious. Therefore, in some cases where precise recovery of signal magnitude is not required, for example, image noise reduction, the soft thresholding is widely used since it can retain the

regularity of signal [167]. Otherwise, hard thresholding is preferred if precise recovery is required.

4.3.2 Popular thresholds

Previously, some paper proposed empirical thresholds which were claimed to be effective in PD recovery [135]. Because the experience based methods inherently have some difficulties in general applications, the automatic threshold estimations are more popular. In this section, three most commonly discussed automatic thresholds: universal threshold, minimax threshold and SURE threshold are introduced based on the brief study of noise variance estimation.

A) Estimation of noise variance

As an important kind of prior information, the variance σ^2 of noise W is closely related to threshold estimation and introduced before thresholds. As aforementioned, orthogonal wavelet transform only generates large-magnitude coefficients near the areas of major spatial activities. When most part of the PD signal f is piecewise regular, most coefficients contribute to the energy of noise and a few of them contributes to the energy of PD. The wavelet coefficients X_B approximate W_B. As the noise still follows Gaussian distribution in orthogonal basis, a robust estimator of variance can be calculated from the median of the fine-scale wavelet coefficients [160]. Different from mean value, median is independent of the magnitude of those few large-magnitude coefficients related with signal. Thus, the variance of white noise can be estimated from the median of absolute wavelet coefficients of fine scales by neglecting the influence from signal f [162]:

$$\tilde{\sigma} = \frac{M_X}{0.6745} \tag{4-9}$$

where M_X is the median of absolute wavelet coefficients X_B.

B) Universal threshold

In orthogonal wavelet transforms, the estimation of white noise is possible when most wavelet coefficients contribute to the variance of noise signal W_B. If the energy of PD signal is quite small and approximates zero $f_B \approx 0$, the wavelet coefficients X_B will have the same distribution, Gaussian distribution, with noise signal W_B. It has been proved that

the maximum amplitude of a vector of N independent Gaussian variables with variance σ^2 have a high probability of being just below $\sqrt{2\log_e N}$ [162].

Thereafter, Donoho and Johnstone [160] assumed this bound to be universal threshold T and proved that the risk of thresholding with universal threshold is small enough to satisfy the requirements of most applications. The universal threshold equals

$$T = \sigma\sqrt{2\log_e N} \tag{4-10}$$

where σ is the estimation of white noise and N is the size of signal and $N \geq 4$. As proved by Donoho and Johnstone in [161], the estimation risk of a thresholding $r_{th}(f)$ with universal threshold is almost the upper bound of possible risks as shown in (4-11):

$$r_{th}(f) \leq (2\log_e N + 1)(\sigma^2 + r_{pr}(f)) \tag{4-11}$$

where $r_{pr}(f)$ is the risk due to wavelet projector. One can conclude from (4-11) that the thresholding risk $r_{th}(f)$ is at most $2\log_e N$ times larger than the risk of a projector $r_{pr}(f)$. The factor $2\log_e N$ cannot be improved anymore by changing the estimators. Since the universal threshold has the largest value, it is possible to remove noises, which often results a nice visual appearance [168].

C) Minimax Threshold

The thresholding risk is often reduced by decreasing the value of threshold, for instance, choosing a threshold smaller than universal threshold. Therefore, according to the inequality in (4-11), minimax threshold is proposed.

According to the inequality in (4-11), the risk of thresholding can be presented in the form of $2\log_e N + 1$ times $(\sigma^2 + r_{pr}(f))$. Since the value of $(\sigma^2 + r_{pr}(f))$ only depends on the characteristics of vector X_B, it is natural and more revealing to look for a more 'appropriate' threshold λ which yields smaller possible constant Λ in place of $2\log_e N + 1$. Thus, the inequality in (4-11) can be rewritten as

$$r_{th}(f) \leq \Lambda(\sigma^2 + r_{pr}(f)) \tag{4-12}$$

Donoho and Johnstone [160] defined the minimax estimator which is designed to find the appropriate Λ that satisfies $\Lambda \le 2\log_e N + 1$, and the threshold $\lambda \le \sqrt{2\log_e N}$. At the same time, the threshold λ is the largest one that achieves the minimum bound of Λ.

Since the calculation of minimax threshold for different signals is a bit difficult and time consuming, an approximate one is commonly used in practical application. It is defined as

$$\lambda = \begin{cases} 0 & (N \le 32) \\ \sigma * (0.3936 + 0.1829 * \log_2 N) & (N > 32) \end{cases} \tag{4-13}$$

Due to the smaller threshold magnitude, the minimax threshold usually cannot generate a recovered signal with proper visual appearances. However, it has the advantage of giving good predictive performance [168].

D) SURE Threshold

Besides minimax threshold, other thresholds were proposed to suit the purpose of reducing the thresholding risk. The most famous one is the SURE threshold proposed by Stein [169]. The basic idea of SURE threshold is to estimate the means of independent Gaussian distributed random variables by using the mean square errors (MSE) as the estimation risk. Then the estimation risk of SURE thresholding could be denoted by [169]

$$r_{th}(f) = E\{Sure(X,T)\} \tag{4-14}$$

where X is the noisy signal and T is the threshold. For the wavelet coefficients in orthogonal basis, if the noise W_B is a Gaussian random vector with zero mean and variance σ^2, the noisy signal X_B equals $f_B + W_B$ and $E\{|X_B[m]|^2\} = |f_B[m]|^2 + \sigma^2$. Thus, the original signal $f_B[m]$ can be estimated from X_B. Then the estimation risk of SURE thresholding becomes

$$r_{th}(f) = E\{Sure(X,T)\} = E\{\| f - \tilde{F} \|^2\} = E\{| X_B + g(X_B) - f_B |^2\} \tag{4-15}$$

Here, $g(X)$ is a weakly differentiable function which is caused by thresholding [162]. The SURE threshold is the T that achieves the minimum estimation risk $E\{Sure(X,T)\}$ in (4-15). The SURE threshold with soft thresholding was proved to be unbiased which means the differences between the expected values and true values are zero. However, it still has

disadvantages: errors will be introduced if the threshold T is too small which occurs when the signal energy is much smaller than that of noise.

The three types of thresholds have their unique characteristics. The universal threshold which is generated with the largest estimation risk usually is larger than the other two kinds of thresholds whose estimation risks are smaller. When compared with minimax threshold, SURE threshold is usually smaller that means more small-amplitude coefficients are kept after thresholding. The values of thresholds commonly follow the inequality $T_{univ} > T_{\min i \max} > T_{sure}$.

4.3.3 Comparison of the thresholds under different conditions

Different thresholds and thresholding functions were proposed to deal with different conditions. Among the aforementioned thresholding functions and thresholds, the most appropriate combination for TEV-detected PD signals is discussed in this section.

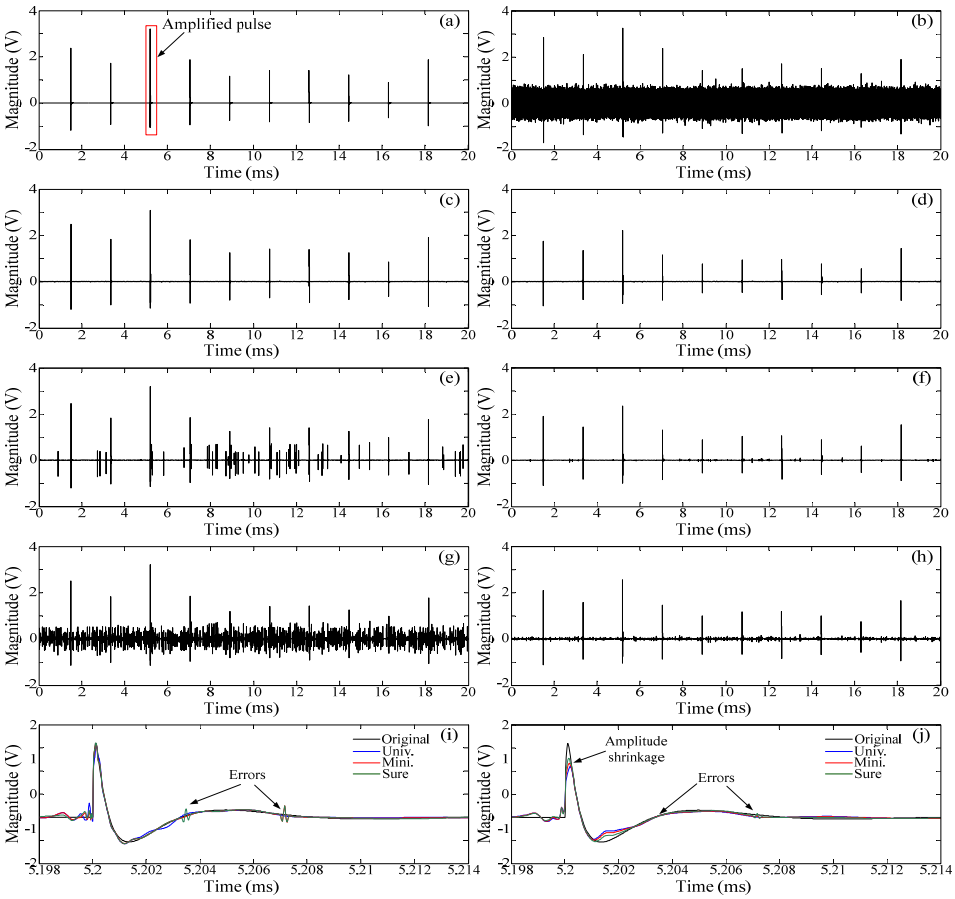

Fig. 4-3 Recovered signals by using different thresholds and thresholding functions, (a) original signal, (b) noised signal (SNR=-14.98dB), (c) hard thresholding with universal threshold (SNR=14.89dB), (d) soft-universal threshold (SNR=9.22dB), (e) hard-minimax threshold (SNR=12.68dB), (f) soft-minimax threshold (SNR=10.85dB), (g) hard-SURE threshold (SNR=3.74dB), (h) soft-SURE threshold (SNR=13.13dB), (i) magnified pulse recovered by hard thresholding, (j) magnified pulse recovered by soft thresholding

To give a clear understanding of the performances of PD recovery, a simulated PD signal and all recovered data by using different thresholding combinations are portrayed in Fig.4-3. This PD signal lasts 20ms and contains 10 pulses which are simulated in the same way as in Section 3.5. The average rise time and pulse width of the ten pulses is 2.5μs and 5.8μs, respectively. A white noise with mean zero and variance 0.2 is added.

It is obvious that the errors induced by minimax threshold and SURE threshold with hard thresholding heavily influence the appearance and SNRs of recovered signals. The errors are much smaller when using soft thresholding. For the single PD pulses shown in Fig.4-3(i), all the recovered pulse via hard thresholding have similar waveforms and amplitude with original pulse while obvious amplitude shrinkages are found in Fig.4-3(j) where soft thresholding is employed.

The PD recovery with only one simulated dataset is not enough to illustrate the performances of different thresholds and thresholding functions. Thus, a number of simulations under different scenarios are studied. Here, 4 groups of PD pulses are adopted and each group includes 10 PD pulses. All these PD pulses have different parameters. Scenarios with different decomposition scale, noise variance, rise time of pulse and pulse width are considered. To show the general performance of different thresholding methods, statistical estimation is employed. Since variances of the samples are unknown, the t-test is used. The 100(1-α) percent confidence interval on the true population mean is

$$\bar{y} - Z_{\alpha/2} S/\sqrt{n} \le \mu \le \bar{y} + Z_{\alpha/2} S/\sqrt{n} \tag{4-16}$$

Here, $\bar{y}$ is the expectation, S is the variance, n is the number of samples, $Z_{\alpha/2}$ is the value of t-distribution. In this research, the probability $\alpha/2$ is 0.05 which suggests the value μ falls in the interval $[\bar{y} - Z_{\alpha/2} S/\sqrt{n},\ \bar{y} + Z_{\alpha/2} S/\sqrt{n}]$ with a probability of 90%.

The estimated SNRs of different kinds of thresholding methods with expected value and confidence intervals are portrayed in Fig.4-4. The expectations are marked by grey triangles and the intervals are denoted by black lines. In some cases the intervals are very wide, for example, the Fig.4-4(b), whereas sometimes the intervals are very small. Since the lower cut-off frequency of TEV measurement system is 100kHz, a minimum decomposition scale 8 is required. Further, as too many decomposition scales will induce large-amplitude ripples, the maximum decomposition scale is 20. For the other three cases, the noise variance ranges from 0.25V to 2V, the average rise time of PD pulses

changes from 1μs to 10.5μs, and the average pulse width varies from 9μs to 80μs. In each scenario, only one parameter varies, and the others are fixed. By studying all estimated SNRs in Fig.4-4, the universal threshold with hard thresholding and SURE threshold with soft thresholidng always perform better than the others. In most cases, these two combinations have similar performances. When the noise level increases, the hard thresholding function with universal threshold shows better de-noising ability than soft thresholding function with SURE threshold. Therefore, both the universal threshold with hard thresholding and SURE threshold with soft thresholding are suitable for de-noising of TEV signals. However, when the noise level increases, the hard thresholding function with universal threshold may be more preferred than the soft one with SURE threshold[39].

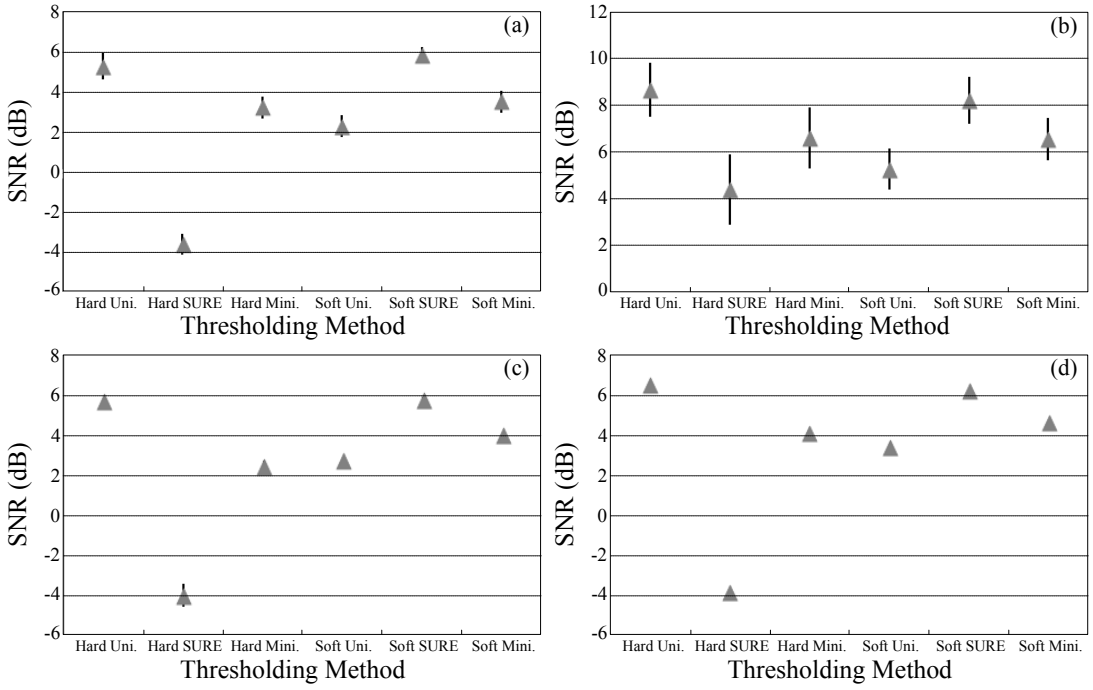

Fig. 4-4 Wavelet de-noising with different hard thresholds under different conditions, (a) different decomposition level, (b) different noise level, (c) different pulse rise time, (d) different pulse widths

4.4 Optimal wavelet selection

As mentioned in Section 4.3.2, the distribution of noise is not influenced by wavelet basis. However, the amplitudes of wavelet coefficients change when different wavelet is employed. Wavelet thresholding explores the ability of wavelet bases to approximate signal f with only a few non-zero coefficients. Therefore, choosing the wavelet bases that generate non-zero coefficients as few as possible is also an important factor in wavelet thresholding.

4.4.1 Properties for choosing a wavelet

The optimal wavelet needs to produce a large number of small coefficients and a few large-amplitude singularities such that the energy of PD which is often denoted by large-amplitude singularities can be easily extracted via thresholding. Choosing such an optimal wavelet depends on the properties of signal and wavelets such as regularity of signal, number of vanishing moments and size of support of wavelet.

A) Vanishing moments

The number of vanishing moments determines what the wavelet doesn't "see"[170]. Usually, the wavelet ψ has p vanishing moments if

$$\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0 \quad \text{for } 0 \le k < p . \tag{4-17}$$

This means ψ is orthogonal to any polynomial of degree $p-1$. Therefore, the wavelet with two vanishing moments cannot see the linear functions; the wavelet with three vanishing moments will be blind to both linear and quadratic functions; and so on. If the signal f is piecewise regular and its signal in a k ψ small interval can be approximated by a Taylor polynomial of degree , the wavelet can generate small coefficients at fine scales 2^j when the polynomial degree k is smaller than the vanishing moments p of wavelet. Although most coefficients at fine scales are close to zeros, the PD pulse also can be reproduced by scale functions by using the large singularities [165].

The measured PD signal X contains the original PD data f and noise signal W which follows Gaussian distribution. Since it is difficult to approximate the random variables of noise W, the lowest degree of Taylor polynomial of PD data f is the crucial factor in selecting number of vanishing moments. According to the TEV model in (3-20), while V_d is a combination of exponential functions, the lowest degree of Taylor polynomial of PD data f or V_{TEV} only depends on the order of Butterworth high-pass filter, which is 3 in our system. Thus, a wavelet with at least 4 vanishing moments can both reduce the amplitudes of noise and keep the energy of PD. However, it is not the higher the better. Too many vanishing moments may represent useful information of PD with smaller energy by quite small coefficients which will be removed after thresholding. Therefore, a proper number of vanishing moments should be a little larger than 3.

B) Size of Support

The size of support is the length of interval in which the wavelet values are non-zero. If the signal f has an isolated singularity at t_0 and if t_0 is inside the support of wavelet $\psi_{j,k}(t) = 2^{-j/2}\psi_{j,k}(2^{-j/2}t - k)$, then the wavelet coefficient $< f, \psi_{j,k} >$ may have a large amplitude. If ψ has a compact support of size N, at each scale 2^j there are N wavelets $\psi_{j,k}$ whose support includes t_0 [165]. In wavelet thresholding application, the signal f is supposed to be represented by a few non-zero coefficients. Thus, the wavelet with a smaller size of support is preferred.

However, the size of support is at least $2p-1$ if an orthogonal wavelet has p vanishing moments. Commonly, when choosing a wavelet from a group of candidates with suitable p , a trade-off between number of vanishing moments and size of support must be considered. If there are a few isolated singularities and the other parts of the signal is regular which means seldom PD occurrence, the wavelet with larger size of support can be employed to produce many small wavelet coefficients. Otherwise, it may be better to choose a wavelet with smaller size of support.

4.4.2 Wavelet families

To choose the appropriate wavelets for TEV-detected PD signals, the features of candidate wavelets should be studied first. As both orthogonal wavelets and biorthogonal wavelets can be used in orthogonal wavelet transform, Daubechies wavelets, symlets, coiflets and biorthogonal wavelets are discussed in this section.

A) Daubechies Wavelets

The Daubechies wavelets have minimum size of support $2p-1$ for a given number of vanishing moments p . However, in order to construct such wavelet, the smoothness as well as the symmetry of the wavelet filter has been sacrificed. The asymmetric filers of Daubechies wavelets cannot obtain linear phase which correspond equal delay to all frequencies and create large coefficients at the borders which lead to boundary distortions [165].. This property is not tolerable for some phase-sensitive applications such as communications. However, for the analysis of PD signal which does not have high requirements on phase information, the Daubechies wavelets were widely used and good performances were also reported [39].

B) Symlets

In order to find the wavelets with minimum support and least asymmetric filter, the Symlets which are also known as the Daubechies least asymmetric wavelets were proposed. The construction of Symlets is very similar to the Daubechies wavelets. They also have minimum size of support $2p-1$ for a given number of vanishing moments p. However, they have more symmetric wavelet filters.

C) Coiflets

Similar with Symlets, the Coiflets are also developed from Daubechies wavelets, but they have better symmetry and their scaling functions also have vanishing moments such that Coiflets were shown to be excellent for the sampling approximation of smooth functions [171]. However, the number of vanishing moments of Coiflets increases to two times of the order of approximation and the size of support extends to $3p-1$ instead of $2p-1$, where p stands for the number of vanishing moments.

D) Biorthogonal wavelets

All the orthogonal wavelets with minimum size of support cannot generate symmetric filters except Haar wavelet or Daubechies 1 wavelet. However, Haar wavelet is not well adapted to approximate smooth functions because it has only one vanishing moments. Therefore, perfect reconstruction is investigated by using biorthogonal wavelets which have minimum support. Biorthogonal wavelet bases are constructed with two pairs of perfect reconstruction filters (h,g) and ($\tilde{h},\tilde{g}$) instead of a single pair of conjugate mirror filters. Compared with orthogonal bases, the design of biorthogonal filters allows more degrees of freedom and it is possible to construct symmetric wavelet functions.

Table 4.1 Properties of different wavelet families

Wavelet name	Order	Number of vanishing moments	Size of support	Symmetry
Daubechies	$N_{\{1\le N<\infty\}}$	N	$2N-1$	Far from
Symlets	$N_{\{2\le N<\infty\}}$	N	$2N-1$	Near from
Coiflets	$N_{\{1\le N\le 5\}}$	$2N$	$6N-1$	Near from
Biorthogonal wavelets	$N_{d\{1\le N\le 8\}}$ for dec. $N_{r\{1\le N\le 6\}}$ for rec.	N_r for dec. N_d for rec.	$2N_d-1$ for dec. $2N_r-1$ for rec.	Yes

Note: 'dec.' is short for decomposition, 'rec.' is short for reconstruction

The properties of four wavelet families mentioned above are listed in Table 4.1 [165]. For the same number of vanishing moments, all the wavelets except Coiflets have minimum size of support. Biorthogonal wavelets are the only family with symmetric filters.

4.4.3 Comparison of the wavelets under different conditions

To choose the best wavelets for PD analysis, all the wavelet families mentioned in Section 4.4.2 are applied in the wavelet thresholding of some simulated signals. In order to explore the influence from vanishing moments, a large number, 15, is employed. Thus, the Daubechies wavelets with order 1 to 15, the Symlet 2 to Symlet 15, and almost all Coiflets and biorthogonal wavelets are included. The effect of the size of support is also studied by using two PD data with different density of pulses: one has 10 pulses per cycle (20ms) and the other has 30 pulses every 2ms. Similar with the comparisons of different thresholds in Section 4.3.2, the comparisons of different wavelets are done under different conditions: different variances of added white noises, different average rise time and average durations of PD pulses. Since the size of sample is large enough and the variance of the sample is unknown, statistical analysis in (4-16) is employed. Here, α is still 0.1 and thus the probability of SNRs falling in the estimated interval is 90%. The estimated SNRs with confidence intervals are shown in Fig.4-5 and Fig.4-6, respectively.

The estimated SNR distributions in Fig.4-5 and Fig.4-6 demonstrate the statements of vanishing moments in Section 4.4.1. The number of vanishing moments should be greater than 3 to filter out the noises but cannot be too large such that useful information is lost and worse de-noising results are generated. As shown in the two figures, the SNRs increase with the number of vanishing moments p and reach their maximums in the interval from 4 to 6, then decrease when p is greater than 6. Thus, the wavelets with 4 to 6 vanishing moments could produce better SNRs than others.

On the other hand, by comparing the estimated SNRs in two figures, it is not a hard job to find the influence from size of support. The estimated SNRs in Fig.4-5(c) and Fig.4-6(c), where PDs with different pulse widths are studied, show the de-noising capabilities of wavelets with larger sizes of support decrease with the increasing density of singularities. For example, the average estimated SNRs is around 28.1dB of coiflets and 27.1dB of Daubechies wavelets in Fig.4-5(c), but in Fig.4-6(c), the values are 27.7dB and 27.5dB, respectively. Thus, for the PDs with low density, all wavelets that have suitable number of vanishing moments can recover PD signal with similar SNRs. However, the Coiflets

with longer support could not perform de-noising as good as the others when pulse density increases or overlapping occurs.

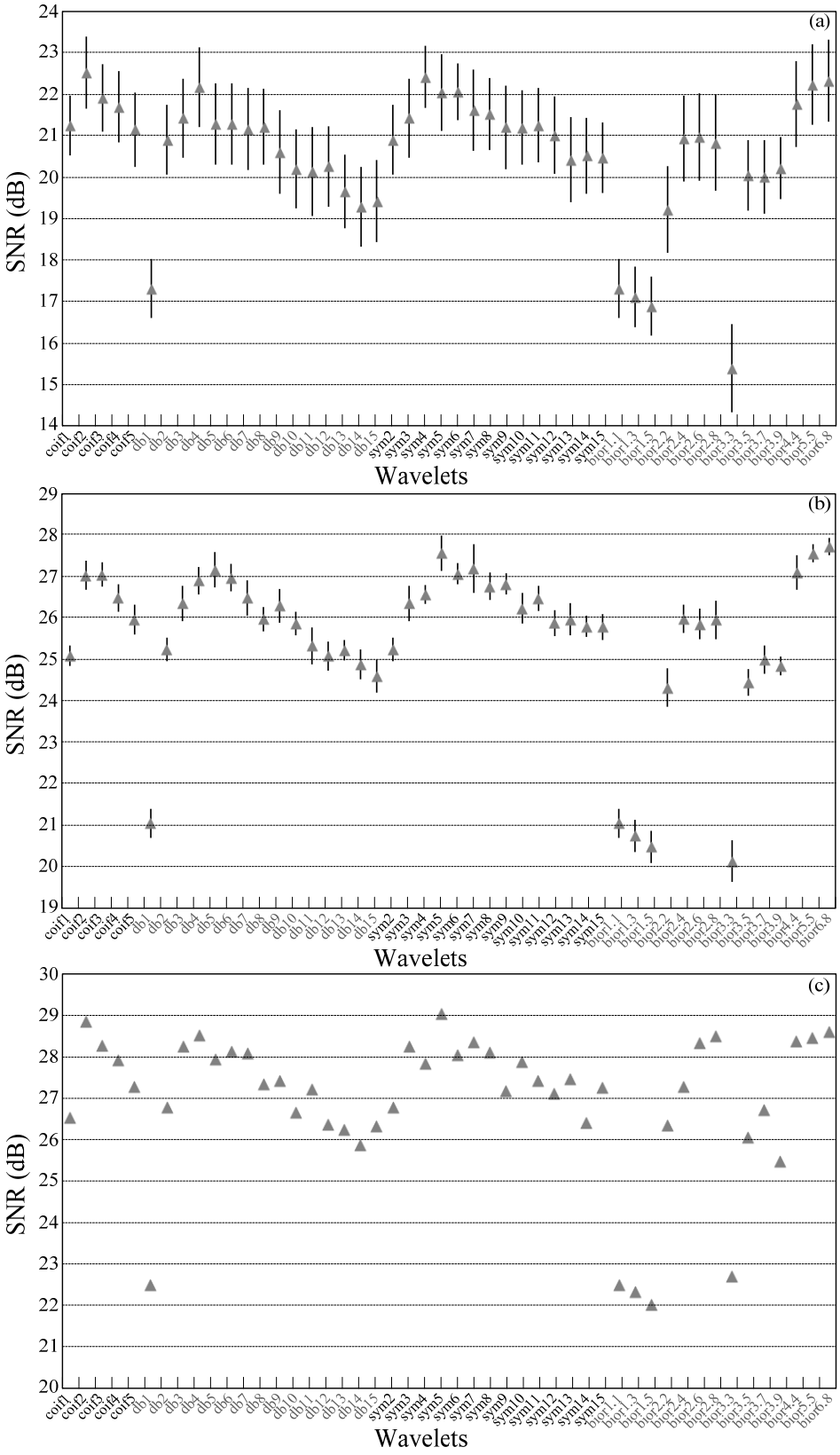

Fig. 4-5 Wavelet thresholding of low-density PD pulses with different wavelets, (a) SNRs of different noise variances, (b) SNRs of different rise times, (c) SNRs of different pulse widthes

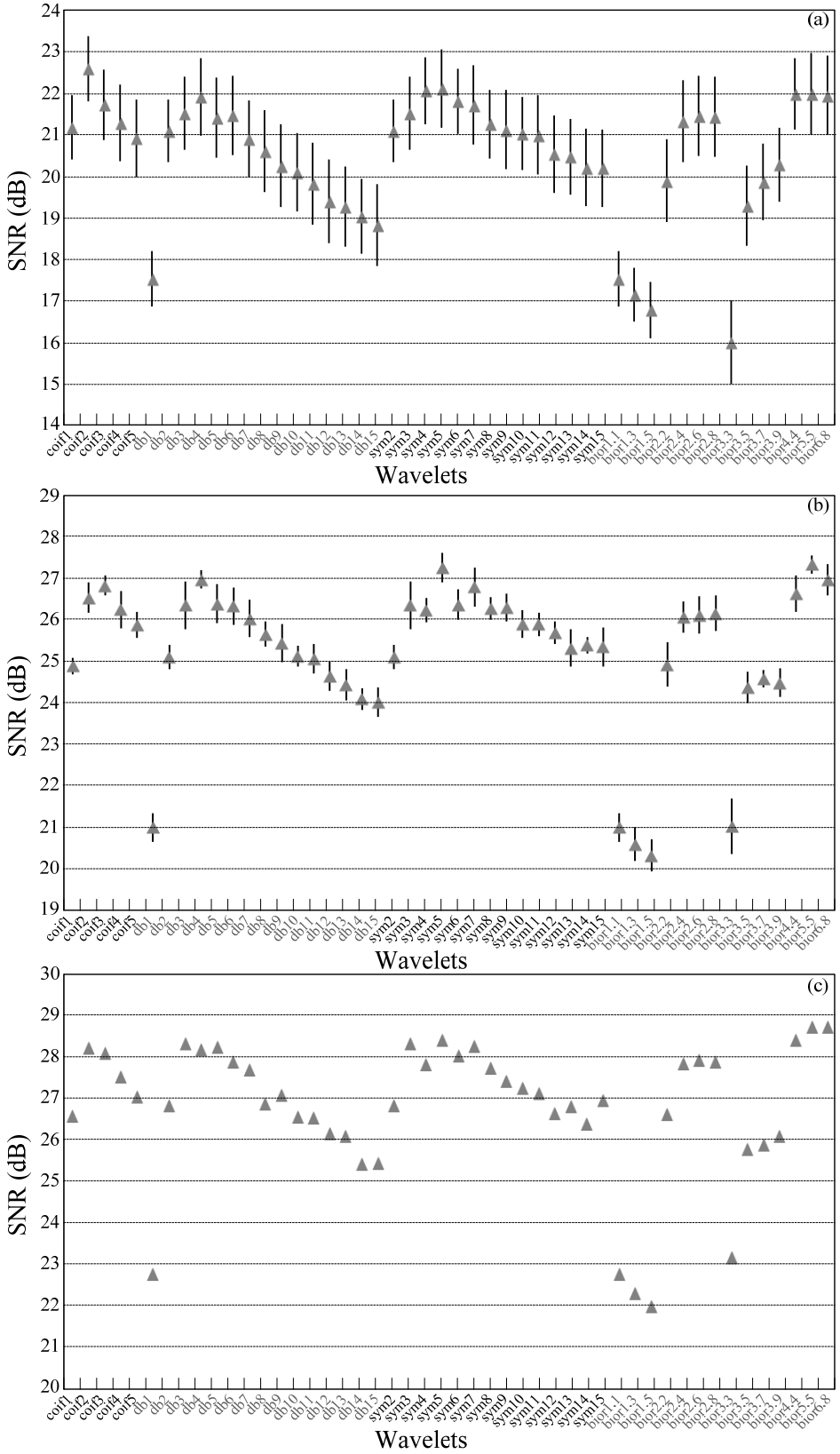

Fig. 4-6 Wavelet thresholding of high-density PD pulses with different wavelets, (a) SNRs of different noise variances, (b) SNRs of different rise times, (c) SNRs of different pulse widths

Furthermore, the more symmetric wavelets produce higher SNRs generally. The estimated SNRs of Daubechies wavelets are smaller than other more symmetrical wavelets when the number of vanishing moments is the same.

Therefore, one can conclude from the above discussions that the wavelets with number of vanishing moments from 4 to 6 are better than other ones, and the wavelets with minimum size of support are better than Coiflets when pulse density increases or overlapping happens. Among all the families, the SNRs by using wavelets with more symmetric filters are more likely to vary less when compared with asymmetrical ones.

4.5 Processing efficiency improvement

With the wide-band non-intrusive sensor and increment of data-acquisition speed, more data samples can be obtained within one cycle (20ms). Although more data points are necessary to analyze features of each PD and differentiate types of PDs, huge size of storage is required in practical applications. Thus, high-speed data processing is often needed to de-noise the measured data and judge the PD existence during on-line PD measurement, and finally reduce the storage requirement. Much previous work has been done in this area: Cheng *et al* used FPGA to speed up the processing procedure [172] and Ma *et al* explored the application of an intelligent DSP based analyzer [173]. In this section, the possibility of employing a software based method, parallelism, is investigated.

4.5.1 Primary considerations

Parallelism is a form of computation in which many calculations are carried out simultaneously [174]. The large problems are divided into small ones that are done concurrently. The main motivation of parallelism is to shorten the time of PD data-processing. It is consistent with the aim of real-time wavelet thresholding algorithms. However, some problems that need to face are the appropriate type of parallelism and suitable software environment.

A) Appropriate type of parallelism

There are different forms of parallelism: bit-level parallelism, instruction-level parallelism, data parallelism (loop-level parallelism), and task parallelism (function parallelism or control parallelism) [175]. Since the main topic of this chapter is focused

on wavelet thresholding of PD signal rather than the analysis of parallel computing, the simplest parallelism, task parallelism which is easiest in realization is considered.

Task parallelism requires the independency of each task. With the increase of sampling rate, more data points are acquired. However, the thresholding method is unchanged. If long signal segment is divided into small sub-segments and PD extractions for all sub-segments are done concurrently, it is possible to reduce overall processing time.

B) Suitable software environments

Currently, most PD analysis is done on the platform of MATLAB. However, the processing of MATLAB is not very fast and the processing speed of a faster environment, C, is investigated. To test the durations, two data segments with different sizes are used. The lengths of the two datasets are 1×10^3 and 1×10^4, respectively. In all the following tests, 'sym4' wavelet and universal threshold which are proved to be appropriate in previous sections are adopted. The decomposition level used here is 10. The durations of wavelet thresholding in different software environments are shown in Table 4.2.

Table 4.2 Durations in different environment

Data length	Durations of MATLAB	Durations of C
1×10^3	328ms – 343ms	< 1ms
1×10^4	343ms – 360ms	< 1ms

As shown by Table 4.2, the durations of program on MATLAB platform is much longer than those in C environment. Thus, the C program should be a better choice than MATLAB in realizing the fast wavelet thresholding algorithm.

4.5.2 Solutions to problems induced by parallelism

When task parallelism is employed, long signal is divided into small segments and thresholding of them are done simultaneously. After noise reduction, the segments are combined together. However, the segmentation may induce two main problems: reduced threshold and boundary distortion.

A) Reduced threshold

According to the definition of universal threshold in (4-10), its magnitude will decrease with the length of data N if the variance of white noise is almost the same. It is thus more likely to induce errors in the reconstructed signals.

To overcome this problem, the universal threshold is still calculated by using the length of original data rather than the revised segments. Therefore, the differences between the sequential and parallelism methods are only caused by the differences between the estimations of noise variances of each segment which are quite small and can be ignored in practical applications.

B) Boundary distortion

Rather than the influence from reduced threshold which often can be ignored, the boundary distortion between any two segments may be the most serious problem in the proposed parallelism based thresholding. Boundary distortions often appear to be obvious discontinuations at the boundary of two neighboring segments. They are often encountered in wavelet thresholding due to the slight errors produced during convolutions of filters and signal.

Some strategies, such as periodic extension and symmetrical extension, are introduced to eliminate this distortion [176]. However, no matter which extension it is, the extended data points near the boundary are only related with the segment itself. For example, in symmetrical extension, the extended data points are copied from data points near the boundaries of signal. However, in parallel computing program, the data segment can be extended when divided. If a proper number of data points of adjoining segment are included in extended segment, the distortions at the boundary of segment will decrease after reconstructions. For example, if the desired length of signal segment in each paralleling task is l, the extended segments should have a length of $l+2x$ with x points extended at both the front and end of the segment. After wavelet thresholding, the middle part of a size l is cut and used as the recovered signal. The boundary distortions of the cut segment will be less obvious than those without extensions.

In practical programming, a more convenient extension is adopted: only one side of first $n-1$ data segments of length l are extended with $2x$ data points, where n is the total number of segments. Then each segment has $l+2x$ points, except the last segment which

has l points. After de-noising process, only the first $l+x$ points of first segment, the last $l-x$ points of last segment and l points from each of the other left segments which are exclusive of their first x points and last x points are recorded to form the de-noised signal. In this way the distortion due to segmenting is minimized or even removed. The de-noised results by using directly-cutting parallelism, extended parallelism algorithm and their differences between results by using sequential method are shown in Fig.4-7. Here, the original PD signal is divided into four parts. The extended length $2x$ equals one tenth of the length of each segment l. By comparing the differences between the signals in Fig.4-7(d) and Fig.4-7(f), one can see that the boundary distortion almost disappears after extension.

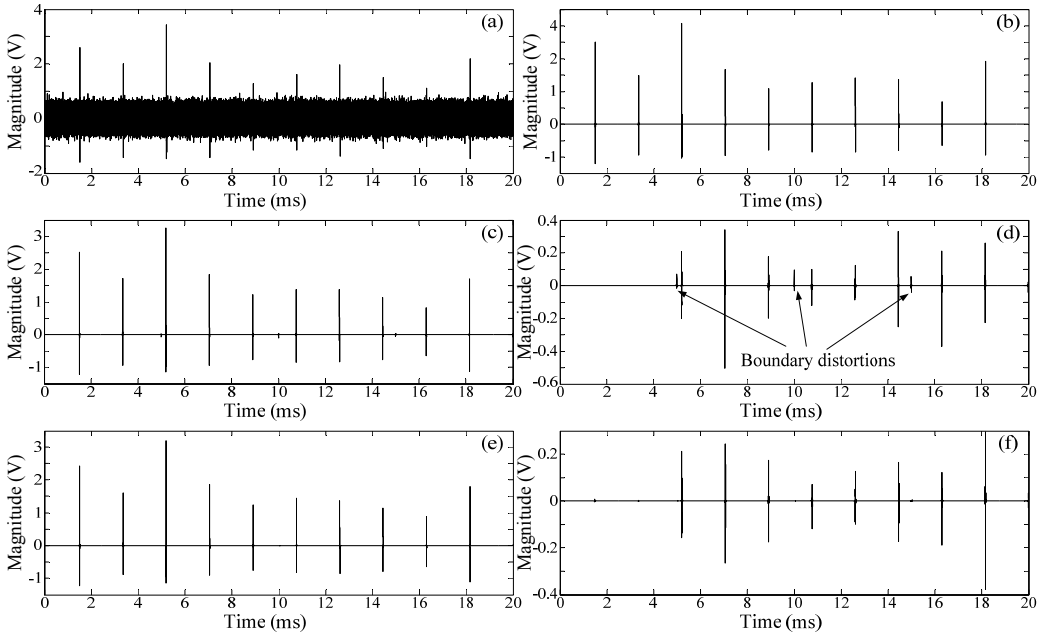

Fig. 4-7 Boundary distortions before and after extension, (a) noised PD signal, (b) recovered PD with sequential algorithm, (c) recovered PD using parallelism without extension, (d) differences between signals in Fig.4-7(b) and (c), (e) recovered PD using parallelism with extension, (f) differences between signals in Fig.4-7(b) and (e)

4.5.3 Comparisons of processing durations

The signals are processed in both MATLAB and C environment using sequential and revised parallel methods. The durations taken by recovering the signals with different methods are listed in Table 4.3.

All of those tests were done on a same computer with an Intel(R) Core(TM) i7 CPU. Here, wavelet 'sym4' is used and the decomposition scale is 10. The original data is divided into ten segments in parallelism. From Table 4.3, one can see that the duration by using parallelism in C environment is significantly shorter than that in other two cases.

Especially, the program with parallelism in C environment can complete the data-processing of 10^6 samples within 20 milliseconds which means real-time noise reduction for data with a sampling rate no higher than 50MHz.

Table 4.3 Durations of wavelet thresholding with different methods

Data length (sampling rate)	Seq. in MATLAB	Seq. in C	Para. in C
1×10^6 (50MHz)	1950ms	110ms	16ms
2×10^6 (100MHz)	2580ms	218ms	32ms

Note: 'Seq.' is short for 'sequential thresholding', 'Para.' is short for 'paralleling thresholding'.

4.6 Conclusion

This chapter discussed the selections of optimal thresholds and wavelets for wavelet thresholding based non-impulsive noise reduction for TEV-detected PDs and proposed a possible efficiency improvement for on-line measurement. Several popular thresholds, and thresholding functions are presented. Also, the wavelet families and their properties that will affect the performance of de-noising are studied. The de-noising capability of all the thresholds, thresholding functions and wavelets are tested by simulated TEV PD signals under different scenarios. The universal threshold with hard thresholding function and the wavelets with 4 to 6 vanishing moments are proved to be more appropriate than other combinations. With the optimal threshold, thresholding function and wavelet, an efficiency improvement method was presented. Since it only suggests a possible direction of high-speed PD processing for on-line measurement, the simplest parallelism algorithm was employed. The comparison of durations illustrates that the parallelism algorithm in C environment is possible to be used in future real-time PD noise reduction.

However, the wavelet thresholding method is only effective in rejecting white noise or sinusoidal harmonics. The rejection of impulsive noises is almost impossible by using thresholdings. Thus, the PD extraction from impulsive noisy background will be discussed in the following chapters.

CHAPTER 5
WAVELET ENTROPY BASED PD RECOGNITION BY USING NEURAL NETWORK

5.1 Introduction

Compared with non-impulsive noises, impulsive interferences are much more difficult to reject due to their high similarities with PD in time and frequency domains. In Chapter 4, wavelet transform has been presented as an effective tool to remove non-impulsive noise such as white noise. Because of its outstanding performance in time-frequency analysis, the capability of wavelet transform in rejecting impulsive interferences is therefore explored and investigated in this chapter.

This chapter presents a wavelet entropy based PD recognition algorithm which classifies PD and noise pulses by using a trained neural network (NN). It is desirable to explore the features of PD and impulsive noise by using wavelet transform with the aim to find appropriate representations of the pulses for neural network. Due to the large amount of wavelet coefficients, entropy which is a measure of disorder is employed to describe those features and reduce the feature dimensions. Next, fundamentals such as neuron models, network structure and training algorithms of neural network are introduced. Finally, the PD recognition system is presented and the performance of the proposed method is demonstrated by experimental results.

5.2 Investigation on signal features

The wavelet analysis which displays the local features in TF domain with real coefficients leads to a better discrimination than pure time or frequency domain methods [37]. The wavelet coefficients contain all the local TF features of a single pulse. However, due to the large amount of data, using wavelet coefficients directly is not a suitable representation of pulses, especially when they are employed as the input of neural network. An operator is thus needed both to effectively characterize the features and to reduce their dimension.

5.2.1 Properties of PD and noises

Wavelet transform is a suitable method in identifying sharp edge transitions [177]. By using the orthogonal decomposition with wavelet bases, the amplitudes of coefficients at fine scales are very small or almost zero when the signal is piecewise regular, and the large-magnitude coefficients only occur exclusively near the areas of major spatial activities [160]. Therefore, wavelet transform is an effective tool to reveal the characteristics of impulsive signals in as few coefficients as possible which leads to a more compact representation [178].

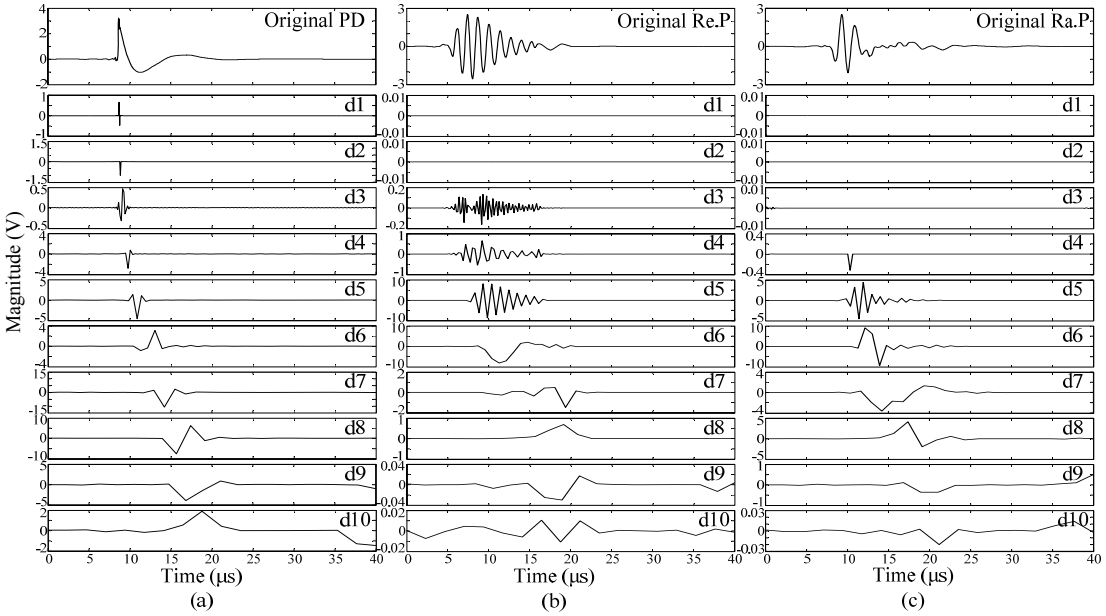

Fig. 5-1 Wavelet coefficients of PD and impulsive noises, (a) PD pulse and its wavelet coefficients, (b) repetitive noise pulse and its wavelet coefficients, (c) random noise pulse and its wavelet coefficients. Note: 'Re.P' is short for 'repetitive pulse', and 'Ra.P' is short for 'random pulse'.

The TEV-detected PD pulses have extremely short rise time, just several microseconds with an ultra-wide-band detection. They are well suited to the use of wavelet transform. The wavelet coefficients of a single PD pulse collected in experimental test via non-intrusive sensor are shown in Fig.5-1(a). The signal is sampled at the rate of 50MSamples/s, the wavelet is 'Coif 2' and the decomposition level is 10. As shown in this figure, large-amplitude singularities can be found in all decomposition levels. It suggests a wide frequency range of PD pulse. By comparing the maximum magnitude of coefficients of each level, it is easy to conclude that most of the energy of the PD pulse is concentrated at lower frequency bands such as the seventh (d7) and eighth (d8) levels.

In order to find the unique PD features, a comparison between PD and impulsive noises is done. Fig.5-1 also portrays the wavelet coefficients of two examples of impulsive noises:

one is a repetitive pulse from electronics equipment and the other is random noise. Both noisy signals are collected from field tests via TEV measuring system.

As the TEV measurement is a local PD detection method and the sensors are mounted on the external surface of cladding, the impulsive noises that propagate alone the metallic tank surface have to travel a relative long way before being captured. Due to the surface resistant of metal, the high frequency energy of pulses has been attenuated greatly during transmission. This results in a narrower frequency spectrum of impulsive noises. The wavelet coefficients of repetitive noise in two higher frequency bands, the first (d1) and second (d2) levels, are almost zero as in Fig.5-1(b) while three levels, d1, d2, and d3 levels of random pulse contain a large number of zeros as in Fig.5-1(c). It is clear that the impulsive noises have narrower frequency spectrums than PD pulses.

Further, in TEV measurements, the PD pulses which are collected with ultra-wide-band sensors often have less oscillating components. However, due to the different mechanisms of different pulse types and the distortions during propagation, the repetitive pulses and random pulses often have oscillating components with large energy. As illustrated in Fig.5-1(b) and (c), the coefficients with largest amplitude of repetitive noise are in the fifth (d5) and sixth (d6) levels, and the largest energy of random pulse is in sixth (d6) level.

The pulses portrayed in Fig.5-1 are typical ones of each type of pulses. Such comparison can provide a general idea about the frequency characteristics of them. According to the comparisons between noises and PDs, the impulsive noises whose sources are usually far away from sensors have narrower frequency spectrum and large-energy contained in higher frequency bands. The differences between repetitive pulses and random noises should also be noticed. First, the large-energy of repetitive noises is usually in higher frequency bands while the large-energy frequencies of random pulse are lower. Second, repetitive pulses are highly similar with each other, but each random pulse is different from the others. Therefore, an appropriate and stable operator is needed in the following content to characterize the different types of pulses.

5.2.2 Wavelet entropy based feature extraction

The discussions in Section 5.2.1 on the wavelet coefficients of PD and noises demonstrate that the wavelet analysis has an inherent capability to describe the signal spatial

characteristics [178]. However, the wavelet coefficients cannot be used directly as the large number of coefficients is not a suitable representation for further classification. An operator is needed to reduce the feature dimensions.

According to the analysis of properties of PD and noise pulses, describing the distributions of wavelet coefficients seems to be a possible direction for pulse characterization. Therefore, entropy which is stable and commonly used in measuring disorder is introduced and its effectiveness of distinguishing pulses is also demonstrated by comparing with the most common way of characterization: energy distribution.

A) Fundamentals of entropy

Entropy is originally a thermodynamic property that describes the available energy in a working system. After long time of development, entropy theory became more and more popular and was introduced into many other areas such as information theory and signal processing where it was used as a measure of uncertainty and disorder. Such wide application enriched the concept of entropy and pushed the development of entropy research in other areas. Entropy was first applied in power system for the control of generators [179]. Thereafter, many different applications of entropy were employed in different aspects of power system, for example, transient signal analysis and power quality evaluation [180]. However, entropy has never been used to characterize a single pulse in PD analysis. Since more chaotic signal generates greater entropy, the energy distribution of a single pulse can be regarded as a dynamic system. When a singularity appears, such system varies from order to disorder. Thus, the concept of entropy could be an ideal candidate to analyze the dynamic variations of a single pulse. In our research, the term entropy refers to Shannon entropy. The entropy H of signal X with possible values $\{x_1, x_2, \cdots x_n\}$ is defined as follow [181]:

$$H(x) = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i) \tag{5-1}$$

where $p(x_i)$ is the probability of x_i, and b is the base of logarithm. Common value of b is 2, and the unit of entropy is 'bit' accordingly. Equation (5-1) shows that the value of entropy only depends on the distribution or probability rather than on amplitude of the coefficients.

When characterizing a single pulse, the entropy values of all decomposition scales are calculated. The wavelet coefficients of each level form the signal X, and the coefficients

in each X produce one entropy value H. To make all the entropy values fall in a single same interval which makes the comparison and illustration more straightforward, the idea of entropy ratio is employed in practical application. As in (5-2), entropy ratio ρ_{Ei} equals the entropy value of each level divided by the norm of entropy vector of all levels.

$$\rho_{Ei} = \frac{H_i}{\| H \|} \tag{5-2}$$

Here, H_i is the entropy of ith level, and H is the entropy vector that constitutes of the entropies of all levels. By using (5-2), the amplitudes of ρ_{Ei} fall in the interval [0 1].

B) Effectiveness comparison

As entropy is a measure of the distribution of wavelet coefficients, its value is unaffected by the magnitude of singular points. To illustrate the advantages of applying entropy as the operator, a comparison between entropy and energy features of wavelet coefficients is presented.

For a generalized comparison, a group of PDs, repetitive pulses and random pulses are included. Each type contains 20 pulses that are collected with TEV sensors from the same sources where the pulses in Fig.5-1 are measured. Furthermore, the same decomposition level and same wavelet are used. The entropy ratio vectors which are the mean values of 20 entropy distributions of different pulse types are portrayed in Fig.5-2(a). The entropy ratio distribution of PDs which decreases gradually from low to high frequency levels is totally different from those of impulsive noises which rise first to reach their peaks at oscillating frequency bands and then decrease at high frequency levels. The differences between entropy ratio vectors of repetitive and random pulses are also obvious: the entropy ratios of random pulse have wider spectrum and the ratios of different levels vary less than energy ratios, but the spectrum of repetitive pulse is narrower and its large-amplitude ratios concentrate in only a few levels.

Also, a simple characterization with energy is used for comparison. Since the wavelet energy reveals the frequency spectrum straightforwardly, it is very common to employ the energy distributions to represent a PD signal [106]. The energy distribution indeed has performed perfectly in some applications. However, the energy value is heavily dependent on the coefficients magnitude. Similar with entropy ratio, the ‘energy ratio’

which normalizes the energy distribution into the same interval [0 1] is calculated with the equation in (5-3).

$$\rho_{engi} = \frac{E_i}{\| E \|} \tag{5-3}$$

In (5-3), the energy of each level equals the Euclidean norm of wavelet coefficients, which equals $E_i = \| C_i \| = \sqrt{\sum_{j=1}^{J} c_{i,j}^2}$ where $c_{i,j}$ is the j th wavelet coefficient on the i th level. As shown in Fig.5-2(b), the trends and distribution of energy ratios of different types are similar: all of them increase from low frequency levels first and decrease after reaching their peaks. Further, their peaks appear at frequency bands that are not near to each other. The peak of PD is on level 8 and the other two are on level 6.

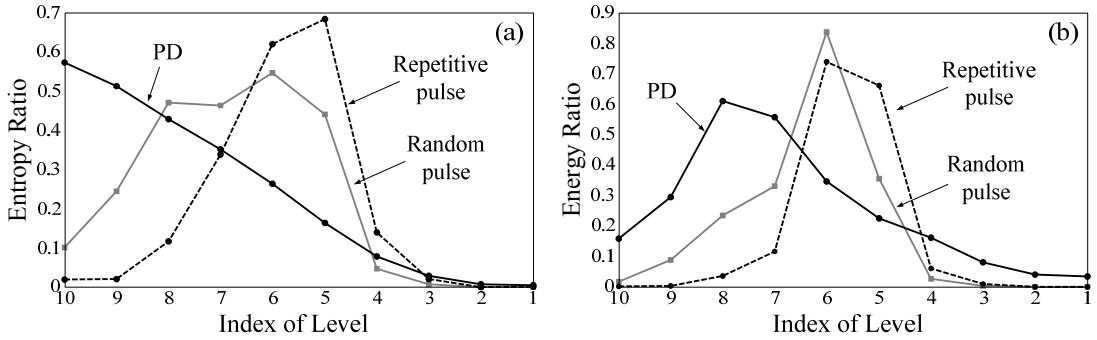

Fig. 5-2 Comparison of entropy with energy, (a) mean entropy ratios of three pulse groups, (b) mean energy ratios of three pulse groups.

More than ratio distributions shown in Fig.5-2, the comparison is quantified by calculating the distances between the ratio vectors of any two of the pulse types. The distance is defined as in (5-4):

$$d = \| A - B \| \tag{5-4}$$

where A and B are the ratio vectors of two different types, and $\| X \|$ is the Euclidean norm or Euclidean distance. The distances of entropy ratios and energy ratios between different types are shown in Table 5.1. The values in Table 5.1 are the estimated distances between the ratio vectors of any two different types. Here, the expectation and confidence intervals are calculated according to equation (4-16) with α=0.1. Since 20 pulses are included in each type, 400 (20×20) distances are thus considered in statistical analysis.

As demonstrated in Table 5.1, all the estimated distances between entropy ratios of different types are greater than those of energy ratios. That means the differences between

entropy ratio vectors are larger than energy ones. The entropy is more effective than energy distribution to represent different pulse patterns.

Table 5.1 Distances between different pulse types

d	PD-Re.P	PD-Ra.P	Re.-Ra.P
Entropy Ratio	$1.016 \leq d \leq 1.023$	$0.7495 \leq d \leq 0.7715$	$0.5655 \leq d \leq 0.5873$
Energy Ratio	$0.9302 \leq d \leq 0.9406$	$0.5349 \leq d \leq 0.5747$	$0.5064 \leq d \leq 0.5288$

5.3 Description of neural network

Just like the brain of human being, the neural network (NN) or artificial neural networks (ANN), provides a brain-like capability for solving problems [182]. The training progress of ANN studies the samples and finds the patterns of inputs and the relationships between inputs and outputs. Due to the excellent classifying and recognizing abilities, NNs are very popular and have been applied in many different tasks [183]. Especially, they become more and more popular in PD diagnosis [184, 185]. The most widely used network, feed-forward back-propagation (BP) network, is employed in the proposed method.

5.3.1 Model of neuron

Similar with human brain which consists a network of a large number of interconnected neurons, the ANN is constructed by many individual cells that can process small amount of information and activate other cells to continue the process [186]. These information-processing units which are called neuron or node in ANN theory are the fundamentals of the neural network operation. A typical model of neuron is shown in Fig.5-3.

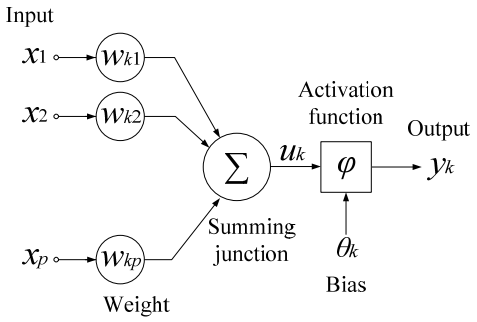

Fig. 5-3 Model of a neuron (after Haykin [187])

A typical neuron model usually contains three basic elements: (i) a set of synapses x_i, each of which is characterized by a weight w_{ki}, (ii) an adder that sums all the input signals such that $u_k = \sum_{j=1}^{p} w_{kj} x_j$, (iii) an activation function φ that limits the amplitude of the output y_k of a neuron. A threshold or bias θ_k is employed to shift the value of output [187]. Therefore, the output of neuron k can be described as

$$y_k = \varphi(u_k - \theta_k) \tag{5-5}$$

The weight w_{ki} and bias θ_k are adjusted during training.

5.3.2 Feed-forward network

Feed-forward neural network is the simplest and earliest type of ANN that has been applied in pattern recognition [188]. In this network, the signal data flows in only one direction, forward, from the input layer, through the hidden layers and to the output nodes. There is no cycles or loops in the network [183]. Therefore, the feed-forward network is also called direct network.

Three kinds of layers are present in feed-forward neural network architectures: input layer, hidden layer and output layer. The networks can be classified into two kinds by including hidden layer or not: single layer network and multi-layer network. The single layer network only contains input and output layers. It can only represent linear separable functions and is seldom used in PD diagnosis where PDs are hard to be discriminated from impulsive noises with linear functions. The feed-forward network that is commonly adopted in PD diagnosis is multi-layer network which includes one or more hidden layers [182, 189]. The network is enabled to extract higher order statistics by adding more hidden layers [187]. However, the network with more than one hidden layer requires a large size of input, which in turn reduces the training efficiency. Therefore, in practical application, the feed-forward network with only one hidden layer is highly recommended [183].

A typical feed-forward network with one hidden layer is portrayed in Fig.5-4. The nodes in the input layer (first layer) provide respective information for the nodes in hidden layer (second layer). Then the outputs of all hidden layer nodes (second layer) are used as the input of next layer, the output layer (third layer). Each node represents one neuron unit as

in Fig.5-3. Therefore, the inputs of neurons in each layer of the network are only related with the outputs of the preceding layer [187]. The output of the whole network is the overall response of the network to the input pattern supplied by nodes in the first layer. Commonly, appropriate determinations of the number of nodes in each layer can help in improving the efficiency and capability of the network and make it more suitable for particular applications.

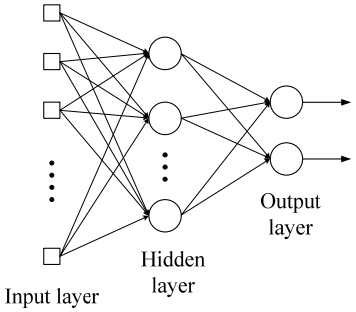

Fig. 5-4 Fully connected feed-forward neural network with one hidden layer

5.3.3 Back-propagation Algorithm

To train the neural network, the calculated outputs from the training network and desired results are compared. As in (5-6), the training error of the node j is defined as the degree that denotes how much the calculated output y_j matches desired output d_j.

$$e_j = y_j - d_j \tag{5-6}$$

Therefore, the training of a network is in fact a procedure to minimize the errors. To train a network, the weights and bias of each neuron unit must be modified, because the output y_k of each neuron depends on the values of weights w_{kj} and bias θ_k for a certain set of input x_j.

The term "back-propagation" suggests the way that a network treats with errors. The values in weight matrix and bias variables are determined by comparing the error signal that propagates backward. Therefore, in feed-forward networks with BP algorithms, there are two data flows in the architecture: signal flow and error flow which transmit in opposite directions. Fig.5-5 depicts the two flows in a part of the feed-forward BP network. The signal flow is actually the input signal that comes from the input layer, propagates through the network neuron by neuron (from left to right), and emerges at the end of the network as an output signal. On the other hand, the error signal flow originates

at the end of the network or the output layer, and propagates backwards through the network (layer by layer) [190].

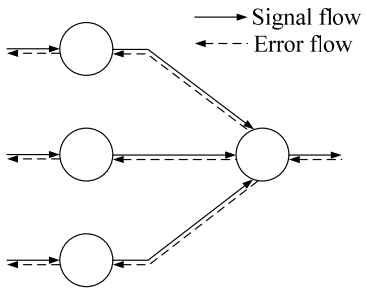

Fig. 5-5 Directions of signal data flow and error signals (after Haykin [190])

The values of weights w_{kj} and bias θ_k are first randomly selected and then revised repetitively in a number of iterations. This procedure stops if the errors satisfy the training requirement. The arithmetic method that the training procedure uses to update the weights and biases is called training function. Although many kinds of training functions were proposed, for example, the most traditional method: gradient descent optimization, only several types of fast algorithms are preferred and often mentioned in practical applications [191]:

A) Quasi-Newton optimization

The weights and bias in each iterations x_{k+1} equal their values of last iteration x_k added with the corrections Δx_{k+1}. The basic step of quasi-Newton optimization is

$$x_{k+1} = x_k - A_k^{-1} g_k \tag{5-7}$$

where A_k is the Hessian matrix which is the partial derivatives of error function, and g_k is the gradient that is a function of errors, weights and activation method [190]. This method is similar to Newton optimization but no calculation of second derivatives is required and only an approximate Hessian matrix is calculated [192]. Thus, it is much faster than the traditional gradient descent method.

B) Levenberg-Marquardt algorithm

Like quasi-Newton optimization, the Levenberg-Marquardt (LM) algorithm also approaches a high-speed optimization without calculating the accurate Hessian matrix. It approximates the Hessian matrix in feed-forward networks to be

$$H = J^{\mathrm{T}} J\,, \tag{5-8}$$

and the gradient to be

$$g = J^{\mathrm{T}} e \tag{5-9}$$

where e is the vector of network errors, and J is the Jacobian matrix that contains first derivatives of network errors with respect to the weights and biases [193]. Therefore, the updated values become

$$x_{k+1} = x_k - [J^T J + \mu I]^{-1} J^T e \tag{5-10}$$

In (5-10), the addition of a unit matrix μI helps improving the converging speed.

C) Bayesian regulation

In the application of neural networks, one of the most serious problems is over-fitting. When training the network, the errors are small enough and satisfy the training requirements. However, the errors increase greatly if new data is employed. In order to modify this problem, regulation of the performance of neural networks is needed. The most famous one is Bayesian regulation proposed by MacKay [194], and this regulation method is often applied in combination with Levenberg-Marquardt algorithm for a better performance.

All of these algorithms are effective and fast in training a neural network. However, which one is best depends on the particular applications they are used for.

5.4 PD recognition system

With the investigations and introductions in Section 5.2 and 5.3, the wavelet entropy is likely to be a suitable characterization of pulse features, and neural network whose parameters and functions are carefully selected might act as an excellent classifier for PDs and impulsive noises. The combination of wavelet entropy and ANN was studied in other research field, but it is the first time adopted in PD analysis. In this section, attentions will be drawn to the PD recognition system based on wavelet entropy and ANN. First, the processing procedure of this system is presented briefly. Then, the details of each processing step are introduced. Especially, the behaviors of network under different conditions, for example, optimal selection of the number of neurons in input and output layers and the size of hidden layer, are discussed such that the network with optimal

architecture is used in PD recognition. Some PD and noise signals are used to test the efficiency of network during training. All the signals used in this section are collected by TEV method with non-intrusive sensor and sampled at a rate of 50 MSamples/s.

5.4.1 Recognition algorithm

The aforementioned pulse features of both PD and impulsive noises, and the investigations of neural network suggest the following PD recognition algorithm, as in Fig.5-6. Although the aim of this chapter is to reject impulsive interferences, the reduction of non-impulsive noises is still included in this system to improve the SNR of recognized signal. Therefore, four main parts are contained in this algorithm:

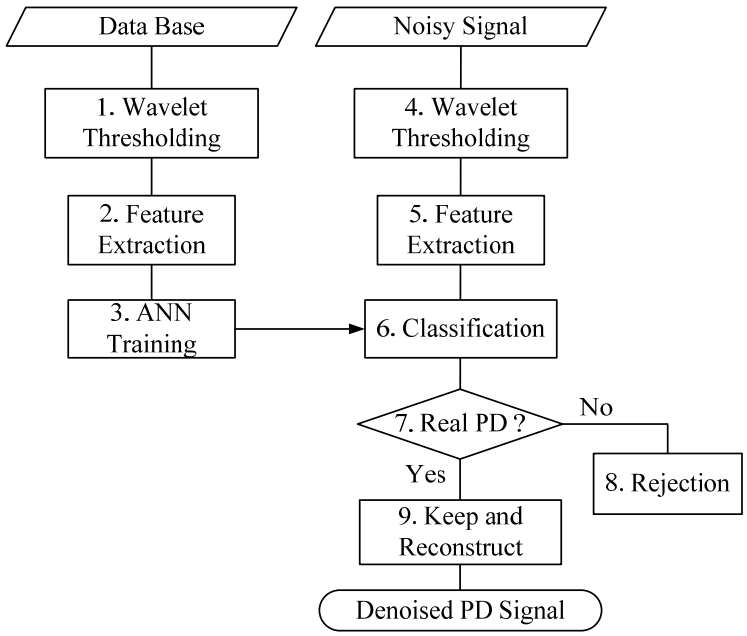

Fig. 5-6 Flowchart of proposed noise rejection method

1. Removing non-impulsive noises by thresholding (step 1 and step 4). The sinusoidal interferences and white noise are removed here. This step is needed to minimize the influence from non-impulsive noise in further analysis.

2. Extracting the features of pulses by using entropy (step 2 and step 5). After thresholding, the entropy only reveals the energy distributions of a single pulse segment at each decomposition scale.

3. Training the neural network with a large dataset of entropy features (step 3). The parameters of network, for example, the nodes of different layers, are carefully selected to ensure its best performance in PD pulse diagnosis.

4. Classifying the extracted features and recognizing real PDs (step 6 to 9). Each pulse-contained segment in the polluted PD signal is recognized by the trained network. The wavelet coefficients of pulses that are classified as real PDs are kept and reconstructed while the others that are regarded as noises are rejected and deleted.

The details of each step are described in the following content.

5.4.2 Wavelet thresholding

The details of wavelet thresholding are discussed in Chapter 4 and will not be repeated here. However, the settings of its parameters are briefly studied. To minimize the influence from non-impulsive noise as much as possible, the universal threshold with hard thresholding which can remove more noises is employed in this algorithm. As mentioned in Chapter 3, the lower cut-off frequency of TEV measurement system is 100kHz. The decomposition scale of wavelet transform should be large enough to ensure all the pulse energy is included in the entropy vector which is calculated from all detail coefficients. As all the signals are sampled at 50MSamples/s, the smallest decomposition scale is 8 and the widest frequency range of lowest approximate coefficients is from DC to 97.65kHz ($50\text{MHz}/2^{8+1}$) accordingly. However, as discussed in Section 4.4.3, SNRs of wavelet thresholding of PD signals sampled at 50 MSamples/s are almost highest and stable when the decomposition scale ranges from 9 to 18. Therefore, the minimum and maximum level is chosen to be 9 and 18 respectively.

5.4.3 Feature extraction

After thresholding, only the large-amplitude coefficients which are related with pulse energy are contained in each decomposition level. Their distributions are characterized by entropy which is introduced in section 5.2.2. The entropy values of all detail levels are calculated with the equation in (5-1) to ensure that all the energy of each pulse segment are included in the input pattern. Each decomposition level generates one entropy value and all entropy values form a vector with a size J, where J is the decomposition scale and $9 \le J \le 18$. The procedure of feature extraction is illustrated in Fig.5-7.

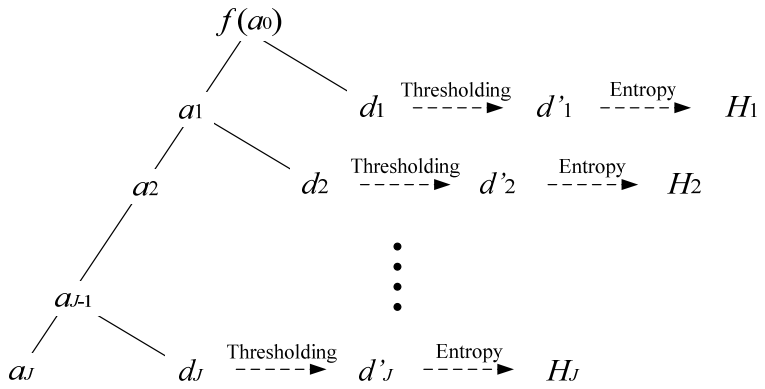

Fig. 5-7 Fundamental of entropy based feature extraction

5.4.4 Training of neural network

Once suitable features are extracted from the wavelet coefficients, a classifier must be constructed based on the extracted features to classify PDs and impulsive noises. The feed-forward BP network with many merits such as simplicity and ease of handling is employed as the classifier.

As mentioned in Section 5.3, the neural network could only perform well for a particular application if appropriate parameters and functions are selected. Therefore, before training the network, the parameters and functions of network are discussed and selected carefully in this section to improve its efficiency and performance in PD pulse recognition.

A) Activation function

According to the descriptions in Section 5.3.1, the output of a neuron depends on the value of weights, bias and the property of activation function if the inputs of this neuron are given. Different from weights and bias that will be modified during network training, the activation function needs selection before training. So far, the sigmoid function is the most commonly used activation function. It is a strictly increasing function that exhibits smoothness [187]. However, training with symmetric activation function is more likely to converge in a shorter duration than asymmetric ones [190]. A simulation of training ANN with different activation functions has been done. The average training duration (average of 20 trainings) of sigmoid function is around 2.63 sec while that of hyperbolic tangent function is around 1.35 sec when all the other parameters are the same. Thus, in our algorithm, the revised sigmoid function that is in the form of hyperbolic tangent function, as shown in Fig.5-8, is selected as the activation function.

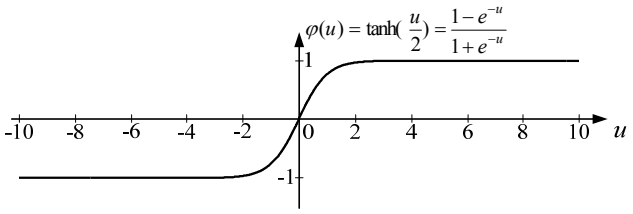

Fig. 5-8 Hyperbolic tangent function

B) Training function

Training function indicates the ways that the neural networks update the values of weights and bias. There are many types of training functions but only certain kinds of fast optimization algorithms are suitable for practical applications, for example, quasi-Newton (QN) optimization, Levenberg-Marquardt (LM) algorithm and Bayesian regulation (BR).

In this section, the optimal training function for PD diagnosis is analyzed. The fast-speed functions mentioned in Section 5.3.3 are evaluated by the mean-squared-errors (MSE) between trained results and desired responses of particular network architectures. The MSE is defined as follow:

$$E = \frac{1}{2}\sum_{j\in C} e_j^2 = \frac{1}{2}\sum_{j\in C}(y_j - d_j)^2 \tag{5-11}$$

where C includes all the nodes at output layer. The mean values of MSWs of four groups are calculated. Each group is formed by some networks with same input and output nodes but different sizes of hidden layer. The training procedure will stop when average MSE of network reaches 10^{-4}, where the average MSE equals E/N and N is the size of output vectors. The mean MSEs of four groups with selected structures are shown in Table 5.2. The structures of neural networks (NNs) are denoted by the size of different layers which are in the order of "input-hidden-output".

Table 5.2 Mean MSEs of different functions with different neural networks

NNs / Functions	$9-x-1$ $(2\le x\le 8)$	$9-x-2$ $(3\le x\le 8)$	$10-x-1$ $(2\le x\le 9)$	$10-x-2$ $(3\le x\le 9)$	Means
QN	0.2825	1.9485	0.0416	0.9908	0.8158
LM	0.0416	0.3784	0.2867	1.5603	0.5667
LM with BR	0.0041	0.0037	0.0033	0.0035	0.0037

Note: x is the numbers of nodes on hidden layer.

The MSEs in Table 5.2 point out the errors of networks trained with LM algorithm and Bayesian regulation are much smaller than the others. The smaller MSEs suggest a better

recognizing and classifying capability. Therefore, the Levenberg-Marquardt algorithm which is improved by Bayesian regulation is adopted in the proposed PD recognition system.

C) Optimal size of network

One of the most important problems of neural network application is choosing its optimal size. As mentioned in Section 5.3.2, the multi-layer feed-forward BP network with only one hidden layer is commonly used in PD pulse diagnosis. Its optimization involves the selections of the size of input layer, size of output layer and most importantly, the size of hidden layer. According to the conditions of wavelet entropy features of pulses and the requirements of PD recognition, the optimal size of NN is discussed.

1) Nodes of input and output layer

The input layer is a conduit through which the external environment, or the extracted features, presents a pattern to neural network. The entropy vector that consists of entropy of all detail coefficients are used as input vector. Each input represents an independent variable that has an influence over the output of the neural network [183]. In order to minimize the influence from amplitude of input values, all the entropy vectors are mapped to the interval [-1 1] such that the minimums and maximums of each entropy vector equal -1 and 1, respectively. The entropy based feature vector has a size of J. As discussed in Section 5.4.2, the possible values of J is from 9 to 18.

On the other hand, the output layer presents a pattern of the pulse types in PD recognition. As PD pulse needs to be discriminated from another two kinds of pulses, the number of nodes in output layer can be considered in two possible cases: 1 and 2. In the first case, only one node is needed. The output "1" represents PD pulse and "0" suggests the noise interferences that include both repetitive pulses and random pulses. In the second case, the output layer contains two nodes. Then, the output "1 1" denotes the PD pulse, "1 0" suggests the repetitive pulse, and "0 0" stands for random pulse.

Therefore, there are in total ten candidate numbers of inputs nodes, 9 to 18, and two possible choices of output sizes. How many nodes in the input and output layers will affect the training and performance of neural network. According to the theory of neural networks, smaller size of network may result in less calculation and lower requirement of memory. The training and testing with small-size network is much easier to be

implemented by hardware. Further, with less computations, the training duration decreases and the response of network increases greatly. When the input layer and output layer can represent the patterns of external environment effectively, the number of nodes should be as few as possible [190]. Therefore, to generate a network that is more effective in PD recognition and less likely to make misjudgment, the neural network with 9 nodes in input layer and single output node is adopted.

2) Nodes of hidden layer

Selection of the number of nodes in hidden layer is very important in the whole network architectures. Although the hidden layer does not interact with the input values directly in feed-forward network and was selected randomly in some research works [195-197], it cannot be ignored that appropriate selection of hidden neurons has a great influence on the final output such that it should be treated carefully.

When selecting the number of neurons in hidden layer, too many or too few numbers are not good. Too few neurons in the hidden layers will result in something called under-fitting. In that case, the few nodes in hidden layer cannot find out the relationship between input and output and the trained network is thus unsatisfied for classification. On the other hand, if the number of neurons in hidden layer is too many, more problems will be caused. First, it may cause over-fitting. The size of inputs may be relatively insufficient when compared with the processing capability of network. Second, even if the size of input pattern is large enough, the time for training increases greatly because of the huge calculation complexity [183]. Therefore, the number of neurons in hidden layer should be neither too many nor too few.

So many methods were proposed to find the rules of determining the correct number of neurons in hidden layer. For example, the number should be less than input targets and greater than output ones [198], or a function of neuron numbers of input and output layers [183]. However, the size of hidden layer selected by using those rules does not always generate optimal results in all applications. A simple and adaptive way is to try the possible numbers one by one and test the trained ANN to determine the optimal number [183]. Since in prior section, the nodes of input layer and output layer are selected to be 9 and 1, the possible candidates for this network are from 2 to 8 which are greater than output targets and smaller than input size. The performance or MSEs of NN with different sizes of hidden layer is shown in Fig.5-9. As the training initial condition of neural network is randomly selected, the MSE value of each network with different size of

hidden layer in Fig.5-9 is the average of 20, 50, and 100 trained networks, respectively. The network training procedure will stop if average MSE of the network is smaller than 10^{-4}.

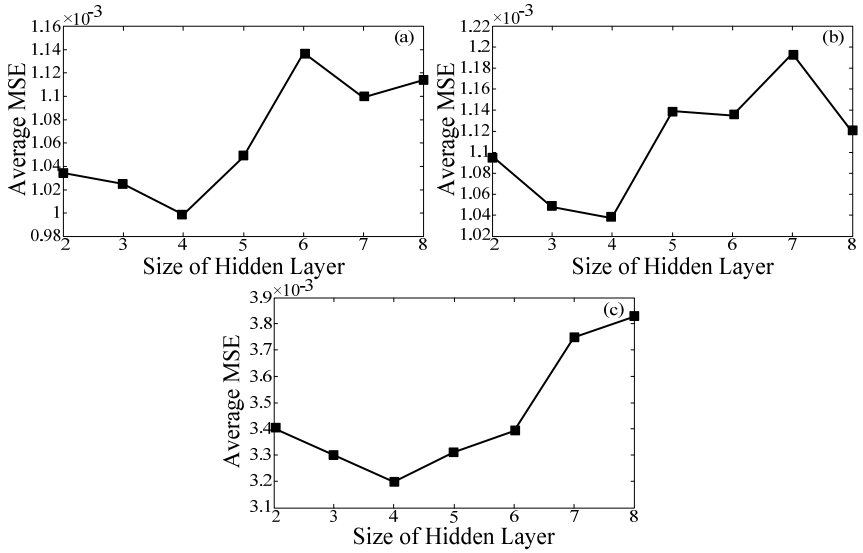

Fig. 5-9 Performance of neural networks with different sizes of hidden layer, (a) average of errors of 20 NNs, (b) average of 50 NNs, (c) average of 100NNs.

As illustrated by Fig.5-9, the MSEs of neural networks with architecture $9-n-1$ are very small if the number of nodes n in hidden layer changes from 2 to 8. As the number of NNs increases, the distribution of average errors becomes more and more regular. However, the distributions always reach their minimum values when the size of hidden layer is 4. Therefore, the number of nodes in hidden layer is set to 4 in this PD recognizing system.

After the discussions on the selections of optimal settings, a feed-forward BP network with a structure of 9-4-1 is employed, and it is trained by LM algorithm with Bayesian regulation. All the data processing procedures are realized on the platform of MATLAB which provides toolboxes of both wavelet analysis and neural network. In addition, the training is terminated when the number of iteration reaches one thousand or the average MSE is smaller than 10^{-7}.

5.4.5 Classification with trained network

Classification and reconstruction of the real PD pulses constitute the last step of PD recognition system. All the pulse-contained segments in the polluted PD signal are analyzed one-by-one. The wavelet entropy features of each pulse segment are classified by the trained neural network which is constructed in Section 5.4.4. If the pulse segment

is judged to include a possible PD pulse, all its wavelet coefficients are kept and reconstructed. Otherwise, the wavelet coefficients are deleted and the pulse segment will not appear in the recovered signal.

5.5 PD recognition results and discussions

In this section, the PD recognitions of some noised PD signals are performed in the proposed system. Here, 6 groups that contain 98 datasets of PD pulse in total are recognized. All of these signals are collected via TEV measurement with non-intrusive sensor. The measured PD signal of one cycle (20 milliseconds) is denoted as one group of PD pulses. Some noisy interference such as repetitive pulse and random pulse is added into the original PD signal. The neural network is constructed and trained as in section 5.4.4. The statistical results of the performance of trained ANN are shown in Table 5.3.

Table 5.3 Recognition results of trained network with test groups

Group No.	Noise type	Concurring pulses	Real PDs	Recognition		Misjudgments	
1	Re.	Yes	16	14	(87.50%)	0	(0%)
2	Ra. + Re.	Yes	18	6	(33.33%)	1	(5.56%)
3	Ra. + Re.	Yes	12	10	(83.33%)	3	(25.00%)
4	Ra. + Re.	Yes	19	15	(78.95%)	2	(10.53%)
5	Ra. + Re.	No	8	7	(87.50%)	3	(37.50%)
6	Ra.	No	25	18	(72.00%)	0	(0%)
Total			98	70	(71.43%)	9	(9.18%)

Among the six groups in Table 5.3, two groups are contaminated by one kind of impulsive noise, either repetitive or random pulses, and the other four are polluted by both kinds of noises. In some groups such as group 1 to 4, a number of impulsive noises occur at the same time with PD pulses. However, in the last two groups, no concurring PD and noise pulses are included. The values in Table 5.3 suggest two conclusions. First, the recognition rates vary greatly when concurrent PD and noise pulses are included: the group 1 where only repetitive noise is added has the highest recognition rate (87.50%), while the group 2 that is polluted by both repetitive and random pulses could only recognize one third of the original PD pulses. Second, the test groups, in which only one type of noises is added, reveal better performance in distinguishing noise pulses. The

group 1 with only repetitive noise and group 6 with only random pulses do not misjudge any noise pulse as PDs.

In order to give a clear understanding of test results, three groups: group 1, group 2 and group 6 are shown in Fig.5-10, Fig.5-11 and Fig.5-12, respectively.

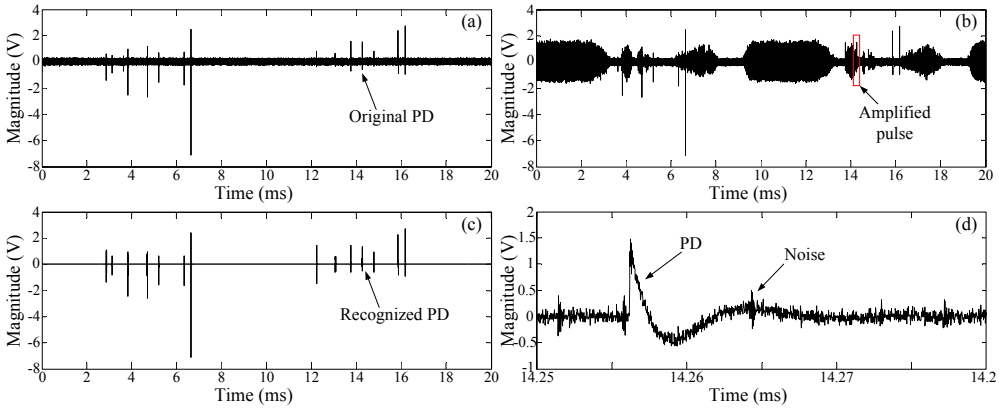

Fig. 5-10 Recognition results of group 1, (a) original PD pulses, (b) noisy signal, (c) recognized PDs, (d) magnified noisy signal segments

Group 1 with the highest recognizing rate and lowest misjudging rate is first analyzed. Since concurring impulsive noises are found in the noised PD signal, the entropy vector which only reveals the features of the pulse segment will distort because of the additional information of impulsive noise. However, the amplitude of original PD pulse of group 1 is much greater than that of noise. This can be demonstrated by the magnified signal segment in Fig.5-10(d). The amplitude of PD pulse peak is about 1.5V, but the maximum peak-to-peak amplitude of noise is only around 0.8V. Although some PD pulses and noise pulses occur at the same time, entropy distribution of pulse segment in Fig.5-10(d) reveal a PD-like pattern rather than a noise-like one. Therefore, the pulse-contained segments are more likely to be recognized as PD pulses if the PD energy is much larger than that of noise.

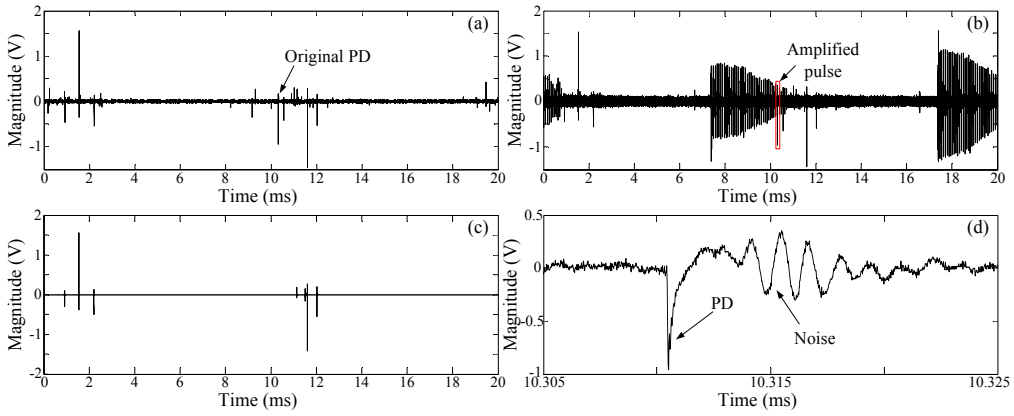

Fig. 5-11 Recognition results of group 2, (a) original PD pulses, (b) noisy signal, (c) recognized PDs, (d) magnified noisy signal segments

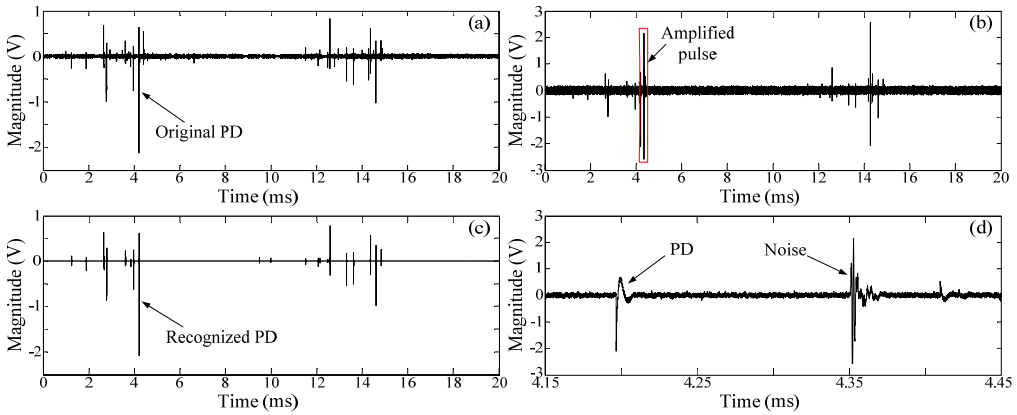

Fig. 5-12 Recognition results of group 6, (a) original PD pulses, (b) noisy signal, (c) recognized PDs, (d) magnified noisy signal segments

However, if the noise amplitude is large enough, the entropy distribution of signal segment will change greatly and be classified as noise. This can be well explained by the test results of group 2. The magnified pulse in Fig.5-11(d) shows the maximum peak-to-peak amplitude (around 0.8V) of noise is close to the maximum amplitude of PD (around 1V). Due to the influence from noise energy, its entropy distribution will be different from that of PD and this pulse cannot be correctly recognized. The concurring noise pulses with similar amplitude with PDs should be the main cause of low recognition rate of group 2.

Besides group 1, the group 6, as shown in Fig.5-12, also has the lowest misjudging rate. From the magnified signal segment, it is easy to find the PD signal and random pulses occur at different times and only one kind of noise is contained. The PD and noise pulses are processed respectively and the noise energy is thus less likely to affect the classification results.

Therefore, the pulse based PD extraction method can generate better performance when PD and noise do not overlap each other.

5.6 Comparisons

To illustrate the effectiveness of proposed PD recognition system, the performance of other methods, for example, non-ANN based method is studied. Although ANN is a good way to recognize PD pulses when their features are known, it is also possible to reject impulsive noises without the help of ANN. The PD pulses and impulsive noises can be classified by using some rules. As aforementioned in 5.2.2, the wavelet entropy distributions of PD, repetitive noise and random noise are different. Comparing the

differences between the wavelet entropy distributions of unknown pulse and the known types is a possible direction to extract PD pulses. Similar with the training procedure of ANN-based method, a data base of the wavelet entropy distributions of three types of pulses is found. The average entropy ratios ρ_{Ek} of twenty pulses from each type are used as the standard distributions. For an unknown pulse, its entropy ratio ρ_{Eu} is first calculated and the differences between ρ_{Eu} and all standard ρ_{Ek} of PD and impulsive noises are then compared. The smallest difference suggests the most similar wavelet entropy distributions. Here, the difference is defined as

$$Dif = \frac{\| \rho_{Eu} - \rho_{Ek} \|}{\| \rho_{Eu} \| \cdot \| \rho_{Ek} \|} \tag{5-12}$$

Since ρ_E stands for the entropy ratio defined in (5-2) and $\| \rho_E \| = 1$, the difference formula in (5-12) can be rewritten to be

$$Dif = \| \rho_{Eu} - \rho_{Ek} \| \tag{5-13}$$

For each unknown pulse, its differences between standard ρ_{Ek} of PDs, repetitive noises and random pulses are denoted by Dif_{PD}, Dif_{re} and Dif_{ra}, respectively. If Dif_{PD} is the smallest among all three differences, the unknown pulse can be classified as a PD.

Table 5.4 Recognition results of test groups by comparing differences *Dif*

Group No.	Real PDs	Recognition		Misjudgments	
1	16	10	(62.50%)	1	(6.25%)
2	18	5	(27.78%)	4	(22.22%)
3	12	6	(50.00%)	9	(75.00%)
4	19	10	(52.63%)	6	(31.58%)
5	8	4	(50.00%)	9	(112.50%)
6	25	11	(44.00%)	0	(0%)
Total	98	46	(46.94%)	29	(29.59%)

Compared with the ANN-based method mentioned before, this difference-based method also bases on the analysis of wavelet entropy distributions. However, it just simply compares the differences between unknown pulse and known features and finds the most similar one. With less calculations and simple judgments, the difference-based method

might not generate as good results as ANN-based method does. To illustrate its performance, the test groups in Table 5.3 are analyzed again. The recognition results are shown in Table 5.4.

As demonstrated in Table 5.4, the PD pulse extraction by comparing the differences can only recover a part of the real PD pulses. The recognition rate is 46.94% which is lower than that of ANN-based method (71.43%) and the misjudgment rate (29.59%) is much higher than the 9.18% of ANN-based method.

5.7 Conclusion

In this chapter, the application of wavelet entropy and neural network in impulsive noise reduction of PD measurement is investigated. Based on the studies of pulse properties and introduction of entropy theory, the wavelet entropy is shown to be effective in characterizing PD and noise pulses and reducing the feature dimension. Furthermore, with careful selection of the parameters and settings, a neural network which is suitable for entropy-based PD recognition is constructed and trained. The test results demonstrate that the proposed wavelet entropy based PD recognition with neural network can recognize most real PD pulses in different cases. However, the recognition rate decreases if concurring noise pulses are included. Further, the training procedure of neural network needs a large number of datasets which is difficult to gather in most field test. Possible solutions of those issues will be discussed and explored in Chapter 6.

CHAPTER 6
TIME-FREQUENCY ENTROPY BASED PD EXTRACTION

6.1 Introduction

When detecting PDs on the external surface of enclosure, one of the major problems that need to be addressed is the interferences from surroundings. As discussed in prior chapters, the non-impulsive noises such as white noise and harmonics can be rejected effectively by wavelet thresholding, and the impulsive interferences could be distinguished by pattern recognition based systems. However, difficulties still exist due to the following challenges:

First of all, it is very hard to extract PD pulses from signals with very low SNR by using wavelet transform. Wavelet analysis could only divide the frequency range into several bands. Such segmentation is not accurate enough to display the slight energy variations when PD's energy is smaller than noises in most frequency bands. For example, the modulated sinusoidal interferences which are discontinuous in time-domain also generate large-amplitude coefficients in fine scales. They are hard to remove if their singularities are larger than thresholds. Also, the wavelet transform cannot discriminate concurrent PD and impulsive noises if their frequency spectrums overlap.

Furthermore, lack of datasets is the major barrier for artificial intelligence (AI) based methods. Artificial intelligence such as ANN has been used to recognize PDs and other pulse-like noises and good performances were reported [37]. A basic premise for these AI based methods is the proper collection of a database with enough size. However, such large database is difficult to collect in most field applications.

This chapter describes an effective PD extracting tool that can extract PD pulse when SNR is very low and without prior knowledge on PD to be extracted. The basic idea of the method is the combination of entropy spectrum and TF analysis. A series of experiments are performed to prove the feasibility and effectiveness of this approach. The organization of this chapter is as follows: First, the fundamentals of TF analysis and details of the application of short-time Fourier transform (STFT) are studied. Then the

generation of entropy spectrum and its feasibility in PD extraction are discussed. Based on analyzing the characteristics of PDs and major noises in TF domain, the TF entropy based algorithm is introduced and its capability of noise reduction is illustrated by some PD signals and interferences.

6.2 STFT based TF analysis

The TF analysis is very useful to discriminate PDs from noisy background, because it allows researchers to observe the frequency spectrum characteristics evolving with time. As the PD signal collected by non-intrusive sensor often has wide frequency bandwidth with its high frequency components being attenuated greatly during propagation, the short-time Fourier transform (STFT) with equal resolution throughout the whole frequency range is capable of revealing the TF distribution of such signal.

6.2.1 Fundamentals of STFT

Short-time Fourier transform (STFT) has often been used to determine the sinusoidal frequency components and phase features of local sections of signal. For any signal f, the resulting STFT is as follows:

$$Sf(u,\xi) = < f, g_{u,\xi} > = \int_{-\infty}^{+\infty} f(t) g(t-u) e^{-i\xi t} dt \tag{6-1}$$

The sliding window $g_{u,\xi}(t) = e^{i\xi t} g(t-u)$ is a real and symmetric window $g(t) = g(-t)$, translated by u and modulated by the frequency ξ. It is normalized $\| g \| = 1$, so that $\| g_{u,\xi} \| = 1$, for any real numbers u and ξ.

Considering discrete signal of period N, the discretization of window and STFT are $g_{m,l}[n] = g[n-m] e^{i2\pi ln/N}$, and

$$Sf[m,l] = < f, g_{m,l} > = \sum_{n=0}^{N-1} f[n] g[n-m] e^{-i2\pi l \quad n/N} . \tag{6-2}$$

When the window $g(t)$ slides along the time axis, the frequency spectrum of the windowed signal is revealed. The spectrum of the whole time range forms a two-dimensional representation of signal which is called time-frequency spectrum [199]. It is denoted by P_s:

$$P_S f(u,\xi) = |Sf(u,\xi)| \tag{6-3}$$

A) Selection of sliding window

The resolution of STFT depends on the spread of the window $g(t)$ in time and frequency, which is measured by its bandwidth $\Delta\omega$ and the maximum amplitude A of the first side-lobes as shown in Fig.6-1 where $\hat{g}(\omega)$ denotes the Fourier transform of $g(t)$.

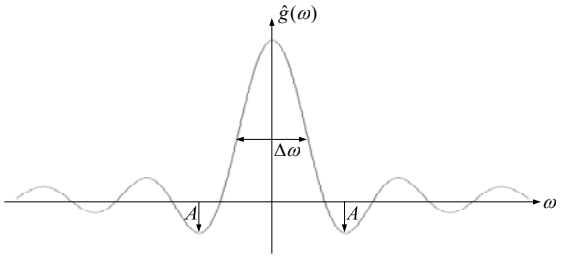

Fig. 6-1 The Fourier transform of window $g(t)$

PD pulse has a quite short duration and wide frequency range. Empirically, a window with larger $\Delta\omega$ and rapid decay is better for TF analysis of PDs. Referring to the window parameters listed in [200], the hanning window is chosen.

B) Size of sliding window

According to Heisenberg uncertainty principle, as the size K of window $g(t)$ increases, the resolution increases in frequency domain, but decreases in time domain [199]. Thus, a trade-off is needed between time and frequency localization. The short duration of PD pulse requires a high resolution in time which suggests a small size K. However, the support of $g(t)$ should at least cover the whole individual PD pulse to reveal all of its frequency components. Furthermore, limited by the frequency response of PD sensor and energy loss during propagation, resolution in frequency should be high enough to differentiate PD from noise. Therefore, the size of window is chosen to be the smallest integer $K = 2M + 1$ such that the longest pulse is covered and M divides N. If the pulse length is very short, a minimum window size of 2μs is used.

For the discrete noisy data F, the resulting STFT is

$$SF[m,k] = < F, g_{m,k} > = \sum_{n=-M}^{M} F[n]g[n-mM]e^{-i2\pi kn/K} \tag{6-4}$$

with $0 \le k \le K$, $0 \le m < N/M$. Here, $M = (K-1)/2$.

6.2.2 TF spectrum of PD

With the selected window in Section 6.2.1, the TF spectrum of PD is studied here. The PD pulse waveform is heavily influenced by insulating materials, PD types, sensor features and physical connections [105]. They often have quite short durations and cover a wide frequency range up to 1GHz. If the frequency response of the sensor is wide enough, an energy strip paralleling to the frequency axis can be found in the TF spectrum as in Fig.6-2. Here, the logarithm of TF spectrum is plotted for clearer visualization.

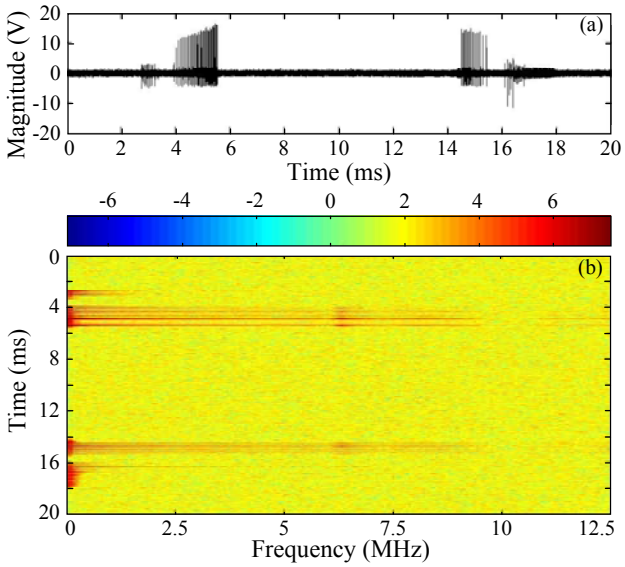

Fig. 6-2 TF spectrum of PD signal, (a) original PD signal, (b) TF spectrum

Compared with the TF spectrums of noises in Section 2.5.1, the TF spectrum of PD is totally different from those of non-impulsive noises: time-axis-paralleling strips of harmonics and squares of modulated sinusoidal interferences. Although the TF spectrums of impulsive noises are also frequency-axis-paralleling strips, there are still some differences such as different frequency range and distributions.

6.2.3 Comparison with wavelet transform

The goal of this paragraph is to compare the TF spectrums of STFT and wavelet transform. The aim of employing STFT is to reveal the energy of signal in equal resolution and explore the slight energy variations in TF domain. This suggests STFT can reveal more information of PD and noises than wavelet analysis.

According to the definition of STFT, a sliding window $g(t)$ moves along the time axis, and the frequency spectrums of signal segments in each window $g(t)$ are revealed to form

the TF spectrum. The atom $g_{u,\xi}(t)$ which is translated by u in time and by ξ in frequency has a spread independent of u and ξ. As shown by Fig.6-3(a), each box $g(t)$ corresponds to a Heisenberg rectangle whose size is independent of its position u,ξ or v,r. The energy of PDs measured by non-intrusive sensor is much higher at lower frequency range while the energy of interferences may appear at any frequency bands. Therefore, equal TF resolution should be much better to reflect the distribution of PD and interferences.

Orthogonal wavelet analysis which is also called multi-resolution analysis is a progress of iterative decomposition of the approximate coefficients by a pair of conjugate mirror filters. Thus, different from STFT, wavelets have a time-frequency resolution that changes. As in Fig.6-3(b), the wavelet $\psi_{j,n}$ has a smaller time support centered at $2^j n$ and wider frequency range centered at $\eta/2^j$, where η is the center of frequency range of ψ. The wavelet $\psi_{j+2,p}$ has a larger time support and higher resolution in frequency. This modification of time and frequency resolution is adapted to represent signals having a component that may vary quickly at high frequencies. Further, the frequency resolution of multi-resolution analysis is lower than that of STFT, only j frequency-indexed vectors are produced by multi-resolution analysis while M vectors by STFT, where j is the number of decomposition scale and $2M+1$ is the size of sliding window $g(t)$. It is more difficult to remove large-magnitude interferences in the frequency bands that contain energy of both PD and noise.

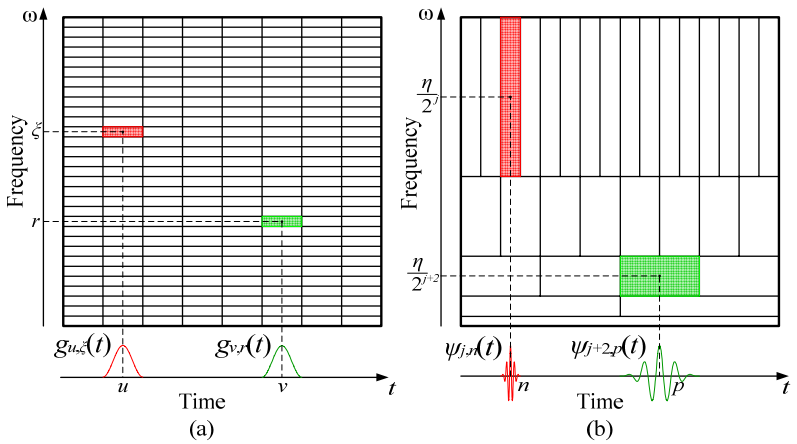

Fig. 6-3 The time-frequency boxes (Heisenberg boxes) of STFT and multi-resolution analysis.

6.3 Study on TF entropy

Although STFT can generate a TF spectrum with equal resolution and provide more information, the influence of power-frequency pick-up, random frequency distribution of

noises, and so on reduces the effectiveness of PD extraction from STFT-generated TF spectrum. To eliminate this influence, TF entropy spectrum is used.

6.3.1 Entropy spectrum generation

TF entropy is actually the entropy spectrum of TF spectrum in (6-3). Each coefficient in the entropy spectrum is the Shannon entropy of a square window that slides along the time and frequency axes of TF spectrum. The definition of Shannon entropy is illustrated in (5-1). The calculation of entropy spectrum includes three steps:

Firstly, as the average magnitude of coefficients varies a lot with frequency, the TF spectrum is normalized. The TF spectrum $SF[k]$ indexed by each frequency k is first subtracted by their minimum value $SF[k]_{\min}$ and then divided by the revised maximum value $SF[k]_{\max} - SF[k]_{\min}$ of this frequency. After normalization, the minimum and maximum value in whole TF spectrum is 0 and 1, respectively.

Secondly, thresholding is used to remove small-amplitude noisy coefficients. In PD spectrum, most noisy coefficients follow Gaussian distribution whose variance was proved to approximate $\varepsilon \approx M_X / 0.6745$. To select as much frequency bands as possible, a threshold with smaller estimation risk is needed. The minimax estimation which has been proved to have a smaller estimation risk and give good predictive performance is used. The minimax threshold is introduced in Section 4.3.2 and its equation is shown in (4-13). The frequency bands with coefficients larger than threshold are regarded to contain large energy.

Finally, a square sliding window with odd dimensions moves along row and columns of TF spectrum to form an entropy spectrum. All the coefficients in this sliding window form an X and produce one entropy value $H(X)$. To keep the TF and entropy spectrum in the same dimension, all four sides and corners of the original TF spectrum are extended. As illustrated by Fig.6-4, when a $(2n+1)\times(2n+1)$ sliding window is employed, the $M \times N$ TF spectrum can only produce an $(M-2n)\times(N-2n)$ entropy spectrum that only represents part of TF spectrum. After extension, the entropy spectrum in Fig.6-4(b) produced by extended TF spectrum has a size of $M \times N$. Here, the boundaries of original TF spectrum $D1$, $D2$, $D3$, and $D4$ are copied to the extended part $E2$, $E1$, $E3$, and $E4$, respectively. The square grid denotes the sliding window. It moves along the rows (frequency, *F*) and columns (time, *T*) of extended TF matrix, one row or column a time.

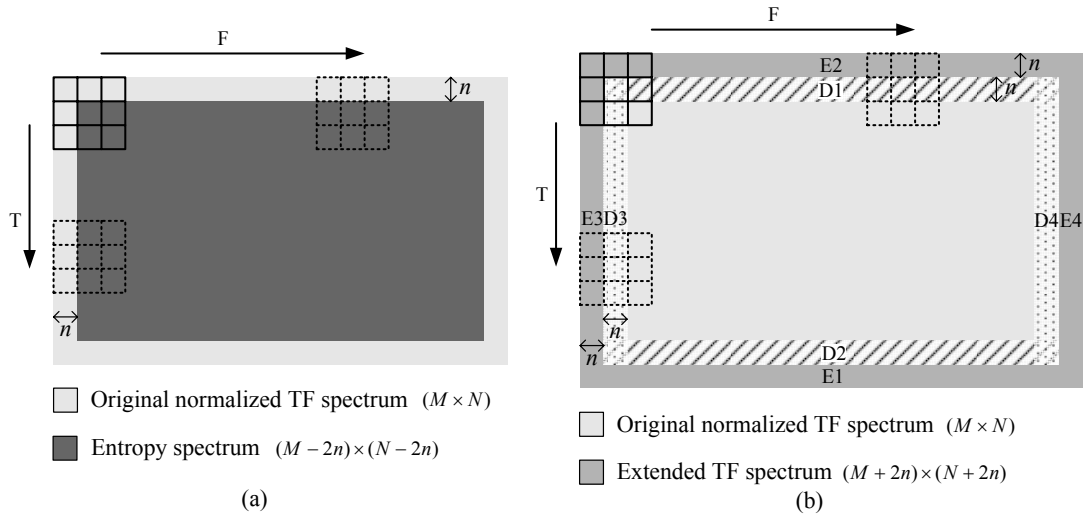

Fig. 6-4 TF spectrum extension and entropy spectrum generation, (a) generating entropy spectrum without extension, (b) generating entropy spectrum with extension

6.3.2 Comparison with TF spectrum

Since entropy only reflects the variation of disorder, entropy spectrum is a robust way to characterize the TF distribution of a signal. This is demonstrated by comparisons with another two simpler normalizations: magnitude normalization and smoothing method. The magnitude-normalized spectrum is actually the normalized TF spectrum before entropy calculation, and smoothed spectrum is the filtered TF spectrum with average filter

The performances of the three types of normalizations are studied on a laboratory-generated noisy PD data. The PD signal and their TF spectrums are shown in Fig.6-5. The values in Fig.6-5(b) equal $20\log_{10} P_s$ and are denoted in dB.

The values in magnitude-normalized and smoothed spectrums rely on the magnitude of TF coefficients. If the magnitude differences between PDs and noises are large enough, obvious components are found in these two spectrums. Otherwise, it is very difficult to discriminate PDs from noises. As shown in Fig.6-5(b), the PD signal detected by non-intrusive sensor has a frequency range up to about 10MHz. The PD energy with frequency below 5MHz has larger amplitude than noises and is obvious in TF spectrum. Both magnitude-normalized and smoothed spectrums can reveal these energy components as in Fig.6-5(c) and (d). However, the PD energies in the black cycles, E1 and E2, are vague because of the small PD amplitude at frequencies above 5MHz. Thus, only E1 is visible in magnitude-normalized spectrum, and both E1 and E2 are lost in smoothed spectrum.

On the other hand, entropy spectrum reveals the disorder of a signal and its values are unaffected by the magnitude of TF coefficients. Rather than the obvious PD energy in the frequency range below 5MHz, the PD energy with small amplitude at higher frequencies in Fig.6-5(b) (the part E1 and E2 in black circles) can also be found in entropy spectrum.

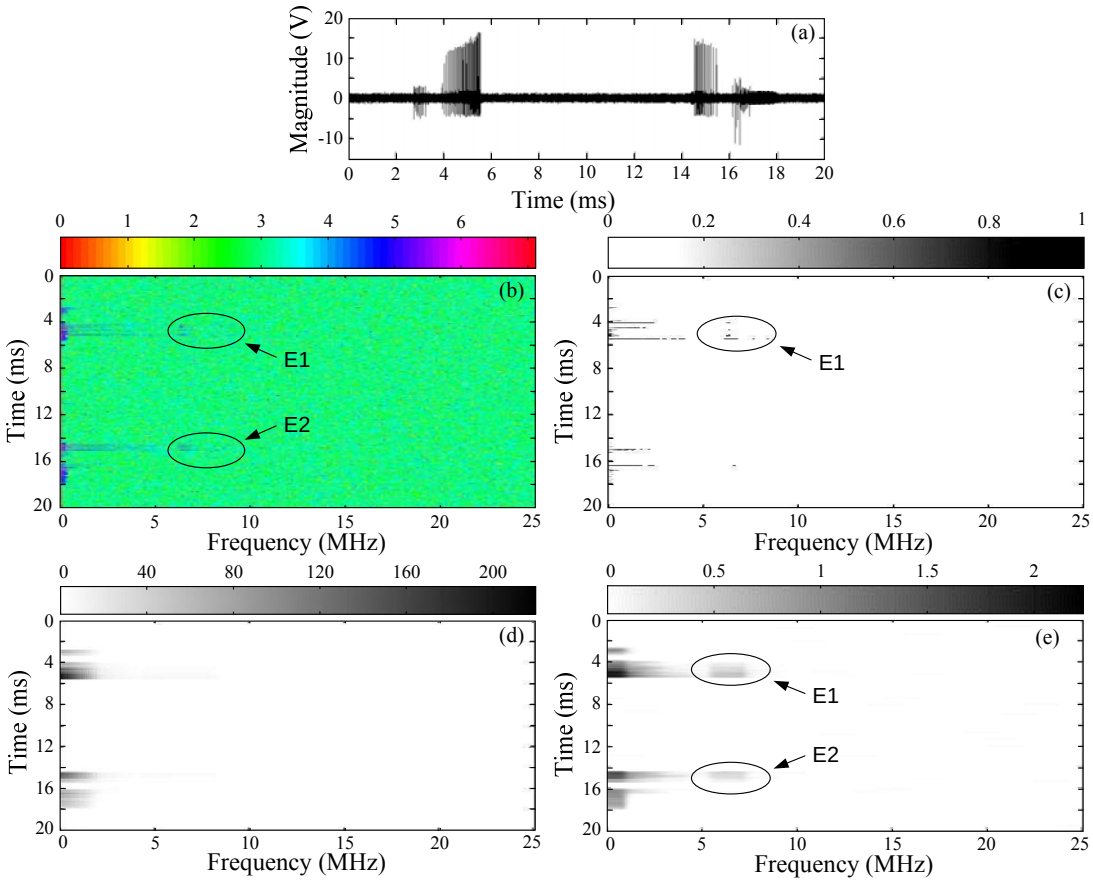

Fig. 6-5 The spectrums produced by different methods, (a) noisy PD signal, (b) TF spectrogram in dB, (c) magnitude-normalized spectrum, (d) smoothed spectrum, (e) entropy spectrum.

6.4 PD extraction algorithm

The entropy and TF analysis described earlier suggest the following PD extraction procedure as shown in Fig.6-6. Five main parts are included in this algorithm:

1) Filter the repetitive pulses with Fourier transform. The repetitive pulses often have large amplitude that sometimes is greater than that of PDs. They are difficult to remove in time-domain if the noise and PD occur at the same time. However, the repetitive pulses from the same source have same features such as frequency distribution which produce many large-amplitude coefficients in frequency domain. Thus, it is possible to remove repetitive pulses by filtering the Fourier coefficients.

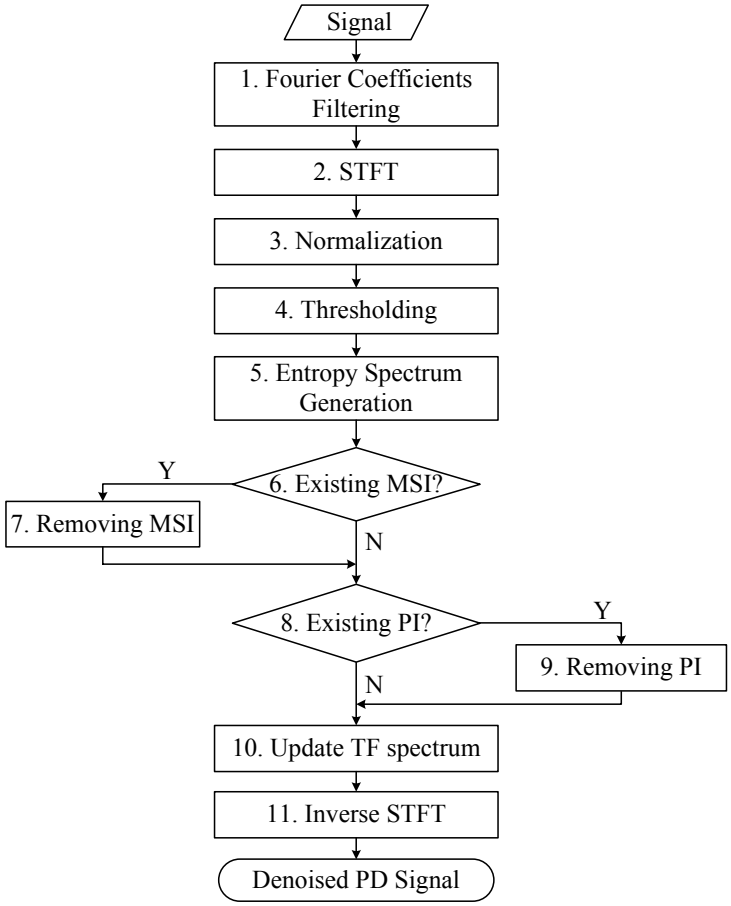

Fig. 6-6 Flow chart of entropy based PD extraction method

2) Perform the TF transform of signal and generate entropy spectrum. First, the TF spectrum is generated via STFT and normalization. As the TF spectrum $SF[k]$ of harmonics with oscillating frequency k is actually a combination of white noise and a positive shift, the distributions of harmonics and white noise are the same after normalization and the influence from harmonics can thus be eliminated after thresholding. Next, threshold in (4-13) is adopted to remove smaller coefficients. Therefore, only the singular points with larger amplitude are shown in entropy spectrum. The noises such as white noise and harmonics cannot be displayed in TF entropy spectrum.

3) Remove modulated sinusoidal interferences (MSI). The sinusoidal noise includes harmonics and MSIs. The harmonics are easily rejected by TF normalization and thresholding in part 2). The MSIs could be rejected by image processing because of their totally different TF spectrums with PDs.

4) Remove random pulse-like interferences. Unlike repetitive pulses, random pulses could not be filtered via Fourier transform. Although these interferences have similar waveform with PDs, their TF spectrums are different in some ways. They are possible to be removed via entropy based TF analysis.

5) Update the TF spectrum according to entropy spectrum and perform inverse STFT to get the recovered PD signal. After removing interferences, the entropy spectrum contains only PD components. The TF spectrum produced by step 4 (thresholding) in Fig.6-6 is compared with entropy spectrum. The coefficients removed in entropy spectrum are also deleted in TF spectrum. Finally, inverse STFT is performed with the updated TF spectrum which includes only PD energy.

The details of interference rejections are discussed in following paragraphs. The de-noising capabilities are illustrated by their performances on some PD and noise signals. All of those signals are collected in the laboratory and field tests via non-intrusive sensor and sampled by 100MSamples/s.

6.4.1 Rejection of repetitive pulses

Since repetitive pulses from the same source have same characteristics (especially waveform and frequency distribution), highly-repeated occurrences of such pulses will generate large-amplitude singularities at their energy peaks which are the same in the frequency domain. Fig.6-7 gives an example of this phenomenon. This repetitive pulse signal is from a high-frequency PFC (power factor correction) convertor in laboratory. As the real and imaginary Fourier coefficients are similar, only real coefficients are displayed. In Fig.6-7(d), most energy of single noise pulse is in the frequency band from 15MHz to 20MHz. This is consistent with the energy distribution of pulse group of one cycle. Meanwhile, the amplitude of Fourier coefficients decreases greatly if the frequency does not correspond to the peak frequencies.

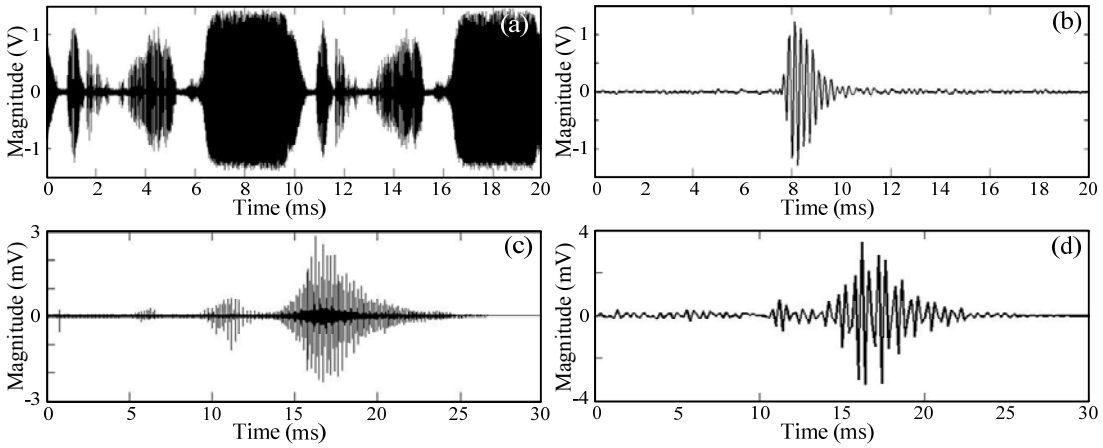

Fig. 6-7 The Fourier coefficients of periodic pulse-like noise, (a) noise of one cycle, (b) single noise pulse,(c) the Fourier coefficients of signal in Fig.6-7(a), (d) the Fourier coefficients of pulse in Fig.6-7(b).

According to this characteristic, the reduction method of repetitive noises is proposed as follow: First, the Fourier coefficients of the noisy data are produced. Next, an empirical threshold 8σ is applied and scans along the frequency axis to detect the singular points. To keep the smooth parts, the frequency axis is divided into many small segments with a bandwidth of 0.5MHz. Here, σ is the estimation of white noise which equals the median of absolute Fourier coefficients of each segment. To reduce the large-amplitude coefficients nearby singularities thoroughly, an inverted-triangular-shaped filter [1, 1-1/n, …, 1/n, 0, 1/n, …, 1-1/n, 1] is employed, where 2n+1 is the width of filter.

The performance of this algorithm is illustrated by a combined data that contains a field-collected PPI and a laboratory-generated PD signal. The small magnitude PPI data is magnified before adding to the PD signal. As in Fig.6-8(b), most energy of the field-collected PPI concentrate around 1MHz. Apparently the large-amplitude coefficients in original data (gray coefficients in Fig.6-8(b)) are removed. Since this filtering method processes signal in frequency domain only, it can effectively distinguish concurrent PD and repetitive pulse. This is demonstrated by three magnified pulses that occur at the same time as in Fig.6-8(e). The combined noisy pulse is separated into a PD pulse and a repetitive pulse successfully.

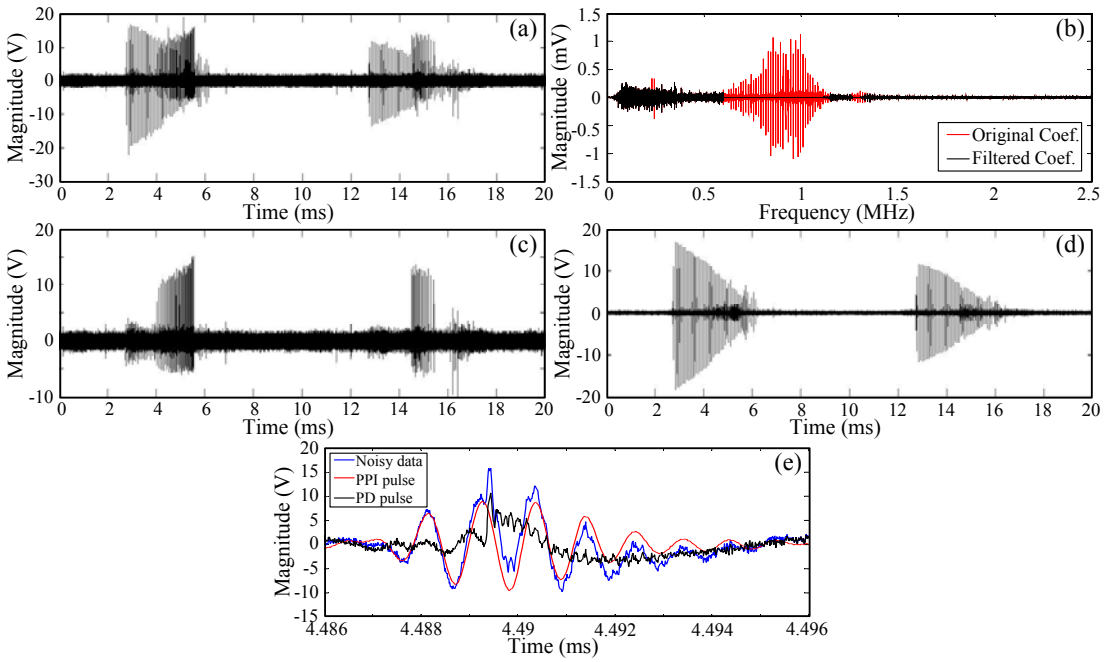

Fig. 6-8 Rejection of periodic pulse-like interferences, (a) noisy data, (b) the Fourier coefficients before and after filtering, (c) recovered PD signal, (d) recovered PPI pulses, (e) magnified single pulse.

6.4.2 Rejection of modulated sinusoidal interferences

The MSIs that vary from time to time often cannot be eliminated by normalization and thresholding. However, the MSI with short duration and narrow frequency range usually concentrates in some small zones in TF plain and appears to be some small squares with sharp edges in TF entropy spectrum. These squares are very easy to detect if the entropy spectrum is regarded as an image. An edge-detecting filter $h=[-0.5 \quad 0 \quad 0.5]$ moves along frequency axis to find the sharp edges of sinusoidal noises first, and then the transposed filter $h'=[-0.5 \quad 0 \quad 0.5]'$ is applied along time axis to find the time range of each 'square'.

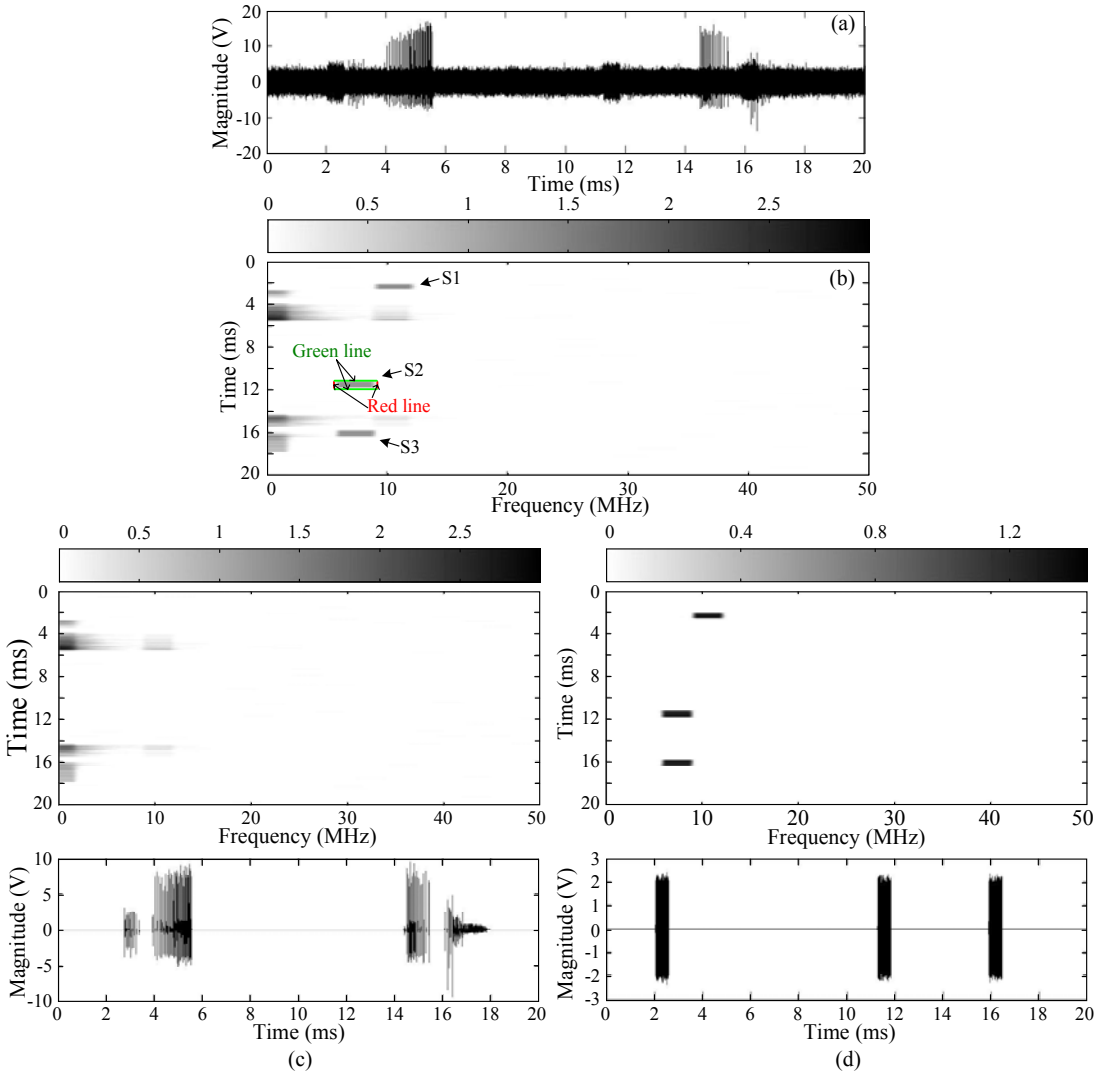

Fig. 6-9 Rejection of sinusoidal noise, (a) polluted PD signal, (b) TF entropy spectrum, (c) entropy spectrum of PD and de-noised PDs, (d) entropy spectrum of sinusoidal interferences and recovered interferences.

A successful PD recovery case is used to show this procedure. The noisy PD data with MSIs and its entropy spectrum are shown in Fig.6-9(a) and (b). The discrete-frequency entropy spectrums of MSIs (squares, S1, S2 and S3) are obviously different from PDs.

The edge-detecting filter firstly moves along frequency axis to locate those large singularities, for instance, the red lines (vertical lines) near S2 in Fig.6-9(b). Then, the transposed filter is applied along time axis. All the sharp edges that are nearest to the red lines are kept, for instance, the green lines (horizontal lines) in Fig.6-9(b). The red lines and green lines envelope all the entropy components of MSI and divide the original entropy spectrum into two: one is PD-related and the other is interference-related. The TF spectrum is updated by referring both entropy spectrums in Fig.6-9(c) and (d). The recovered PD signal and sinusoidal interferences are portrayed too.

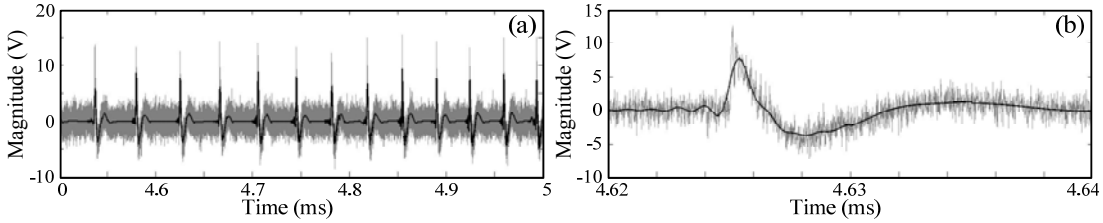

Fig. 6-10 The magnified noisy and de-noised PD signal, (a) PD series last 0.5ms, (b) single PD pulse. Gray line: noisy PD signal, black line: de-noised signal.

Fig.6-10 shows the magnified PD pulses of signal in Fig.6-9(c). The de-noised PD pulses have high similarity with original PD signal. This is quite helpful in further analysis such as condition evaluation or PD type recognition.

6.4.3 Rejection of random pulses

After rejections of repetitive pulses and MSI, random pulse interferences are still found in TF entropy spectrum. Such pulses cannot produce Fourier coefficients large enough to be rejected via filtering Fourier coefficients. Furthermore the random pulses and PDs have similar properties. For example, their TF energy both change gradually with frequency such that sharp edges are rarely found in their entropy spectrums and the edge-detecting filter thus cannot generate satisfactory results. However, there are also some differences between PDs and random pulses, for example, the frequency range and the energy distributions. Some random pulses have similar frequency distribution with PDs, for example, the random pulses appear near the PD source. This case is not considered here. Only the random pulses with great attenuations during propagation are discussed in this research.

Considering those properties, the random pulse-removing algorithm is designed to include four steps:

1) Use frequency-indexed variance to find the PD frequency range.

Since in TF entropy spectrum, large amplitude coefficients suggest the existence of signals such as pulses, if a frequency band contains energy of pulses, it must have greater variance, and vice versa. Therefore, the variances of all frequencies are first calculated and then filtered by the threshold in (4-13) to select the frequency bands that contain pulses. Because the frequency response of non-intrusive sensor is very small in the frequency range beyond 30MHz where white noise dominates, the variance of frequency bands higher than 30MHz is used as the estimation of white noise in threshold calculation. As the PD energy concentrates at lower frequency bands, the pulses with large energy at high frequency bands are regarded to be noise-related such as the pulses with energy in the frequency range larger than 9MHz in Fig.6-11(c).

2) Use time-indexed entropy to describe the energy distribution of each pulse.

As in STFT the size of sliding window is selected to be the width of longest pulse in noisy signal, each time-indexed vector in entropy spectrum includes all the energies of one single pulse. Therefore, the energy distribution of each pulse can be described by entropy of all time-indexed vectors in PD-related entropy spectrum. The pulses of the same type should have similar energy distribution and thus similar entropy. The PD-contained vector which includes more large-amplitude coefficients should have greater time-indexed entropy.

3) Select the possible PDs by considering the location of largest entropy.

In order to discriminate the pulses with similar time-indexed entropy but different energy distributions, further analysis is needed. Considering the energy loss during propagation and the frequency response of non-intrusive sensor, the largest PD energy concentrates at lower frequency bands. However, the largest energy of pulse-like noises which usually have oscillating components should be in higher frequency band than that of PDs. Therefore, the time-indexed entropy generated in step 2) is divided by the relative location of largest entropy which equals the frequency with largest entropy divided by the maximum frequency of entropy spectrum. For the PD pulses and interferences with similar time-indexed entropy, the revised entropy of interference should be smaller because of its larger relative location of largest-entropy.

4) Select the PD pulses with larger revised time-indexed entropy by thresholding.

Due to the large amount of zero-values in time-indexed entropy, median is no longer suitable for noise estimation in this case. Here, the variance based on experimental experience is used. The vectors with entropy greater than variance are used to index PD pulses.

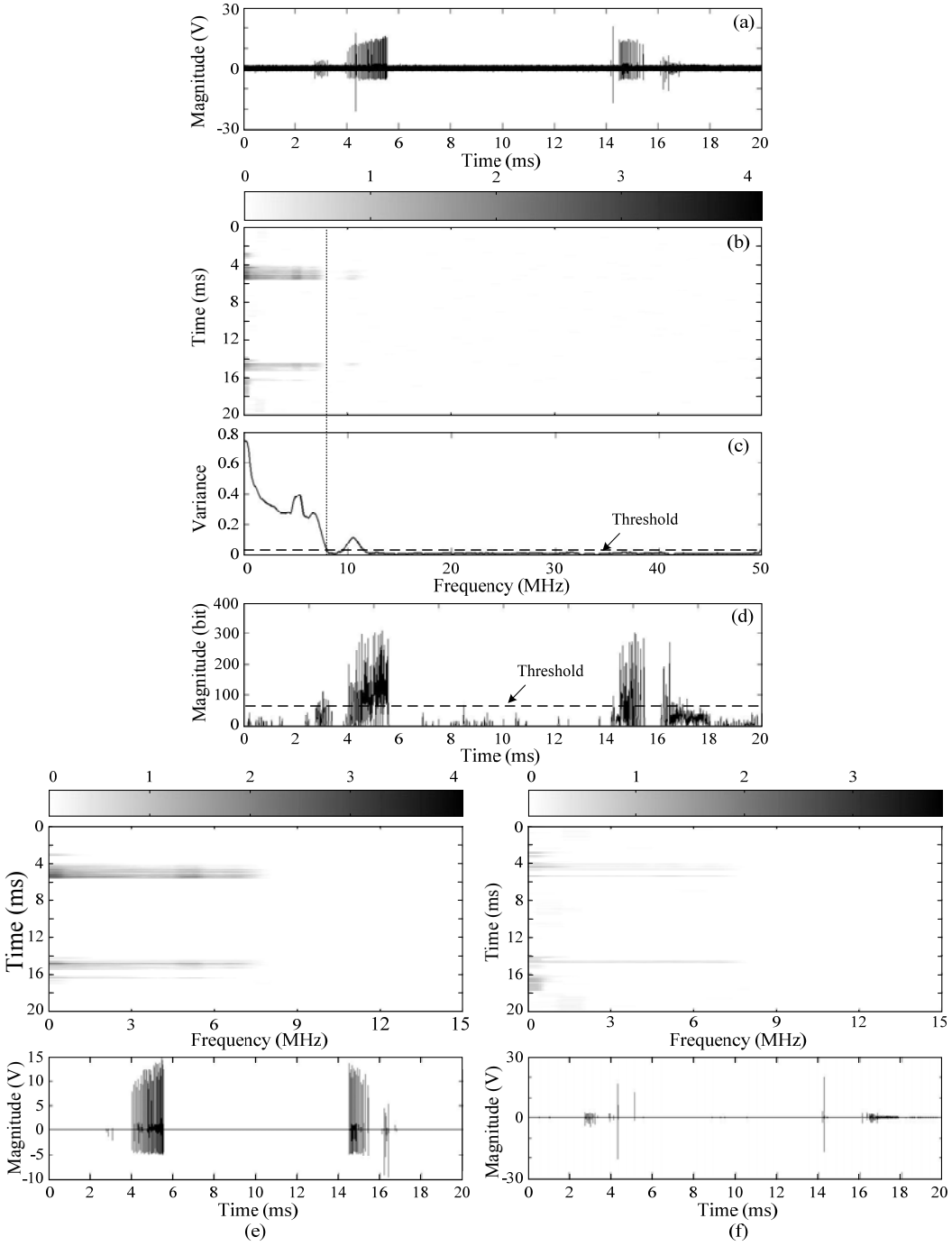

Fig. 6-11 Rejection of pulse-like noise, (a) noisy data, (b) TF entropy spectrum, (c) entropy variance of each frequency, (d) revised time-indexed entropy, (e) entropy spectrum of PDs and de-noised PD, (f) entropy spectrum of noise and recovered pulse-like interference.

Fig.6-11 shows a PD recovery case with random pulses. As shown in Fig.6-11(b) and (c), the PD-contained spectrum is chosen by thresholding the frequency-indexed variance of

entropy spectrum. Fig.6-11(d) portrays the revised time-indexed entropy of PD-contained spectrum. The results in Fig.6-11(e) and (f) show that this entropy-based method can discriminate PD pulses from pulse-like noises.

6.5 Applications

A number of PD signals and noises are collected in both laboratory and field tests. From those datasets, three different cases are considered here to demonstrate the performance of the proposed PD extraction algorithm. All the data used in this section are detected by non-intrusive sensor and sampled at the rate of 50MHz per second. Case 1 and case 2 are the combined noisy PD signals which contain PD pulses, white noise, repetitive pulses, modulated sinusoidal interferences and random pulses. The signal used in case 3 is a field collected signal on an operating apparatus.

Case 1:

In case 1, the original PD signal is generated by a cavity discharge sample, the repetitive and random pulses are collected in a field test of switchgear, and the modulated sinusoidal interference from cell phone is measured in laboratory. All those signals are added to form the noised PD signal. Fig.6-12 shows the original, polluted and recovered signals of a semi-simulated noisy PD signal. After filtering the large-amplitude Fourier coefficients, removing the sharp-edge areas in entropy spectrum and thresholding the small-amplitude revised time-indexed entropy, the repetitive pulses which are obvious in Fig.6-12(b), the modulated sinusoidal interference as in Fig.6-12(f) and the random pulses as in Fig.6-12(i) are separated and the real PDs are recovered. Compared with the original PD signal, almost all PD pulses are extracted. However, due to the revisions of TF spectrum which is updated with the entropy spectrum, some PD energy is lost after inverse STFT. The magnitudes of some recovered PD pulses decreases greatly as in Fig.6-12(j).

Case 2:

To further investigate the performance of proposed algorithm, another PD extraction case of synthesized noisy PD signal is studied. Similar with conditions in case 1, the original PDs are also cavity discharges, and the modulated sinusoidal interference is a laboratory-detected communication signal from cell phone. However, the impulsive noise is from a high-frequency PFC convertor. As demonstrated in Fig.6-13(c), the large-amplitude

singularities in frequency domain concentrate in the range from 10MHz to 20MHz. There still exists some energy of repetitive pulses after filtering Fourier coefficients. However, that residual energy which can be clearly found in entropy spectrum in Fig.6-13(e) could not affect the results of random pulse rejection, because the noise-related frequency bands are removed before calculating time-indexed entropy. Finally, almost all PD pulses in Fig.6-13(a) are extracted in the recovered signal in Fig.6-13(j).

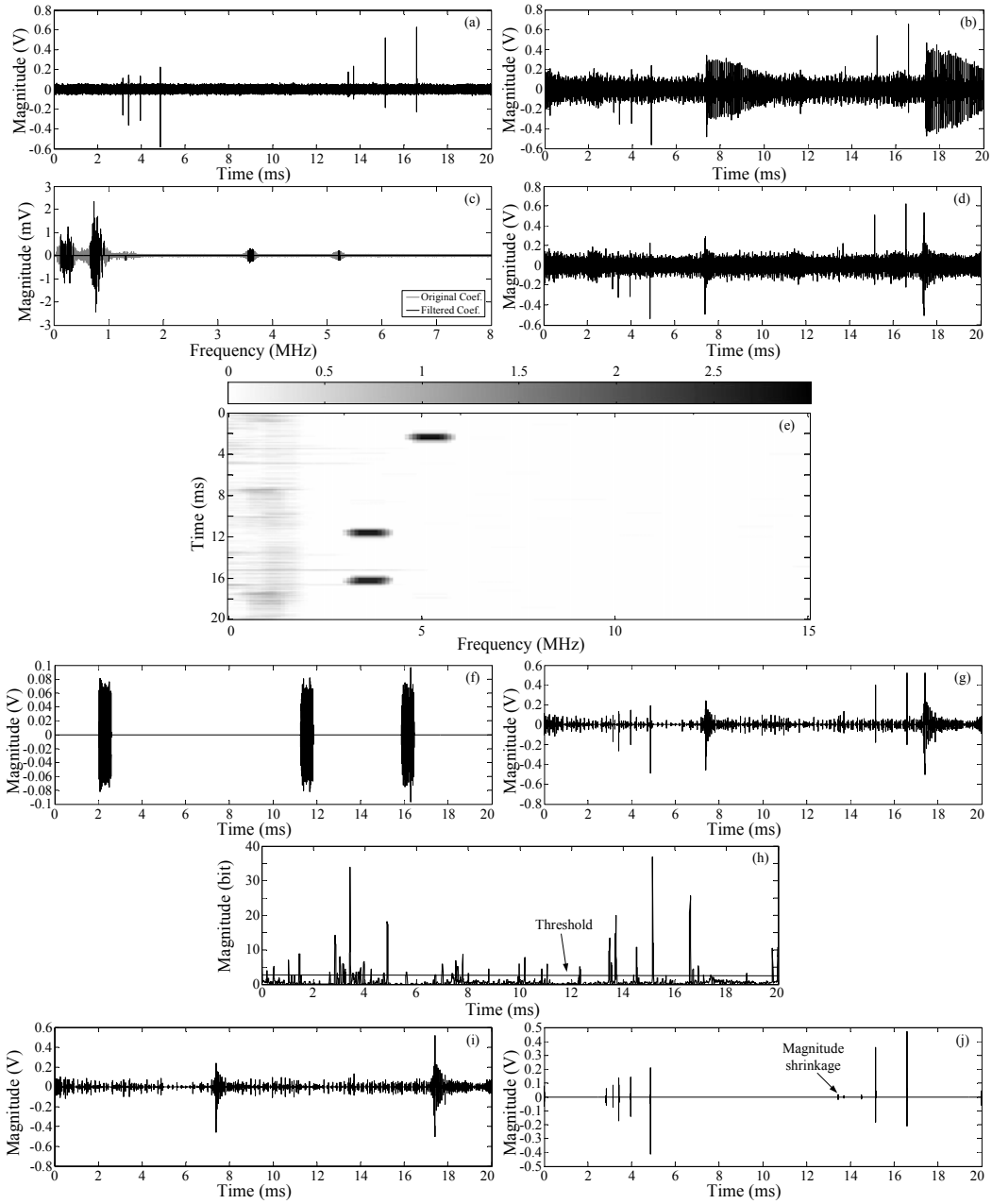

Fig. 6-12 PD extraction from noisy background (case one), (a) original PD signal, (b) noised signal, (c) real Fourier coefficients before and after filtering, (d) noised PD signal after removing repetitive pulses, (e) entropy spectrum of signal in (d), (f) recovered modulated sinusoidal interferences, (g) noisy PD signal after removing modulated sinusoidal interferences, (h) time-indexed entropy and threshold, (i) recovered random pulses, (j) recovered PD pulses

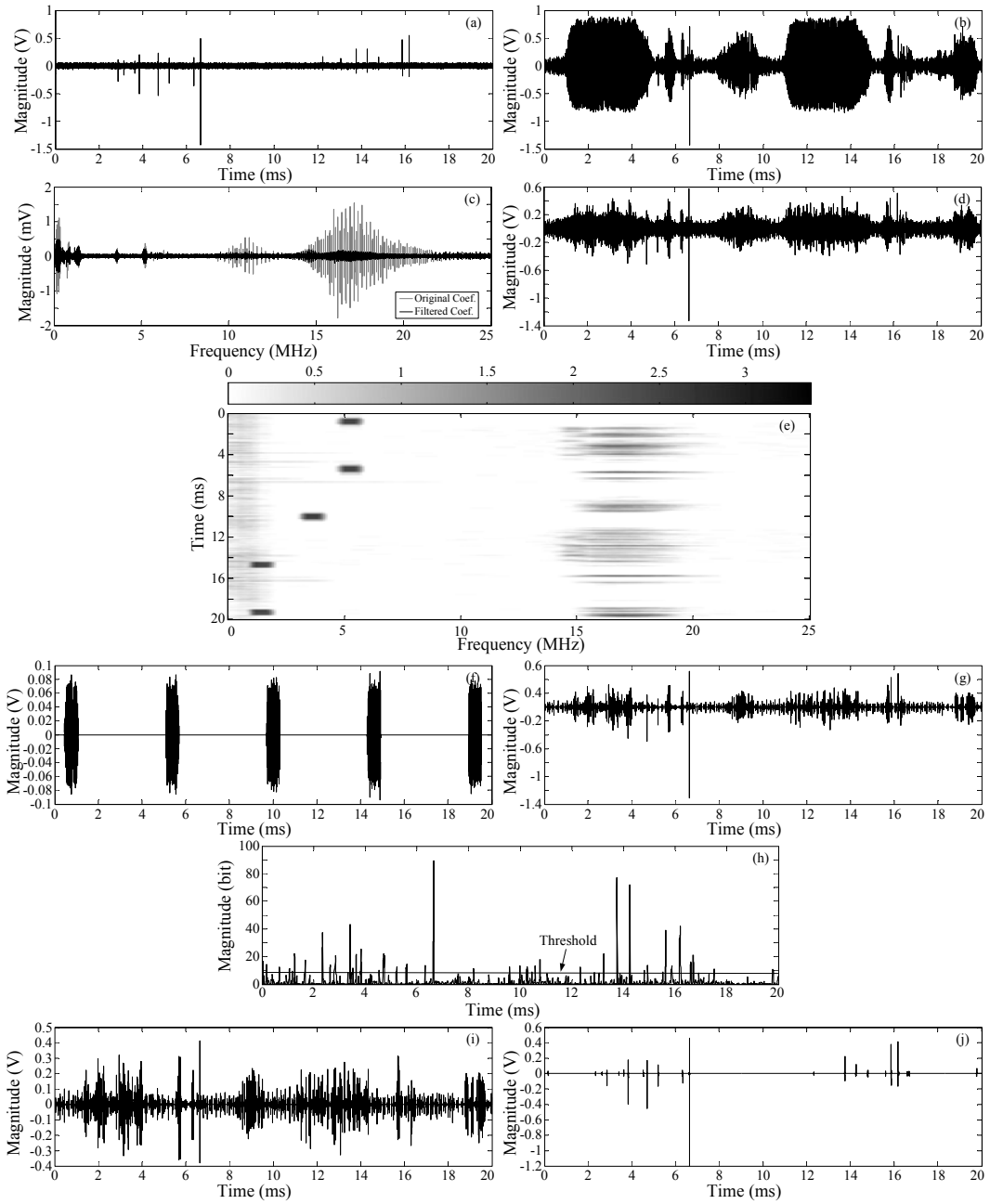

Fig. 6-13 PD extraction from noisy background (case two), (a) original PD signal, (b) noised signal, (c) real Fourier coefficients before and after filtering, (d) noised PD signal after removing repetitive pulses, (e) entropy spectrum of signal in (d), (f) recovered modulated sinusoidal interferences, (g) noisy PD signal after removing modulated sinusoidal interferences, (h) time-indexed entropy and threshold, (i) recovered random pulses, (j) recovered PD pulses

Case 3:

The proposed entropy based time-frequency analysis is also applied to analyze signals collected from field tests. An on-site measurement was carried out in Singapore Shaw Tower. The sensor was placed on the external surface of the metallic enclosure of an oil-insulated transformer. The original signal, the entropy spectrum and de-noised results are displayed in Fig.6-14. In Fig.6-14(a), the impulsive data was collected by the non-intrusive sensor and the sampling rate is 50MSamples/s, and the sinusoidal data is

collected at the output of transformer. They are drawn in the same figure with the help of simulation tool. Some pulses are extracted by using the proposed method as shown in Fig.6-14(c). It would be much better to compare the extracted pulses with traditional methods. More practical data is needed to illustrate the effectiveness of proposed method. However, with the studies in previous cases, the extracted signals in Fig.6-14(c) should be correctly recognized with high confidence.

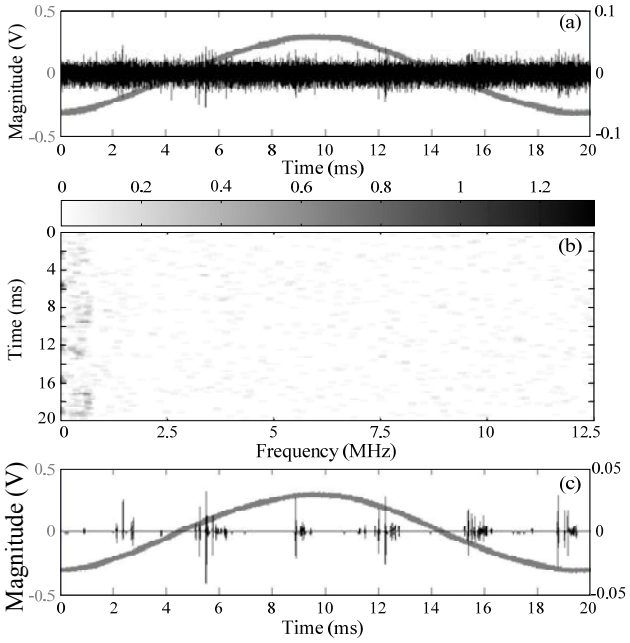

Fig. 6-14 The extracted PDs for field test, (a) original measured data, (b) entropy spectrum, (c) de-noised result.

6.6 Conclusion

Motivated by the PD extraction problem of non-intrusive PD measurement, we presented entropy based time-frequency analysis as an efficient tool to identify real PD pulses in the presence of high levels of noises. The efficiency and adaptability of this proposed method are demonstrated by analyzing PD data as well as some noise data. This thesis shows that entropy based time-frequency analysis can successfully supersede other existing PD de-noising methods in solving the problem of PD pulse extraction. It has several advantages over existing methods. For example, artificial intelligence is not required: the PD pulses are extracted by studying their time-frequency characteristics directly; the entropy based time-frequency spectrum is better than other commonly used methods to represent the features of different pulses; and this method can easily reject noises which overlap the PD pulses and still keep the original waveform of PDs.

CHAPTER 7
CONCLUSIONS AND RECOMMENDATIONS

7.1 Conclusions

This thesis has documented the investigation of TEV based PD measurement in terms of hardware design and noise reduction. The research includes a review of fundamentals of PD phenomenon, popular PD measurement methods, and the research achievements relating to noise reduction; design of a TEV based PD measurement system with non-intrusive sensors and high-pass filter; selection of the optimal settings of wavelet thresholding for TEV signals and enhancement of the processing efficiency of thresholding algorithm; rejection of impulsive noise by using wavelet entropy and ANN when the noise and PD pulses does not overlap; and development of a noise rejection system which could remove both non-impulsive and impulsive noises even if they occur at the same time.

An increasing demand for non-intrusive and precise PD measurement emerges in recent years. To design a reliable system, the fundamentals of PD phenomenon, popular measurement methods, and existing noise rejections have been reviewed. The analysis of PD mechanism and characteristics provides theoretical bases for system design. The investigations of popular PD measurement methods pointed out the merits and drawbacks of coupling capacitor method and UHF method, and proved the potential of employing non-intrusive TEV measurement. The study of noises commonly encountered in PD measurement and the existing de-noising methods gives a clear understanding of noise and is helpful for the development of measuring system and noise reduction methods.

With investigations of the requirements of effective PD measurement, a non-intrusive TEV measuring system is proposed. This system includes a non-intrusive sensor and a high-pass filter. The sensor has a wide frequency response range to capture most of the energy of PDs and the cut-off frequency of high-pass filter ensures their impulsive waveforms. The experimental results show that the proposed system is sensitive enough to capture all PD pulses occurring inside the enclosure. Further, according to the features

of TEV signals and measuring system, the TEV signal has been simulated. By comparing simulated and experimental PD pulses, it is clear the TEV signal can be accurately simulated and the simulations provide important information for further theoretical analysis of TEV.

Wavelet thresholding is the most popular method in removing non-impulsive noises. The three most commonly used thresholds: universal threshold, minimax threshold and SURE threshold, and two thresholding functions: hard and soft thresholding are studied and analyzed. The wavelets from different families are also selected. The de-noised results of simulations point out the universal threshold with hard thresholding function and the wavelets with 4 to 6 vanishing moments are more appropriate than other combinations. The processing efficiency of thresholding algorithm with optimal settings is enhanced by using parallelism. The processing durations of different data show that the parallelism method can greatly speed up the de-noising procedure.

Entropy is a measure of disorder and its value is independent of the magnitude of factors. The application of wavelet entropy in rejecting impulsive noise has been proven to be effective. The comparisons between wavelet entropy distribution and energy distributions demonstrate the advantages of wavelet entropy in charactering PD and noise pulses and reducing the feature dimension. With careful selection of parameters, an ANN suitable for wavelet entropy was constructed and trained. The test results demonstrated that the proposed wavelet entropy based PD recognition with neural network can recognize most real PD pulses in different cases.

With the good performance of wavelet entropy, a time-frequency entropy based noise reduction method was proposed to reject both impulsive and non-impulsive noises even when they overlap each other. The noises are rejected according to their types: non-impulsive noises, repetitive noise, modulated sinusoidal noises and random impulsive noises. The non-impulsive noises are removed by thresholding, the repetitive noises are removed by Fourier transform, and the other two kinds of noises are rejected by time-frequency entropy. The de-noised results of two experimental signals and one field-collected signal show that the proposed noise reduction system is effective in rejecting noises of TEV signals and has potential in practical applications.

7.2 Recommendations

Based on the contributions of the completed project, further researches are recommended in following areas.

1. Enhancement of non-intrusive TEV sensor.

The non-intrusive sensor is the most important part of the PD measurement system with TEV method. In the proposed system, the non-intrusive sensor is either placed on the top of metal box or hand-held when measuring signals on side surfaces. However, such design is difficult to maintain a stable measurement since the sensor may not be fully contacted to the side surface when it is held by human. Therefore, the design of non-intrusive sensor can be improved by adding a clamping device. Consequently, the conductor of non-intrusive sensor can be in contact with the enclosure surface firmly.

Besides the improvements on sensor structure, study on the frequency response is also a good direction. As the lower and upper cut-off frequencies of proposed system are fixed to 100kHz and around 10MHz, respectively. The influences from variations of cut-off frequencies should be studied. Also, more research could be done in exploring new materials or designs which have wider frequency response range.

2. Research on additional kinds of PD other than cavity discharges and needle-to-plane discharges.

In our research only two most commonly encountered PD types: cavity discharges and needle-to-plane discharges were investigated. However, many other kinds of PDs occur in field tests, for example, treeing discharges, surface discharges, discharges caused by floating objects and so on. Their characteristics vary with the insulating materials and sources. The capability of proposed TEV detection and de-noising system should be explored when other kinds of are encountered.

3. Derivation of precise TEV signal of rectangular enclosure.

The theoretical analysis of TEV signal on the external surface of enclosure is crucial for further research. Since the derivation of TEV signal in this thesis neglected losses and during propagation, only an approximate model was developed. However, a precise model is more helpful for simulations of TEV signal.

4. Application on more practical equipment.

Due to the difficulties to access practical equipment with different kinds of noises, the conclusions in this research are mainly based on simulated or combined signals. More practical data should be collected to test the performance of our system, and the applications of our system in different environments should also be studied. Furthermore, the effectiveness of proposed methods needs more analysis when the sampling rate of signals changes.

AUTHOR'S PUBLICATIONS

1. Luo G.M., Zhang D.M., Koh Y.K., Ng K.T., and Leong W.H.. "Time-frequency entropy based partial discharge extraction for non-intrusive measurement", IEEE Trans. Power Delivery, Vol.27, No.4, 2012, 1919-1927.

2. Luo G.M., Zhang D.M., Tseng K.J.. "Recognition of Partial Discharge in TEV Measurements, Using Wavelet Entropy and Neural Network", IEEE Trans. Power Delivery, submitted.

3. Luo G.M., Zhang D.M.. "Recognition of Partial Discharge Using Wavelet Entropy and Neural Network for TEV Measurement", in IEEE PES Int. Conf. Power System Technology, Auckland, POWERCON, 2012, pp: 1-6.

4. Luo G.M., Zhang D.M., Koh Y.K., Ng K.T., and Leong W.H.. "An advanced time-frequency domain method for PD extraction with non-intrusive measurement", in Int. Conf. Power Electronics and Power Engineering, Kuala Lumpur, ICPEPE, 2012, pp: 167-173.

5. Luo G.M., Zhang D.M.. "Entropy application in partial discharge analysis with non-intrusive measurement", in 2nd Int. Congr. Computer Applications and Computational Science, Bali, CACS, 2011, pp: 319-324.

6. Luo G.M., Zhang D.M.. "Efficiency improvement for data-processing of partial discharge signals using parallel computing", in 10th IEEE Int. Conf. Solid Dielectrics, Potsdam, ICSD, 2010, pp: 1-4.

7. Luo G.M., Zhang D.M.. "Study on performance of HFCT and UHF sensors in partial discharge detection", in Conf. Proc. of 9th Int. Power & Energy Conference, Singapore, IPEC, 2010, pp: 630-635.

8. Luo G.M., Zhang D.M.. "Study on performance of developed and industrial HFCT sensors", in 20th Australasian Universities Power Engineering Conf., Christchurch, AUPEC, 2010, pp: 1-5.

9. Luo G.M., Zhang D.M.. “Application of wavelet transform to study partial discharge in XLPE sample”, in Australasian Universities Power Engineering Conf., Adelaide, AUPEC, 2009, pp: 1-6.

10. Luo G.M., Zhang D.M., “Wavelet denoising”, in Wavelet Transform, 1st ed, Dumitru Baleanu, Ed. Croatia: InTech Open, 2012, ISBN 979-953-307-385-8, pp:59-80.

BIBLIOGRAPHY

[1] F. L. Flanders, "Branch Discussion," *Transactions of the American Institute of Electrical Engineers,* vol. XXII, pp. 733-734, 1903.

[2] E. L. Nichols, "Discussion," *Transactions of the American Institute of Electrical Engineers,* vol. XIX, pp. 1063-1073, 1902.

[3] E. Jennings and A. Collinson, "A partial discharge monitor for the measurement of partial discharges in a high voltage plant by the transient earth voltage technique," in *International Conference on Partial Discharge*, 1993, pp. 90-91.

[4] M. Hikita, S. Ohtsuka and S. Matsumoto, "Recent trend of the partial discharge measurement technique using the UHF electromagnetic wave detection method," *IEEE Transactions on Electrical and Electronic Engineering,* vol. 2, pp. 504-509, Sep 2007.

[5] R. Bartnikas, "Partial discharges. Their mechanism, detection and measurement," *IEEE Transactions on Dielectrics and Electrical Insulation* vol. 9, pp. 763-808, 2002.

[6] British Standards Institution., *IEC 60270: High-voltage test techniques - partial discharge measurements*. London: BSI, 2001.

[7] P. H. F. Morshuis and J. J. Smit, "Partial discharges at DC voltage: their mechanism, detection and analysis," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 12, pp. 328-340, 2005.

[8] J. C. Devins, "The 1984 J. B. Whitehead Memorial Lecture the Physics of Partial Discharges in Solid Dielectrics," *IEEE Transactions on Electrical Insulation,* vol. EI-19, pp. 475-495, 1984.

[9] H. Raether, *Electron avalanches and breakdown in gases*. London: Butterworths, 1964.

[10] J. M. Meek, "The vital spark," *Electronics and Power,* vol. 14, pp. 431-433, 1968.

[11] R. Bartnikas and J. P. Novak, "On the character of different forms of partial discharge and their related terminologies," *IEEE Transactions on Electrical Insulation,* vol. 28, pp. 956-968, 1993.

[12] V. Nikonov, R. Bartnikas and M. R. Wertheimer, "The influence of dielectric surface charge distribution upon the partial discharge behavior in short air gaps," *IEEE Transactions on Plasma Science,* vol. 29, pp. 866-874, 2001.

[13] R. Bartnikas, "Some Observations on the Character of Corona Discharges in Short Gap Spaces," *IEEE Transactions on Electrical Insulation,* vol. EI-6, pp. 63-75, 1971.

[14] C. Hudon, R. Bartnikas and M. R. Wertheimer, "Surface conductivity of epoxy specimens subjected to partial discharges," in *Conference Record of the 1990 IEEE International Symposium on Electrical Insulation, 1990*, 1990, pp. 153-155.

[15] C. Hudon, R. Bartnikas and M. R. Wertheimer, "Analysis of degradation products on epoxy surfaces subjected to pulse and glow type discharges," in *Conference on Electrical Insulation and Dielectric Phenomena, CEIDP. 1991 Annual Report.* , 1991, pp. 237-243.

[16] P. Swarbrick, "Characteristics of an arc discharge in sulphur hexafluoride," *Proceedings of the Institution of Electrical Engineers,* vol. 114, pp. 657-660, 1967.

[17] N. Gherardi and F. Massines, "Mechanisms controlling the transition from glow silent discharge to streamer discharge in nitrogen," *IEEE Transactions on Plasma Science,* vol. 29, pp. 536-544, 2001.

[18] T. Ihara, T. Kiyan, S. Katsuki, T. Furusato, M. Hara, and H. Akiyama, "Positive Pulsed Streamer in Supercritical Carbon Dioxide," *IEEE Transactions on Plasma Science,* vol. 39, pp. 2650-2651, 2011.

[19] M. S. Naidu and V. Kamaraju, "Conduction and breakdown in gases," in *High voltage engineering*, 2nd ed New Delhi: Tata McGraw-Hill, 1995, pp. 12-40.

[20] E. Kuffel, W. S. Zaengl and J. Kuffel, "Electrical breakdown in gases, solids and liquids," in *High voltage engineering : fundamentals*, 2nd ed Oxford ; Boston: Butterworth-Heinemann, 1984, pp. 297-419.

[21] M. S. Naidu and V. Kamaraju, "Conduction and breakdown in liquid dielectrics," in *High voltage engineering*, 2nd ed New Delhi: Tata McGraw-Hill, 1995, pp. 41-53.

[22] S. Jayaram and J. D. Cross, "Influence of impurities on electroconvection in insulating liquids," *IEEE Transactions on Electrical Insulation,* vol. 27, pp. 255-270, 1992.

[23] A. Clark, R. J. Dewhurst, P. A. Payne and C. Ellwood, "Degassing a liquid stream using an ultrasonic whistle," in *IEEE Ultrasonics Symposium, 2001*, 2001, pp. 579-582 vol.1.

[24] M. S. Naidu and V. Kamaraju, "Breakdown in solid dielectrics," in *High voltage engineering*, 2nd ed New Delhi: Tata McGraw-Hill, 1995, pp. 54-76.

[25] M.-J. Chen, H.-T. Huang, J.-H. Chen, C.-W. Su, C.-S. Hou, and M.-S. Liang, "Cell-based analytic statistical model with correlated parameters for intrinsic breakdown of ultrathin oxides," *IEEE Electron Device Letters,* vol. 20, pp. 523-525, 1999.

[26] E. Kuffel, W. S. Zaengl and J. Kuffel, *High voltage engineering : fundamentals*, 2nd ed. Oxford ; Boston: Butterworth-Heinemann, 2000.

[27] F. H. Kreuger, *Partial discharge detection in high-voltage equipment*. London ; Boston: Butterworths, 1989.

[28] R. Bruetsch, "High Voltage Insulation Failure Mechanisms," in *Conference Record of the 2008 IEEE International Symposium on Electrical Insulation, ISEI 2008.* , 2008, pp. 162-165.

[29] W. A. Thue, *Electrical power cable engineering*. New York: Marcel Dekker, 1999.

[30] R. Patsch, "Electrical and water treeing: a chairman's view," *IEEE Transactions on Electrical Insulation,* vol. 27, pp. 532-542, 1992.

[31] R. Patsch, "On Tree-Inhibition in Polyethylene," *IEEE Transactions on Electrical Insulation,* vol. EI-14, pp. 200-206, 1979.

[32] T. Miyashita, "Deterioration of Water-Immersed Polyethylene-Coated Wire by Treeing," *IEEE Transactions on Electrical Insulation,* vol. EI-6, pp. 129-135, 1971.

[33] J. Sletbak, "The mechanical damage theory of water treeing-a status report," in *Proceedings of the 3rd International Conference on Properties and Applications of Dielectric Materials, 1991*, 1991, pp. 208-213 vol.1.

[34] J. Saetbak, "A Theory of Water Tree Initiation and Growth," *IEEE Transactions on Power Apparatus and Systems,* vol. PAS-98, pp. 1358-1365, 1979.

[35] J. Yang, "Study on electrical characteristics of XLPE insulation of high voltage cable," Doctor of Philosophy, School of Electrical and Electronics, Nanyang Technological University, Singapore, 2009.

[36] R. A. Fouracre, S. J. MacGregor and F. Teuma, "Some properties of surface discharges," in *IEE Colloquium on Atmospheric Discharges for Chemical Synthesis (Ref. No. 1998/244)*, 1998, pp. 3/1-3/2.

[37] C. S. Chang, J. Jin, C. Chang, T. Hoshino, M. Hanai, and N. Kobayashi, "Separation of corona using wavelet packet transform and neural network for detection of partial discharge in gas-insulated substations," *IEEE Transactions on Power Delivery,* vol. 20, pp. 1363-1369, 2005.

[38] C. Hudon and R. H. Rehder, "Recognition of phase resolved partial discharge patterns for internal discharges and external corona activity," in *Proceedings of the 1995 IEEE 5th International Conference on Conduction and Breakdown in Solid Dielectrics, ICSD'95*, 1995, pp. 386-392.

[39] X. Ma, C. Zhou and I. J. Kemp, "Interpretation of wavelet analysis and its application in partial discharge detection," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 9, pp. 446-457, 2002.

[40] R. E. James, Q. Su and I. o. E. a. Technology, *Condition assessment of high voltage insulation in power system equipment*. London: Institution of Engineering and Technology, 2008.

[41] S. Whitehead, *Dielectric breakdown of solids*. Oxford,: Clarendon Press, 1951.

[42] E. Kuffel and W. S. Zaengl, "Non-Destructive Insulation Test Techniques," in *High-voltage engineering : fundamentals*, 1st ed Oxford (Oxfordshire) ; New York: Pergamon Press, 1984, pp. 422-461.

[43] J. T. Tykociner, H. A. Brown and E. B. Paine, *Oscillations due to ionization in dielectrics and methods of their detection and measurement: a report of an investigation conducted by the Engineering experiment station*: University of Illinois at Urbana Champaign, College of Engineering. Engineering Experiment Station, 1933.

[44] R. Bartnikas, "Detection of partial discharges (corona) in electrical apparatus," *IEEE Transactions on Electrical Insulation,* vol. 25, pp. 111-124, Feb 1990.

[45] C. S. Lai, Y. S. Chou, P. W. Wu and T. L. Peng, *Acoustic partial discharge detection for dry type transformer*, 2004.

[46] Y. Luo, S. Ji and Y. Li, "Phased-ultrasonic receiving-planar array transducer for partial discharge location in transformer," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control,* vol. 53, pp. 614-622, 2006.

[47] S. Ren, X. Yang, R. Zhu, B. Xi, X. Man, and X. Cao, "Ultrasonic localization of partial discharge in power transformer based on improved genetic algorithm," in *International Symposium on Electrical Insulating Materials, (ISEIM 2008)*, 2008, pp. 323-325.

[48] R. Sarathi and P. G. Raju, "Diagnostic study of electrical treeing in underground XLPE cables using acoustic emission technique," *Polymer Testing,* vol. 23, pp. 863-869, Dec 2004.

[49] S. S. Bamji and A. T. Bulinski, "Electroluminescence - an optical technique to determine the early stages of polymer degradation under high electric stresses," in

Conference Digest 2002 Conference on Precision Electromagnetic Measurements, 2002. , 2002, pp. 106-107.

[50] S. S. Bamji, A. T. Bulinski and R. J. Densley, "Final Breakdown Mechanism Of Water Treeing," in *Annual Report. Conference on Electrical Insulation and Dielectric Phenomena, CEIDP 1991.* , 1991, pp. 298-305.

[51] S. S. Bamji, A. T. Bulinski and R. J. Densley, "Degradation of polymeric insulation due to photoemission caused by high electric fields," *IEEE Transactions on Electrical Insulation,* vol. 24, pp. 91-98, 1989.

[52] Y. Ehara, M. Tsuno, H. Kishida and T. Ito, "Optical and electrical detection of single pulse of partial discharge on electrical treeing," in *Proceedings of 1998 International Symposium on Electrical Insulating Materials, 1998.* , 1998, pp. 639-642.

[53] H. Kaneiwa, Y. Suzuoki and T. Mizutani, "Partial discharge characteristics and tree inception in artificial simulated tree channels," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 7, pp. 843-848, 2000.

[54] J. M. Bryden, I. J. Kemp, A. Nesbitt, J. V. Champion, S. J. Dodd, and Z. Richardson, "Correlations among light emission and partial discharge measurements made during electrical tree growth," in *Eighth International Conference on Dielectric Materials, Measurements and Applications, 2000. (IEE Conf. Publ. No. 473)*, 2000, pp. 513-518.

[55] M. Morita, K. Wu, F. Komori and Y. Suzuoki, "Investigation of electrical tree propagation from water tree by utilizing partial discharge and optical observation," in *Proceedings of the 7th International Conference on Properties and Applications of Dielectric Materials, 2003.* , 2003, pp. 891-894 vol.3.

[56] M. Nishizaka, H. Kawabata, C. S. Kim and T. Mizutani, "Change in partial discharge characteristics by tree propagation from an artificial-simulated tree channel," in *Proceedings of the 7th International Conference on Properties and Applications of Dielectric Materials, 2003.*, 2003, pp. 887-890 vol.3.

[57] M. D. Noskov, M. Sack, A. S. Malinovski and A. J. Schwab, "Measurement and simulation of electrical tree growth and partial discharge activity in epoxy resin," *Journal of Physics D-Applied Physics,* vol. 34, pp. 1389-1398, May 2001.

[58] S. Okabe, S. Kaneko, T. Minagawa and C. Nishida, "Detecting characteristics of SF6 decomposed gas sensor for insulation diagnosis on gas insulated switchgears," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 15, pp. 251-258, Feb 2008.

[59] K. Spurgeon, W. H. Tang, Q. H. Wu, Z. J. Richardson, and G. Moss, "Dissolved gas analysis using evidential reasoning," *IEE Proceedings-Science Measurement and Technology,* vol. 152, pp. 110-117, May 2005.

[60] M. Duval, "A review of faults detectable by gas-in-oil analysis in transformers," *IEEE Electrical Insulation Magazine,* vol. 18, pp. 8-17, 2002.

[61] S. A. Boggs, "Partial discharge. II. Detection sensitivity," *IEEE Electrical Insulation Magazine,* vol. 6, pp. 35-42, 1990.

[62] National Electrical Manufacturers Association., *NEMA 107-1964 (R1971) : Methods of Measurement of Radio Influence Voltage (RIV) of High Voltage Apparatus*. Virginia: NEMA, 1971.

[63] G. H. Vaillancourt, R. Malewski and D. Train, "Comparison of Three Techniques of Partial Discharge Measurements in Power Transformers," *IEEE Transactions on Power Apparatus and Systems,* vol. PAS-104, pp. 900-909, 1985.

[64] K. Itoh, Y. Kaneda, S. Kitamura, K. Kimura, A. Nishimura, T. Tanaka, H. Tokura, and I. Okada, "New noise rejection techniques on pulse-by-pulse basis for on-line partial discharge measurement of turbine generators," *IEEE Transactions on Energy Conversion,* vol. 11, pp. 585-594, 1996.

[65] S. R. Campbell and G. C. Stone, "Investigations into the use of temperature detectors as stator winding partial discharge detectors," in *Conference Record of the 2006 IEEE International Symposium on Electrical Insulation, 2006*, 2006, pp. 369-375.

[66] R. T. Harrold and T. W. Dakin, "The Relationship Between the Picocoolomb and Microvolt for Corona Measurements on HV Transformers and Other Apparatus," *IEEE Transactions on Power Apparatus and Systems,* vol. PAS-92, pp. 187-198, 1973.

[67] E. A. Franke and E. Czekaj, "Wide-Band Partial Discharge Detector," *IEEE Transactions on Electrical Insulation,* vol. EI-10, pp. 112-116, 1975.

[68] W. N. English, "Photon pulses from point-to-plain corona," *Physical Review,* vol. 77, pp. 850-850, 1950.

[69] T. W. Dakin and J. Lim, "Corona Measurement and Interpretation," *Power Apparatus and Systems, Part III. Transactions of the American Institute of Electrical Engineers* vol. 76, pp. 1059-1065, 1957.

[70] H. Suzuki and T. Endoh, "Pattern recognition of partial discharge in XLPE cables using a neural network," in *Proceedings of the 3rd International Conference on Properties and Applications of Dielectric Materials, 1991*, 1991, pp. 43-46 vol.1.

[71] C. A. Bailey, "A Study of Internal Discharges in Cable Insulation," *IEEE Transactions on Electrical Insulation,* vol. EI-2, pp. 155-159, 1967.

[72] P. Morshuis, "Assessment of dielectric degradation by ultrawide-band PD detection," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 2, pp. 744-760, 1995.

[73] G. C. Stone, H. G. Sedding, N. Fujimoto and J. M. Braun, "Practical implementation of ultrawideband partial discharge detectors," *IEEE Transactions on Electrical Insulation,* vol. 27, pp. 70-81, 1992.

[74] V. Nassisi and A. Luches, "Rogowski coils - theory and experimental results," *Review of Scientific Instruments,* vol. 50, pp. 900-902, 1979.

[75] H. Borsi, "A PD measuring and evaluation system based on digital signal processing," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 7, pp. 21-29, 2000.

[76] R. Bartnikas, "Corona pulse probability density-function measurements on primary distribution power-cables," *IEEE Transactions on Power Apparatus and Systems,* vol. PA94, pp. 716-723, 1975.

[77] Transformers Committee of the IEEE Power Engineering Society., *IEEE Std C57.113-1991: IEEE Guide for Partial Discharge Measurement in Liquid- Filled Power Transformers and Shunt Reactors*, 1992.

[78] J. S. Johnson and M. Warren, "Detection of Slot Discharges in High-Voltage Stator Windings During Operation," *Transactions of the American Institute of Electrical Engineers,* vol. 70, pp. 1998-2000, 1951.

[79] I. A. Metwally, "Status review on partial discharge measurement techniques in gas-insulated switchgear/lines," *Electric Power Systems Research,* vol. 69, pp. 25-36, 2004.

[80] A. G. Sellars, O. Farish and B. F. Hampton, "Characterising the discharge development due to surface contamination in GIS using the UHF technique," *IEE Proceedings - Science, Measurement and Technology,* vol. 141, pp. 118-122, 1994.

[81] O. Farish, M. D. Judd, B. F. Hampton and J. S. Pearson, "SF6 insulation systems and their monitoring," in *Advances in high voltage engineering*, ed London: Institution of Electrical Engineers, 2004, pp. 37-76.

[82] B. F. Hampton and R. J. Meats, "Diagnostic measurements at UHF in gas insulated substations," *IEE Proceedings on Generation, Transmission and Distribution,* vol. 135, pp. 137-145, 1988.

[83] M. D. Judd, O. Farish and B. F. Hampton, "The excitation of UHF signals by partial discharges in GIS," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 3, pp. 213-228, 1996.

[84] J. S. Pearson, B. F. Hampton and A. G. Sellars, "A continous UHF monitor for gas-insulated substations," *IEEE Transactions on Electrical Insulation,* vol. 26, pp. 469-478, Jun 1991.

[85] M. D. Judd, L. Yang and I. B. B. Hunter, "Partial discharge monitoring for power transformers using UHF sensors Part 1: Sensors and signal interpretation," *IEEE Electrical Insulation Magazine,* vol. 21, pp. 5-14, Mar-Apr 2005.

[86] M. D. Judd, B. M. Pryor, S. C. Kelly and B. F. Hampton, "Transformer monitoring using the UHF technique," in *Eleventh International Symposium on High Voltage Engineering*, 1999, pp. 362-365.

[87] G. S. Smith, *An introduction to classical electromagnetic radiation*. Cambridge, U.K. ; New York, NY, USA: Cambridge University Press, 1997.

[88] D. Fabiani, A. Cavallini and G. C. Montanari, "A UHF Technique for Advanced PD Measurements on Inverter-Fed Motors," *IEEE Transactions on Power Electronics,* vol. 23, pp. 2546-2556, 2008.

[89] M. D. Judd, O. Farish, J. S. Pearson and B. F. Hampton, "Dielectric windows for UHF partial discharge detection," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 8, pp. 953-958, 2001.

[90] J. S. Pearson, O. Farish, B. F. Hampton, M. D. Judd, D. Templeton, B. W. Pryor, and I. M. Welch, "Partial discharge diagnostics for gas insulated substations," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 2, pp. 893-905, 1995.

[91] M. D. Judd, O. Farish and B. F. Hampton, "Broadband couplers for UHF detection of partial discharge in gas-insulated substations," *IEE Proceedings - Science, Measurement and Technology,* vol. 142, pp. 237-243, 1995.

[92] J. Lopez-Roldan, T. Tang and M. Gaskin, "Optimisation of a sensor for onsite detection of partial discharges in power transformers by the UHF method," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 15, pp. 1634-1639, 2008.

[93] T. Pinpart and M. D. Judd, "Experimental comparison of UHF sensor types for PD location applications," in *IEEE Electrical Insulation Conference (EIC 2009)*, 2009, pp. 26-30.

[94] M. D. Judd, O. Farish and J. S. Pearson, "UHF couplers for gas-insulated substations: a calibration technique," *IEE Proceedings - Science, Measurement and Technology,* vol. 144, pp. 117-122, 1997.

[95] P. J. Moore, I. E. Portugues and I. A. Glover, "Radiometric location of partial discharge sources on energized high-Voltage plant," *IEEE Transactions on Power Delivery,* vol. 20, pp. 2264-2272, 2005.

[96] Q. Shan, I. A. Glover, P. J. Moore, I. E. Portugues, M. Judd, R. Rutherford, and R. J. Watson, "TEM horn antenna for detection of impulsive noise," in *International Symposium on Electromagnetic Compatibility (EMC Europe)*, 2008, pp. 1-6.

[97] T. Hoshino, K. Kato, N. Hayakawa and H. Okubo, "A novel technique for detecting electromagnetic wave caused by partial discharge in GIS," *IEEE Transactions on Power Delivery,* vol. 16, pp. 545-551, 2001.

[98] T. Hoshino, K. Kato, N. Hayakawa and H. Okuno, "Frequency characteristics of electromagnetic wave radiated from GIS apertures," *IEEE Transactions on Power Delivery,* vol. 16, pp. 552-557, 2001.

[99] S. Matsumoto, N. Shiroi, I. Suzuki, T. Akiba, M. Sobataka, K. Kasajima, Y. Shibuya, Y. Murooka, and T. Kawamura, "Three-axis loop antenna for the detection of partial discharge signal," in *International Symposium on Electrical Insulating Materials (ISEIM 2008)*, 2008, pp. 28-31.

[100] L. Hamada, N. Otonari and T. Iwasaki, "Measurement of electromagnetic fields near a monopole antenna excited by a pulse," *IEEE Transactions on Electromagnetic Compatibility,* vol. 44, pp. 72-78, 2002.

[101] S. Kaneko, S. Okabe, M. Yoshimura, H. Muto, C. Nishida, and M. Kamei, "Detecting characteristics of various type antennas on partial discharge electromagnetic wave radiating through insulating spacer in gas insulated switchgear," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 16, pp. 1462-1472, 2009.

[102] Z. Jin, C. Sun, C. Cheng and J. Li, "Two Types of Compact UHF Antennas for Partial Discharge Measurement," in *International Conference on High Voltage Engineering and Application (ICHVE 2008)*, 2008, pp. 616-620.

[103] S. Tenbohlen, D. Denissov, S. Hoek and S. M. Markalous, "Partial discharge measurement in the ultra high frequency (UHF) range," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 15, pp. 1544-1552, 2008.

[104] A. Cavallini, A. Contin, G. C. Montanari and F. Puletti, "Advanced PD inference in on-field measurements. Part I: Noise rejection," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 10, pp. 216-224, Apr 2003.

[105] H. Zhang, T. R. Blackburn, B. T. Phung and D. Sen, "A novel wavelet transform technique for on-line partial discharge measurements. Part 1: WT de-noising algorithm," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 14, pp. 3-14, Feb 2007.

[106] X. Zhou, C. Zhou and I. J. Kemp, "An improved methodology for application of wavelet transform to partial discharge measurement denoising," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 12, pp. 586-594, 2005.

[107] I. Blokhintsev, M. Golovkov, A. Golubev and C. Kane, "Field experiences with the measurement of partial discharges on rotating equipment," *IEEE Transactions on Energy Conversion,* vol. 14, pp. 930-938, 1999.

[108] B. Dong, M. Han, L. Sun, J. Wang, Y. Wang, and A. Wang, "Sulfur Hexafluoride-Filled Extrinsic Fabry-Pérot Interferometric Fiber-Optic Sensors for Partial Discharge Detection in Transformers," *IEEE Photonics Technology Letters,* vol. 20, pp. 1566-1568, 2008.

[109] H. G. Sedding, S. R. Campbell, G. C. Stone and G. S. Klempner, "A new sensor for detecting partial discharges in operating turbine generators," *IEEE Transactions on Energy Conversion,* vol. 6, pp. 700-706, 1991.

[110] D. Wenzel, H. Borsi and E. Gockenbach, "Partial discharge measurement and gas monitoring of a power transformer on-site," in *Seventh International Conference on Dielectric Materials, Measurements and Applications (Conf. Publ. No. 430)*, 1996, pp. 255-258.

[111] D. Pommerenke, T. Strehl, R. Heinrich, W. Kalkner, F. Schmidt, and W. Weissenberg, "Discrimination between internal PD and other pulses using directional coupling sensors on HV cable systems," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 6, pp. 814-824, 1999.

[112] B. A. Fruth and D. W. Gross, "Partial discharge signal generation transmission and acquisition," *IEE Proceedings - Science, Measurement and Technology,* vol. 142, pp. 22-28, 1995.

[113] R. Itoh, Y. Kaneda, S. Kitamura, K. Kimura, K. Otoba, T. Tanaka, H. Tokura, and I. Okada, "On-line partial discharge measurement of turbine generators with new noise rejection techniques on pulse-by-pulse basis," in *Conference Record of the 1996 IEEE International Symposium on Electrical Insulation*, 1996, pp. 197-200 vol.1.

[114] H. Zhang, T. R. Blackburn, B. T. Phung, J. Hanlon, P. Taylor, and IEEE, "A novel on-line differential technique for partial discharge measurement of MV/HV power cables," in *ICPASM 2005: Proceedings of the 8th International Conference on Properties and Applications of Dielectric Materials, Vols 1 and 2*, 2006, pp. 641-644.

[115] M. Kurtz and J. F. Lyles, "Generator insulation diagnostic testing," *IEEE Transactions on Power Apparatus and Systems,* vol. 98, pp. 1596-1603, 1979.

[116] G. Stone, "Importance of bandwidth in PD measurement in operating motors and generators," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 7, pp. 6-11, Feb 2000.

[117] E. Gulski, "Discharge pattern recognition in high voltage equipment," in *International Conference on Partial Discharge*, 1993, pp. 36-38.

[118] T. Hucker and H. G. Krantz, "Requirements of automated PD diagnosis systems for fault identification in noisy conditions," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 2, pp. 544-556, 1995.

[119] H. G. Kranz, "Partial Discharge Evaluation of Polyethylene Cable-Material by Phase Angle and Pulse Shape Analysis," *IEEE Transactions on Electrical Insulation,* vol. EI-17, pp. 151-155, 1982.

[120] E. Gulski, P. H. F. Morshuis and F. H. Kreuger, "Conventional and time-resolved measurements of partial discharges as a tool for diagnosis of insulating materials," in *Proceedings of the 4th International Conference on Properties and Applications of Dielectric Materials, 1994*, 1994, pp. 666-669 vol.2.

[121] N. Hozumi, T. Okamoto and T. Imajo, "Discrimination of partial discharge patterns using a neural network," *IEEE Transactions on Electrical Insulation,* vol. 27, pp. 550-556, 1992.

[122] J. M. Braun, S. Rizzetto, N. Fujimoto and G. L. Ford, "Modulation of partial discharge activity in GIS insulators by X-ray irradiation," *IEEE Transactions on Electrical Insulation,* vol. 26, pp. 460-468, 1991.

[123] G. Wu, X. Jiang, H. Xie and D.-H. Park, "The experimental study on tree growth in XLPE using 3D PD patterns," in *Proceedings of the 6th International Conference on Properties and Applications of Dielectric Materials*, 2000, pp. 558-561 vol.1.

[124] E. Gulski, "Computer-aided measurement of partial discharges in HV equipment," *IEEE Transactions on Electrical Insulation,* vol. 28, pp. 969-983, Dec 1993.

[125] A. Contin and S. Pastore, "Classification and separation of partial discharge signals by means of their auto-correlation function evaluation," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 16, pp. 1609-1622, 2009.

[126] T. Kalicki, J. M. Braun, J. Densley and H. G. Sedding, "Pulse-shape characteristics of partial discharge within electrical trees in polymeric materials," in *Conference on Electrical Insulation and Dielectric Phenomena (Annual Report)* 1995, pp. 380-383.

[127] X. Wang, D. Zhu, F. Li and S. Gao, "Analysis and rejection of noises from partial discharge (PD) on-site testing environment," in *Proceedings of the 7th International Conference on Properties and Applications of Dielectric Materials*, 2003, pp. 1104-1107.

[128] U. Kopf and K. Feser, "Rejection of narrow-band noise and repetitive pulses in on-site PD measurements," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 2, pp. 433-446, 1995.

[129] K.-H. Kim, J.-H. Sun, C.-G. Kim, J.-K. Lee, and C.-W. Kang, "Development of on-line partial discharge detector," in *Proceedings of the 6th International Conference on Properties and Applications of Dielectric Materials*, 2000, pp. 745-748.

[130] V. Nagesh and B. I. Gururaj, "Evaluation of digital filters for rejecting discrete spectral interference in on-site PD measurements," *IEEE Transactions on Electrical Insulation,* vol. 28, pp. 73-85, 1993.

[131] Y.-H. Lin, "Using K-Means Clustering and Parameter Weighting for Partial-Discharge Noise Suppression," *IEEE Transactions on Power Delivery,* vol. 26, pp. 2380-2390, 2011.

[132] A. Contin, A. Cavallini, G. C. Montanari, G. Pasini, and F. Puletti, "Digital detection and fuzzy classification of partial discharge signals," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 9, pp. 335-348, 2002.

[133] P. D. Agoris, S. Meijer, E. Gulski and J. J. Smit, "Threshold selection for wavelet denoising of partial discharge data," in *Conference Record of the 2004 IEEE International Symposium on Electrical Insulation*, 2004, pp. 62-65.

[134] O. Altay and O. Kalenderli, "Noise reduction on partial discharge data with wavelet analysis and appropriate thresholding," in *International Conference on High Voltage Engineering and Application (ICHVE)*, 2010, pp. 552-555.

[135] H. Zhang, T. R. Blackburn, B. T. Phung and D. Sen, "A novel wavelet transform technique for on-line partial discharge measurements. Part 2: On-site noise rejection application," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 14, pp. 15-22, 2007.

[136] K. L. Wong, "Electromagnetic emission based monitoring technique for polymer ZnO surge arresters," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 13, pp. 181-190, 2006.

[137] X. Wang, B. Li, Z. Liu, H. T. Roman, O. L. Russo, K. K. Chin, and K. R. Farmer, "Analysis of partial discharge signal using the Hilbert-Huang transform," *IEEE Transactions on Power Delivery,* vol. 21, pp. 1063-1067, 2006.

[138] Y.-W. Tang, C.-C. Tai, C.-C. Su, C.-Y. Chen, J.-C. Hsieh, and J.-F. Chen, "Partial discharge signal analysis using HHT for cast-resin dry-type transformer," in

International Conference on Condition Monitoring and Diagnosis (CMD 2008) 2008, pp. 521-524.

[139] C. Caironi, D. Brie, L. Durantay and A. Rezzoug, "Interest and utility of time frequency and time scale transforms in the partial discharges analysis," in *Conference Record of the 2002 IEEE International Symposium on Electrical Insulation*, 2002, pp. 516-522.

[140] Y. H. M. Thayoob, Z. Zakaria, M. R. Samsudin, P. S. Ghosh, and M. L. Chai, "Preprocessing of acoustic emission signals from partial discharge in oil-pressboard insulation system," in *IEEE International Conference on Power and Energy (PECon)*, 2010, pp. 29-34.

[141] EEP electrical engineering portal. (2009). *Gas-Insulated Switchgear Type 8DN8*. Available: http://www.csanyigroup.com/gas-insulated-switchgear-type-8dn8

[142] Awesome Inc. (2012). *Electricity by photos*. Available: http://emadrlc.blogspot.sg/2012/06/330kv132kv-transformer-armidale.html

[143] Omicron. (2012). *Rotating Machines Diagnosis*. Available: http://www.omicron.at/en/products/app/rotating-machines/

[144] West coast switchgear Inc. (2012). *Switchgear*. Available: http://www.westcoastswitchgear.com/products.aspx

[145] EA technology. (2012). *Partial Discharge Monitoring In MV Substations*. Available: http://www.eatechnology.com/news/ea-technology-offers-energy-related-degree/partialdischargemontioringinmvsubstations

[146] G. C. Stone, "Partial discharge diagnostics and electrical equipment insulation condition assessment," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 12, pp. 891-904, 2005.

[147] S. Boggs and J. Densley, "Fundamentals of partial discharge in the context of field cable testing," *IEEE Electrical Insulation Magazine,* vol. 16, pp. 13-18, 2000.

[148] A. Cavallini, X. Chen, G. C. Montanari and F. Ciani, "Diagnosis of EHV and HV Transformers Through an Innovative Partial-Discharge-Based Technique," *IEEE Transactions on Power Delivery,* vol. 25, pp. 814-824, 2010.

[149] D.-J. Kweon, S.-B. Chin, H.-R. Kwak, J.-C. Kim, and K.-B. Song, "The analysis of ultrasonic signals by partial discharge and noise from the transformer," *IEEE Transactions on Power Delivery,* vol. 20, pp. 1976-1983, 2005.

[150] N. Davies, T. Yuan, J. C. Y. Tang and P. Shiel, "Non-intrusive partial discharge measurements of MV switchgears," in *Condition Monitoring and Diagnosis, 2008. CMD 2008. International Conference on*, 2008, pp. 385-388.

[151] M. Ren, M. Dong, Z. Ren, H. D. Peng, and A. C. Qiu, "Transient Earth Voltage Measurement in PD Detection of Artificial Defect Models in SF6," *IEEE Transactions on Plasma Science,* vol. PP, pp. 1-7, 2012.

[152] J. Zhao, C. D. Smith and B. R. Varlow, "Substation monitoring by acoustic emission techniques," *IEE Proceedings Science, Measurement and Technology,* vol. 148, pp. 28-34, 2001.

[153] M. N. O. Sadiku, "Maxwell's equations," in *Elements of electromagnetics*, 4th ed New York: Oxford University Press, 2007, pp. 385-420.

[154] M. N. O. Sadiku, "Electromagnetic wave propagation," in *Elements of electromagnetics*, 4th ed New York: Oxford University Press, 2007, pp. 429-500.

[155] Y. Li, Y. Wang, G. Lu, J. Wang, and J. Xiong, "Simulation of transient earth voltages aroused by partial discharge in switchgears," in *International Conference on High Voltage Engineering and Application (ICHVE)*, 2010, pp. 309-312.

[156] G. Luo, D. Zhang, Y. Koh, K. Ng, and W. Leong, "Time-Frequency Entropy-Based Partial-Discharge Extraction for Nonintrusive Measurement," *IEEE Transactions on Power Delivery,* vol. 27, pp. 1919-1927, 2012.

[157] D. Schlichthärle, "Analog Filters," in *Digital filters : basics and design*, ed New York: Springer, 2000, pp. 19-83.

[158] A. Antoniou, "Analog-filter approximation," in *Digital filters: analysis, design, and applications*, 2nd ed New York: McGraw-Hill, 1993, pp. 138-172.

[159] H. Okubo and N. Hayakawa, "A novel technique for partial discharge and breakdown investigation based on current pulse waveform analysis," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 12, pp. 736-744, 2005.

[160] D. L. Donoho and I. M. Johnstone, "Ideal Denoising In An Orthonormal Basis Chosen From A Library Of Bases," *Comptes Rendus De L Academie Des Sciences Serie I-Mathematique,* vol. 319, pp. 1317-1322, Dec 1994.

[161] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika,* vol. 81, pp. 425-455, Sep 1994.

[162] S. G. Mallat, "Denoising," in *A Wavelet Tour of Signal Processing : The Sparse Way*, Sparse ed Amsterdam ; Boston: Elsevier /Academic Press, 2009, pp. 535-610.

[163] S. G. Mallat, "Sparse Representations," in *A Wavelet Tour of Signal Processing : The Sparse Way*, Sparse ed Amsterdam ; Boston: Elsevier /Academic Press, 2009, pp. 1-30.

[164] I. Daubechies, *Ten lectures on wavelets*. Philadelphia, Pa.: Society for Industrial and Applied Mathematics, 1992.

[165] S. G. Mallat, "Wavelet Basis," in *A Wavelet Tour of Signal Processing : The Sparse Way*, Sparse ed Amsterdam ; Boston: Elsevier /Academic Press, 2009, pp. 263-376.

[166] I. Shim, J. J. Soraghan and W. H. Siew, "Detection of PD utilizing digital signal processing methods. Part 3: Open-loop noise reduction," *IEEE Electrical Insulation Magazine,* vol. 17, pp. 6-13, 2001.

[167] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association,* vol. 90, pp. 1200-1224, Dec 1995.

[168] S. Sardy, "Minimax threshold for denoising complex signals with waveshrink," *Ieee Transactions on Signal Processing,* vol. 48, pp. 1023-1028, Apr 2000.

[169] C. M. Stein, "Estimation of the mean of a multivariate normal-distribution," *Annals of Statistics,* vol. 9, pp. 1135-1151, 1981.

[170] B. B. Hubbard, "The world according to wavelets : the story of a mathematical technique in the making," 2nd ed Wellesley, Mass: A.K. Peters, 1998, pp. 244-245.

[171] D. Wei, A. C. Bovik and B. L. Evans, "Generalized coiflets: a new family of orthonormal wavelets," in *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems & Computers*, 1997, pp. 1259-1263 vol.2.

[172] Y. Cheng, X. Hu, X. Chen and P. Li, "Partial discharge on-line monitoring system based on FPGA," in *Proceedings of 2005 International Symposium on Electrical Insulating Materials (ISEIM 2005)*, 2005, pp. 486-489 Vol. 2.

[173] X. D. Ma, C. Zhou and I. J. Kemp, "DSP based partial discharge characterisation by wavelet analysis," in *Proceedings. ISDEIV. XIXth International Symposium on Discharges and Electrical Insulation in Vacuum*, 2000, pp. 780-783 vol.2.

[174] G. S. Almasi and A. Gottlieb, *Highly parallel computing*, 2nd ed. Redwood City, Calif.: Benjamin/Cummings Pub. Co., 1994.

[175] S. V. Adve, V. S. Adve, G. Agha, M. I. Frank, and M. J. Garzaran, "Parallel Computing Research at Illinois," in *The UPCRC Agenda*, ed: Department of Computer Science, Department of Electrical and Computer Engineering, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 2008.

[176] G. Luo and D. Zhang, "Efficiency improvement for data-processing of partial discharge signals using parallel computing," in *2010 10th IEEE International Conference on Solid Dielectrics (ICSD 2010), 4-9 July 2010*, Piscataway, NJ, USA, 2010, p. 4 pp.

[177] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Transactions on Information Theory,* vol. 38, pp. 617-643, 1992.

[178] D. Evagorou, A. Kyprianou, P. L. Lewin, A. Stavrou, V. Efthymiou, A. C. Metaxas, and G. E. Georghiou, "Feature extraction of partial discharge signals using the wavelet packet transform and classification with a probabilistic neural network," *IET Science, Measurement & Technology,* vol. 4, pp. 177-192, 2010.

[179] R. Liu, X. Sun and Z. Li, "On the application of entropy in excitation control," in *2004 International Conference on Power System Technology, (PowerCon 2004)* 2004, pp. 952-956 Vol.1.

[180] Z. Li, W. Li and R. Liu, "Applications of Entropy Principles in Power Systems: A Survey," in *IEEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific*, 2005, pp. 1-4.

[181] S. Verdú, S. W. McLaughlin and IEEE Information Theory Society., "Fifty Years of Shannon Theory," in *Information theory : 50 years of discovery*, ed New York: IEEE Press, 2000.

[182] E. Gulski and A. Krivda, "Neural networks as a tool for recognition of partial discharges," *IEEE Transactions on Electrical Insulation,* vol. 28, pp. 984-1001, 1993.

[183] J. Heaton, "Understanding backpropagation," in *Introduction to neural networks with Java*, ed St. Louis: Heaton Research, 2005, pp. 125-154.

[184] M. M. A. Salama and R. Bartnikas, "Determination of neural-network topology for partial discharge pulse pattern recognition," *IEEE Transactions on Neural Networks,* vol. 13, pp. 446-456, 2002.

[185] Z. Wang, Y. Liu and P. J. Griffin, "Neural net and expert system diagnose transformer faults," *IEEE Computer Applications in Power,* vol. 13, pp. 50-55, 2000.

[186] J. Heaton, "Overview of artificial intelligence," in *Introduction to neural networks with Java*, ed St. Louis: Heaton Research, 2005, pp. 31-48.

[187] S. S. Haykin, "Introduction," in *Neural networks : a comprehensive foundation*, 2nd ed Upper Saddle River, NJ: Prentice Hall, 1999, pp. 1-44.

[188] A. K. Jain, R. P. W. Duin and J. Mao, "Statistical pattern recognition: a review," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 22, pp. 4-37, 2000.

[189] R. Candela, G. Mirelli and R. Schifani, "PD recognition by means of statistical and fractal parameters and a neural network," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 7, pp. 87-94, 2000.

[190] S. S. Haykin, "Multilayer Perceptrons," in *Neural networks : a comprehensive foundation*, 2nd ed Upper Saddle River, NJ: Prentice Hall, 1999, pp. 138-235.

[191] B. Widrow and M. A. Lehr, "30 years of adaptive neural networks: perceptron, Madaline, and backpropagation," *Proceedings of the IEEE,* vol. 78, pp. 1415-1442, 1990.

[192] K. Guney, C. Yildiz, S. Kaya and M. Turkmen, "Artificial neural networks for calculating the characteristic impedance of air-suspended trapezoidal and rectangular-shaped microshield lines," *Journal of Electromagnetic Waves and Applications,* vol. 20, pp. 1161-1174, 2006.

[193] K. Guney, C. Yildiz, S. Kaya and M. Turkmen, "Neural models for the V-shaped conductor-backed coplanar waveguides," *Microwave and Optical Technology Letters,* vol. 49, pp. 1294-1299, Jun 2007.

[194] D. J. C. MacKay, "Bayesian Interpolation," *Neural computation,* vol. 4, pp. 415-447, 1992.

[195] G.-B. Huang, L. Chen and C.-K. Siew, "Universal approximation using incremental constructive feedforward networks with random hidden nodes," *IEEE Transactions on Neural Networks,* vol. 17, pp. 879-892, 2006.

[196] S. Jagannathan and F. L. Lewis, "Multilayer discrete-time neural-net controller with guaranteed performance," *IEEE Transactions on Neural Networks,* vol. 7, pp. 107-130, 1996.

[197] N.-Y. Liang, G.-B. Huang, P. Saratchandran and N. Sundararajan, "A Fast and Accurate Online Sequential Learning Algorithm for Feedforward Networks," *IEEE Transactions on Neural Networks,* vol. 17, pp. 1411-1423, 2006.

[198] A. Blum, *Neural networks in C++ : an object-oriented framework for building connectionist systems*. New York: Wiley, 1992.

[199] S. G. Mallat, "Time Meets Frequency," in *A Wavelet Tour of Signal Processing : The Sparse Way*, Sparse ed Amsterdam ; Boston: Elsevier /Academic Press, 2009, pp. 89-150.

[200] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proceedings of the IEEE,* vol. 66, pp. 51-83, 1978.