Partial Discharge Denoising for Power Cables

A dissertation submitted to the Department of Electronic and Electrical Engineering and the committee for postgraduate studies of the University of Strathclyde in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

by

Jiajia Liu

2019

Declaration

This thesis is the results of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signature:

Date:

Acknowledgement

I could not have this PhD project done without the support and help from people around me. In appreciation and gratitude, I wish to give my sincere thanks to all these people. Thanks for your valuable discussions, your constructive advice, and your kind encouragement.

I would like to express my deepest thanks to Prof W.H. Siew and Prof J. Soraghan for their excellent supervision during this journey. I owe tremendous thanks to Professor Siew who made it possible for me to undertake my PhD study through the award of a University Scholarship. He gave me a lot of advice to keep me on the right track to the final destination. The most important advice from him that helps me is how to be a qualified researcher. My project mainly focuses on partial discharge denoising, which is highly dependent on the knowledge of signal processing. Before I started this project, I knew not very much about signal processing. I wish to give my sincere thanks to Prof Soraghan, who gave me a lot of technical suggestions and support in signal processing. His unending support has helped me to progress my study at a reasonable pace and hence to complete my study.

Apart from the support from supervisors, I also wish to give my sincere thanks to Dr Faisal P. Mohamed. I really appreciate his valuable advice on partial discharge denoising. When he was in the research group, I received tremendous help from him for learning signal processing from scratch, and then, how to apply these fundamentals to my PhD project. I also wish to give my gratitude to Dr Xiao Hu and Dr Xiaosheng Peng for their generous support of real PD data to test proposed denoising algorithms in my project. My appreciation goes to all my colleagues during the period for my PhD project for their efforts to make a pleasant working environment for me.

I would like to give my thanks to my family. I wish to thank my wife, Dr Ziyi Wei, for her love and encouragement through this long journey. I also wish to thank my parents-in-law, Mr. Fanping Wei and Mrs. Dai Sun for their financial and moral support along this journey, and their caring for my little son, Jonathan Lau. Finally, I would like to express my gratitude to my mum, Mrs. Yinzhen Qin, for her love and understanding.

Abstract

Partial discharge (PD) diagnostics is considered a major and effective tool for the monitoring of insulating conditions of power cables. As such, a large amount of off-line or online PD measurements have been deployed in power cables during the past decades. However, challenges still exist in PD diagnostics for power cables. Noise is one of the challenges involved in PD measurement. This thesis develops new algorithms based on the characteristics of both PD signals and noise to improve the effectiveness of wavelet-based PD denoising. In the meantime, it presents new findings in the application of empirical mode decomposition (EMD) in PD denoising.

Wavelet-based technique has received high attention in the area of PD denoising, it still faces challenges, however, in wavelet selection, decomposition scale determination, and noise estimation. It is therefore the first area of interest in this thesis to improve the effectiveness of existing wavelet-based technique in PD detection by incorporating proposed algorithms. These new algorithms were developed based on the difference of entropy between transformed PD signals and noise, and the sparsity of transformed PD signals corrupted by noise.

One concern commonly expressed by critics of wavelet-based technique is a pre-defined wavelet is applied in wavelet-based technique. EMD is an algorithm that can decompose a signal based on the signal itself. Thus, the second area of interest in this thesis is to further investigate the application of EMD in PD denoising; a technique that does not require the selection of a pre-defined signal to represent the “unknown” signal of interest. A new method for relative mode selection (RMS) was proposed based on the entropy of each intrinsic mode function (IMF). Although this new method cannot outperform the existing ones, it reveals that RMS is not as important as claimed in the application of EMD in signal denoising. Also, PD signals, especially those with lower magnitudes, can receive serious distortion through EMD-based denoising.

Finally, comparisons between wavelet-based and EMD-based denoising were implemented in the following aspects, i.e., executing time, distortion, effectiveness, adaptivity and robustness.

Results unveil that improved wavelet-based technique is more preferable as it can present better performance in PD denoising.

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

AC	Alternating Current
AE	Acoustic Emission
AM	Amplitude Modulation
CBM	Condition-based Monitoring
CBWSS	Correlation-based Wavelet Selection Scheme
CRMS	Correlation-based Relative Mode Selection
CT	Current Transformer
CWT	Continuous Wavelet Transform
DAC	Damped Alternating Current
DC	Direct Current
DGA	Dissolved Gas Analysis
DSI	Discrete Spectral Interference
DWT	Discrete Wavelet Transform
EBWSS	Energy-based Wavelet Selection Scheme
EERMS	Energy Entropy-based Relative Mode Selection
EHV	Extra High Voltage
EMD	Empirical Mode Decomposition
EMD-DT	Empirical Mode Decomposition Direct Thresholding
EMD-IT	Empirical Mode Decomposition Interval Thresholding
EMD-TR	Empirical Mode Decomposition Traditional Denoising
EPR	Ethylene Propylene Rubber
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FM	Frequency Modulation
GIL	Gas Insulated Lines
HFCT	High Frequency Current Transformer
HPLC	High Performance Liquid Chromatography
HV	High Voltage

HVAC	High Voltage, Alternating Current
HVDC	High Voltage, Direct Current
IIR	Infinite Impulse Response
LV	Low Voltage
ME	Magnitude Error
MSE	Mean Square Error
MV	Medium Voltage
NMRMS	Noise Model-based Relative Mode Selection
OVW	Oscillating Voltage Waves
PD	Partial Discharge
PE	Polyethylene
PILC	Paper Insulated Lead Covered
PRPD	Phase Resolved Partial Discharge
PVC	Poly-Vinyl Chloride
QMF	Quadrature Mirror Filters
RFCT	Radio Frequency Current Transducer
RFI	Radio Frequency Interference
RMS	Relative Mode Selection
SCR	Significant Coefficients Ratio
SNR	Signal to Noise Ratio
STFT	Short-Time Fourier Transform
SURE	Stein’s Unbiased Risk Estimate
TDR	Time Domain Reflectometry
UHF	Ultra-High Frequency
VLF	Very Low Frequency
WEBWSS	Wavelet Entropy-based Wavelet Selection Scheme
WN	White Noise
WSS	Wavelet Selection Scheme
WT	Wavelet Transform
XCORR	Cross-Correlation Coefficients

XLPE	Cross-link Polyethylene

List of Publications

Journal:

1. Liu, J., Siew, W. H., Soraghan, J. J., Hu, X., Peng, X. & Morris, E. A., "A novel wavelet selection scheme for partial discharge signal detection under low SNR condition", 25 Nov 2018, (Accepted/In press) In: CIGRE Science and Engineering.

Conference:

2. Liu, J., Siew, W. H., Soraghan, J. J. & Morris, E. A., "A novel wavelet selection scheme for partial discharge signal denoising", IEEE Conference on Electrical Insulation and Dielectric Phenomenon (CEIDP), Cancun, Mexico, Oct 2018.
3. Liu, J., Siew, W. H. & Soraghan, J. J., "On-line PD Diagnostics in HVDC Cables", 9th Universities High Voltage Network Colloquium (UHVnet), Cardiff, UK, Jan 2016.

1 Introduction

1.1 Research Background

The development of electrical power networks has been accelerated by the ever-increasing demand of worldwide electricity consumption. This trend seems to continuously increase in the near future, resulting in the need to improve current electrical power networks. As electrical apparatuses have been continuously penetrated into a wide spectrum of electrical power networks, the expected improvement is primarily dictated by the reliability of these apparatuses. Insulation fault is one of the primary causes of electrical apparatus failures. To cut the cost of operation and maintenance incurred by insulation faults, power utilities continuously resort to diagnostic techniques in practice. In relation to power cable systems, PD is considered the initial indicator of the insulation fault and may lead to eventual insulation breakdown in power cables [1]–[7]. PD measurement is therefore of tremendous importance to prevent the insulation failure of power cables and consequent outage of power supply.

In the past decades, a large number of PD measurements has been deployed in power cable systems to investigate the characteristics of PD as well as its negative effect on insulation systems. Although PD measurement is a major and effective tool for the monitoring of insulation degradation of power cables, it still has existing challenges for its application in practice. Figure 1-1 briefly illustrates these challenges for the deployment of online PD measurement in power cable systems,

- *Attenuation*: High-frequency PD signal can suffer serious attenuation when it propagates along the power cable to both ends of the cable, where it may be picked up by sensors. The degree of attenuation is proportional to the travelling distance between the origins of PD signals and the locations of sensors placed. That is, a PD signal will suffer more serious attenuation if its origin is far from the locations of those sensors placed to pick it up. This attenuation is primarily due to dielectric loss of the solid dielectrics and the radial displacement current travelling through the resistance of the semiconducting layers

of power cables [8].

- *Noise*: The magnitude of PD signals generated from defects in power cables can be extremely small, which makes PD measurement vulnerable to noise [9]. Noise can reduce the credibility of PD detection as a diagnostic tool, particularly when PD signals are completely buried into it. Thus, how to effectively extract PD signals of interest from noise is of great importance for the application of PD measurement in power cable systems.
- *PD sensor placement*: Various PD sensors have been investigated for PD signal detection in power cables. High frequency current transformer (HFCT) is a predominant sensor due to its convenient installation on the earth strap at the cable termination. However, it may be a problem for underground cables as the access to earth strap is often limited due to constructional reasons [10]. Also, the number of PD sensors placed for PD detection depends on noise and the length of the cable under measurement.
- *PD location*: Time domain reflectometry (TDR) is a common method used for PD location in power cable systems. However, its accuracy of locating PD defects is highly dependent on the effect of attenuation and noise. It becomes even worse when more than one defect exists.

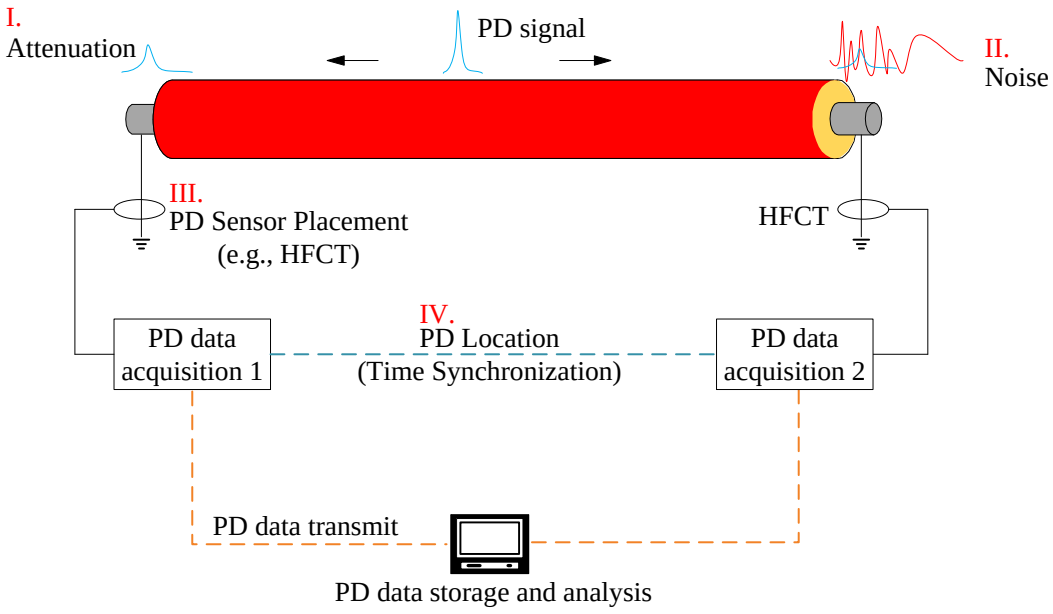

Figure 1-1 Challenges for the application of online PD measurement in power cable systems

Based on the discussion above, noise is one of most important challenges for PD monitoring in practical applications. As a result, techniques for noise reduction in PD measurement have been widely studied in past decades. The extent to which noise that can be suppressed is a major concern when a denoising strategy employed to mitigate the effect of noise on PD detection. In this project, the noise-induced challenges will be investigated and corresponding techniques will then be proposed to reduce its effects on PD measurement in power cables.

1.2 Research Motivation

In PD measurement, noise that contaminates PD signals may include discrete spectral interference (DSI), or radio frequency interference (RFI) in some references, pulse-shape noise and white noise [11]–[13]. To alleviate the effect of noise, signal processing techniques, e.g., Digital filters [14]–[16], Matched filters [17], [18],wavelet transform [12], [19]–[25], and empirical mode decomposition (EMD) [13], [26]–[29], have been investigated in PD detection during the past decades.

For many applications in signal denoising, it is difficult to have a perfect strategy for noise cancellation as there is no access to the instantaneous value of noise or the noise-only signal [30]. As such, noise cannot be completely eliminated, but it may be reduced based on the statistics of the signal and the noise process in an average sense. In relation to the area of PD detection the denoising techniques mentioned above can reduce noise that accompanies with PD signals, although each of them has its own limits in noise reduction.

Digital filters are only applicable to certain type of noise, e.g., stationary noise, the variance of which is independent of time. This is due to the working mechanism of digital filters, which relies on the Fourier Transform (FT) to analyze signals in an infinite interval. FT only transfers the information of a signal from its time domain to frequency domain. Due to the transient and non-periodic nature, FT cannot reveal the important information, e.g., time of arrival and pulse duration, carried by PD signals [9]. Also, the application of digital filters for PD denoising requires prior knowledge of noise, which is not the case in practice as noise varies with time. It is therefore extremely difficult to determine the noise presented in PD

measurement, particularly in online PD measurement.

Matched filters are a basic signal processing tool for the extraction of a known waveform from a signal that has been interfered by noise. Provided the information of PD signals is known, weak PD signals can be detected since matched filters can maximize the signal-to-noise ratio (SNR) of the output signal. The factors, such as the type and location of PD sources, detecting circuits and the travelling path of PD signals, lead to the variety of PD signals in practical measurements, which impede the application of matched filters for the requirement of prior knowledge of PD signals. Also, the effectiveness of matched filters for PD noise reduction is affected by noise types. Matched filters are considered effective only if white noise is presented in PD measurement [9].

Wavelet transform (WT) has received high attention in signal denoising due to its appealing attribute in signal representations. It can simultaneously represent signal information both in time and frequency domain, which is particularly important for the analysis of non-stationary signals, e.g., PD signals. Compared to digital filters and matched filters, wavelet-based technique can be more flexible and robust in separation of PD signals from severe noise disturbance. The success for the application of wavelet-based technique in PD denoising, however, is dependent on three major challenging aspects, i.e., mother wavelet selection, decomposition scale or level determination, and threshold estimation.

Although wavelet-based technique has been widely used in PD denoising, it is often criticized for its non-adaptivity due to the use of a predefined wavelet. As such, empirical mode decomposition (EMD), an adaptive and data driven algorithm, has been applied to PD denoising in recent years. A noisy PD signal can be decomposed into a few intrinsic mode functions (IMFs) by a sifting iteration. The relevant mode selection (RMS) is claimed to be important for EMD-based denoising, as it can distinguish the IMFs that are dominated by noise or not. As a result, the effectiveness of noise reduction can be improved if threshold is only applied to the IMFs that primarily contain signal information. Note that the threshold in EMD-based denoising is estimated by a white noise model. As aforementioned, noise that encounters in PD measurements is probably not white noise only, which may lead to inaccuracy in PD denoising via EMD-based technique.

As mentioned above, the performance of wavelet-based technique still relies on mother wavelet selection, decomposition scale determination, and threshold estimation. Thus, the first motivation of this research is to improve the performance of wavelet-based technique in PD denoising through these three aspects with further study of the statistical characteristics of PD signals and the associated noise. Then, EMD-based technique is further investigated as an alternative for PD denoising, since it does not require the selection of a pre-defined signal to represent the "unknown" signal of interest. Finally, an appropriate denoising strategy for PD detection is recommended based on the assessment of wavelet-based and EMD-based techniques for PD denoising.

1.3 Research Contributions

The research contributions in this thesis are subsequently outlined in Chapter 4 – 8, and the summary of the contributions are as follows:

- Wavelet selection is the first, and most, significant step in the application of wavelet-based technique for PD denoising. The match of PD signals and the wavelet used can maximize the reduction of noise accompanied with detected PD signals. Investigations have been implemented on the existing techniques, i.e., correlation-based and energy-based technique, for wavelet selection from a predefined wavelet library. The wavelets in this library are from Daubechies family ('db2' – 'db25'). These wavelets have been used in the existing techniques. As such, wavelets selected from Daubechies family are used to directly compare the denoising results with the existing techniques. The number at the postfix of 'db' represents the number of vanishing moments of the wavelet, e.g., the wavelet 'db2' has two vanishing moments (A vanishing moment limits the wavelet's ability to represent polynomial behavior or information in a signal). It is found that correlation-based technique is vulnerable to noise, and the criterion for wavelet selection in energy-based technique is not strictly true for PD signals, especially for damped oscillating pulse- (DOP-) type PD signals. Based on the analysis of the statistical characteristics of PD signals and the associated noise, two novel scale-dependent schemes have been proposed for the selection of appropriate wavelets

respectively. One scheme is based on the high-order statistics, i.e., kurtosis, of the transformed PD signals. The other is based on wavelet entropy, which is inspired by Shannon entropy and the associated information cost function in information theory. Results of simulated PD signals demonstrate that the kurtosis-based wavelet selection scheme is less effective than the existing schemes, while the entropy-based scheme outperforms the existing schemes for both simulated and real PD signals. As such, the entropy-based criterion has been adopted as the first improvement in wavelet-based technique for PD denoising.

- Decomposition scale in wavelet-based technique has a major effect on denoising results. Low-scale decomposition cannot sufficiently remove the noise while high-scale decomposition can lead to the loss of signal information. However, decomposition scale in wavelet-based technique is normally determined by trials and errors, or by experiences, or by an empirical formula (termed wavelet length method in this thesis). None of these methods has supports from a solid theory. In this thesis, a novel method has been proposed to choose an optimal decomposition scale for wavelet-based technique through sparse nature of PD signals in wavelet domain. The novel method is dependent on the significant coefficients ratio (SCR) of transformed PD signals. The investigations on simulated PD signals and a real PD signal with single pulse show that the scale selected by the sparsity-based method can have the best or close to the best value denoising results. In turn, the sparsity-based method for decomposition scale determination is extended to real PD signals, together with the newly proposed wavelet selection scheme. The denoising results show that the effectiveness of wavelet-based technique for PD denoising has been substantially improved as compared to the wavelet length method.
- Threshold estimation is a key step in wavelet-based denoising. The threshold used determines the effectiveness of noise suppression. Underestimated or overestimated threshold can lead to severe distortion of denoised signals. The investigation of threshold estimation is initiated from the existing threshold estimations integrated in MATLAB. Denoising results of simulated PD signals show that the (revised) universal threshold is close to the noise level within PD signals. The universal threshold is estimated based on

the variance of noise, however. To reduce the dependence on noise for threshold estimation, a novel threshold estimation has been proposed. It is derived from the SCR of the transformed PD signals. Based on the value of SCR, the transformed PD signals are divided into four categories, i.e., noise-only, noise-dominated, signal-dominated and signal-only signals. Then, the new threshold has been estimated through the number of significant coefficients and SCR for each category. Denoising results of simulated and real PD signals demonstrate that the new threshold can improve the performance of wavelet-based technique for PD denoising.

- A new thresholding function proposed in the denoising of hydrologic series data has been referenced in this application of wavelet-based technique for PD denoising. The traditional hard- and soft-thresholding have their inherent limits: the discontinuous nature of hard-thresholding function and the derivative of soft-thresholding function is not desired for optimization problems. The new thresholding function can overcome the drawback of traditional thresholding functions. With its application in wavelet-based denoising of simulated and real PD signals, the effectiveness of wavelet-based technique has been remarkably improved.
- Relevant mode selection (RMS) is claimed to play a significant role in EMD-based denoising. RMS can distinguish an IMF if it is noise-only or dominated by signal. A noise-only model or an empirical value is required in the current existing RMSs. From a practical point of view, a new RMS is proposed based on the signal itself. The new RMS is derived from the concept of energy entropy. During the test of this new RMS, it is found that it can only present better denoising results when PD signals are not heavily contaminated by noise. It is also interesting to find that the RMS is not as important as claimed in the denoising of PD signals. Based on numerous tests, all the IMFs retained for denoising can present the best denoising results. In the meantime, a series of existing EMD-based denoising techniques, i.e., EMD-TR, EMD-DT, and EMD-IT, have been reviewed. EMD-IT can provide the best denoising results among these techniques with a newly developed code.

  As a reference method, EMD-based technique is investigated to highlight the advantages

of wavelet-based technique in PD denoising. The factors, such as inaccurate interpretation of IMFs and noise estimation, are analyzed to reveal the drawbacks of EMD-based technique for PD signal extraction. Based on this, the improved wavelet-based technique is recommended for denoising of PD detection of power cables in practice due to its strengths in effectiveness improvement.

1.4 Organizations of the Thesis

This thesis is comprised of nine chapters, and the structure of the thesis is described as follows.

Chapter 1:

Chapter 1 is the introduction of this thesis, including research background, research motivation, research contributions, and the thesis structure.

Chapter 2:

Chapter 2 introduces the literature review related to PD phenomenon in practice. The definition of PD and its underlying mechanism are introduced. The most commonly forms of PD signals in practice are presented, which is followed by the description of PD characteristics. The understanding of PD characteristics is beneficial for the interpretation of PD data, e.g., pattern recognition. As the research contributions in this thesis are expected to be applied to power cables, a brief introduction of power cable theory is presented in this chapter. It then moves to the discussion of PD detection. Various methods for detection based on PD-induced phenomena have been introduced. Following this, noise as one of the challenges for PD measurement has been discussed based on different noise sources.

Chapter 3:

Chapter 3 focuses on the topic of denoising techniques, e.g., digital filters, matched filters, wavelet-based technique, and EMD-based technique, which are frequently used in the area of PD denoising. The fundamentals, principles, scope of application, together with the strengths

and weaknesses of these techniques are fully introduced. As wavelet-based technique and EMD-based technique are the primary strategies investigated in this thesis, these two techniques are presented with more details. In the meantime, the issues raised in wavelet-based and EMD-based technique are briefly discussed. These will be main topics in the following chapters.

Chapter 4:

Chapter 4 proposes a novel wavelet selection scheme for wavelet-based technique. The motivation of this new scheme is originated from the discussion of limits in the current existing wavelet selection schemes, i.e., correlation-based and energy-based schemes. In terms of further study of statistical characteristics of PD signals and the associated noise, a kurtosis-based and a wavelet entropy-based criterion are proposed for the selection of an appropriate wavelet at each decomposition scale. Denoising results of simulated PD signals suggest that kurtosis-based criterion is not as effective as current existing techniques, while wavelet entropy-based criterion presents its advances in PD denoising.

Chapter 5:

Chapter 5 aims to provide a novel method for decomposition scale determination in wavelet-based technique. The new method is inspired by the sparseness measurement method. To overcome the requirement of an empirical value in the sparseness measurement method, a new idea is derived from the concepts of sparsity in compressive sensing and Shannon entropy in information theory. Wavelet entropy introduced in last chapter is adopted to determine the number of significant coefficients of transformed PD signals. In turn, the sparsity can be reflected by the ratio of the number of significant coefficients to the length of that transformed PD signal. The scale selected by the sparsity-based method can have the best or close to the best SNR for denoised PD signals. This has been demonstrated by numerous simulated PD signals. Following this, the combination of the novel wavelet selection scheme proposed in last chapter and the sparsity-based method for decomposition scale selection is applied to simulated PD signals to embody its advance in PD denoising as compared to the

wavelet length method.

Chapter 6:

Chapter 6 investigates the effectiveness of current existing threshold estimations in MATLAB for wavelet-based denoising. The universal threshold can present best denoising results for PD signals among these existing threshold estimations. However, the universal threshold is estimated based on the variance of noise. To reduce the dependence on noise, a new scale-dependent threshold is estimated by a new formula comprised of the SCR value and the number of significant coefficients. Denoising results of simulated PD signals show that the new threshold estimation can improve the effectiveness of wavelet-based denoising than the universal threshold. In the meantime, a new thresholding function proposed in denoising of hydrologic series data is referenced to overcome the discontinuity of hard-thresholding function and derivative of soft-thresholding function. Denoising results indicate that the performance of wavelet-based technique can be increased by the use of this new thresholding function.

Chapter 7:

Chapter 7 discusses the role of RMS in EMD-based technique for PD denoising. A new adaptive RMS is proposed based on the energy entropy of IMFs. However, denoising results of simulated PD signals show that it only functions well when the SNR of original PD signal is relatively high. During the investigation of RMS, it is interesting to find that RMS is not as important as claimed in EMD-based denoising. With numerous tests, all the IMFs remained for further denoising can provide better denoising results than any RMS. Following this, a threshold estimated from a white noise model is adopted to each IMF to remove the noise. The performance test of EMD-based denoising is implemented on simulated PD signals. Although PD pulses can be extracted, the pulses with small magnitudes suffer serious distortion after denoising. Also, the tolerance of EMD-based denoising to DSI is very low, which may be due to the threshold estimation applied.

Chapter 8:

Chapter 8 constitutes a critical chapter in this thesis. The proposed methods for wavelet-based denoising and EMD-based denoising are tested through simulated PD signals in previous chapters. The feasibility of all these methods for real PD data needs to be evaluated. In this chapter, real PD signals from laboratory experiment and on-site measurement are adopted to test the improved wavelet-based technique and EMD-based technique. Denoising results demonstrate that the improved wavelet-based technique still works well for real PD data, and the effectiveness is remarkably improved. However, EMD-based technique shows its ineffectiveness in real PD denoising. The reasons are discussed through the analysis of the underlying mechanism of EMD-based expansion. The comparison between wavelet-based technique and EMD-based technique is then fully discussed through executing time, distortion, effectiveness, adaptivity and robustness. It is therefore recommended that the improved wavelet-based denoising is more preferable in noise reduction due to its strength in the extraction of PD pulses, especially those with small magnitudes.

Chapter 9:

Conclusions are drawn in this chapter based on the contributions of this project. Recommendations on the potential future research directions are also discussed.

2 Partial Discharge-based Literature Review

2.1 Introduction

The reliability of high voltage (HV) power apparatus in power networks is a major concern of power utilities. Rigorous quality controls have been implemented through the design, manufacturing, and installation of HV apparatus. However, failure of these apparatuses continues while in service, which imposes pressure on power utilities. Electrical insulation is a key component of HV apparatus, and investigations reveal that, in most cases, insulation fault is the main cause of the failure of power apparatuses. The implementation of condition-based monitoring (CBM) becomes an ever-increasing demand of power utilities, since CBM can help them drive down the related high maintenance and operational costs. For the monitoring of insulation condition, partial discharge (PD) measurement is considered an indispensable, non-destructive, sensitive and powerful tool [16], [23]. It is possible to have an overall assessment of insulation conditions of power apparatus using PD data obtained through PD measurement.

In this chapter, the definition of PD is detailed in Section 2.2. PD is a common phenomenon in power apparatus, it is necessary to have a good understanding of its underlying mechanism, which is therefore introduced in Section 2.3, including the conditions for the occurrence of PDs and the recurrent nature of PDs. Section 2.4 presents the most commonly seen forms of PDs in practice. Following this, the characteristics of PDs, which are important for the analysis of PD phenomena in practice, are presented in Section 2.5. Due to power cables as a research object, related power cable theory is introduced in Section 2.6. Various methods of PD measurement are detailed in Section 2.7 from non-electrical to electrical ones. Also, PD detection methods for power cables are separately discussed in this section, including off-line and on-line approaches and PD sensor options. Section 2.8 introduces the challenges that affect the application of PD measurement in electrical apparatus, particularly in power cables. It primarily concentrates on the types of noise that frequently occur in PD measurement. Conclusions regarding the fundamentals of PDs are presented in Section 2.9.

2.2 Partial Discharge Definition

In IEC 60270 [31] or BS EN 60270 [32], partial discharge is defined as a localized electrical discharge that only partially bridges the insulation between conductors and which may or may not occur adjacent to a conductor. PDs can result from breakdown of gaseous medium in a cavity or void, or breakdown at interfaces or surfaces, or breakdown between a conductor and a floating metal component, or breakdown of gas in an electrical tree channel, and to name a few [33], [34].

The occurrence of PDs is due to the change of electric field configuration in gaseous, solid, and liquid dielectrics. Cavity-induced partial discharges have been seen as a major source that can cause progressive deterioration in the insulation systems of electrical apparatus, e.g., transformers and power cables subjected to electrical stresses at power frequencies [35], [36]. This progressive deterioration of electrical insulation can shorten the service life of electrical apparatus. In relation to power cables, polymeric insulating materials, e.g., polyethylene (PE), cross-link polyethylene (XLPE) and ethylene propylene rubber (EPR), have been widely adopted to increase their resistance to PD-induced progressive deterioration and temperature tolerance during operation. In spite of the rigorous implementation of quality control, cavities may be still formed during the period of the manufacturing, installation, and operation of power cables, and the resultant cavity-induced PDs can cause irreversible damages to their polymeric insulating materials. As such, the introduction of polymer insulated power cables stimulates a rapid development of PD diagnostics to assess their reliability for practical use.

2.3 Partial Discharge Mechanism

Partial discharge has been a hot topic in the area of electrical insulation for many decades. Various PD processes have been found in electrical insulation due to the diversity of insulating materials used in power apparatus and the associated cavity geometries within these materials. Cavities, as the most common source of partial discharges, cannot be seen since they are buried inside the electrical insulation. As such, detecting and locating PDs induced by such defects requires frequent use of PD diagnostics in practice [37]. Electron

avalanche is considered a critical process involved in the development of PDs. From a physical point of view, self-sustaining electron avalanches may only take place in gaseous dielectrics or gaseous inclusions (e.g., cavity) in solid and liquid dielectrics [38]. Based on this, a gas-filled cavity is used as the most technologically important source of PDs to reveal its underlying mechanism.

2.3.1 Preconditions for PD Occurrence

A gas-filled cavity within electrical insulation generally has a lower permittivity and a lower breakdown strength than the adjacent insulating material. The lower permittivity results in an enhanced electric field in the gas-filled cavity. The relationship between this enhanced electric field and the electric field of the adjacent insulating material is determined by the permittivity of the insulating material and the cavity geometry. Cavities with different geometries within the insulating material, depicted in Figure 2-1 [39], are used to unveil this relationship.

Given the permittivity of the insulating material is ε and the electric field of the insulating material is E_d , then the electric field of the cavity, E_c, in some cases, can be approximated by the following equations [39]:

- cavity (a) in Figure 2-1, which is a flat cavity and perpendicular to the electric field, the stress in the cavity can be calculated by

$$E_c = \varepsilon \cdot E_d \tag{2-1}$$

- cavity (b) in Figure 2-1, which is a spherical cavity, the stress in the cavity can be calculated by

$$E_c = 3\varepsilon/(1 + 2\varepsilon) \cdot E_d \tag{2-2}$$

- cavity (c) in Figure 2-1, which is long and parallel to the electric field, then the stress in the cavity tends to be equal to that in the dielectric, as shown in the following equation,

$$E_c = E_d \tag{2-3}$$

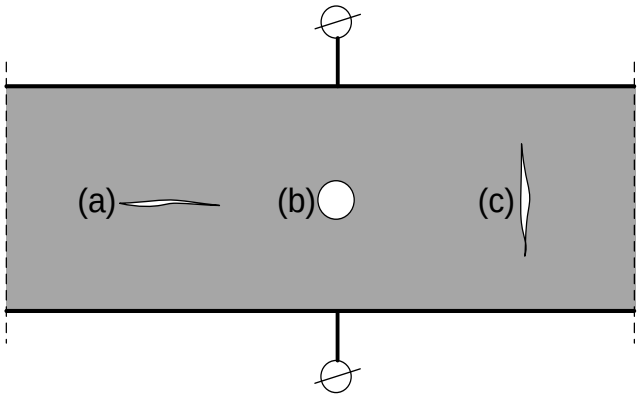

Figure 2-1 Cavities with different geometries within the insulating material [39]

Based on (2-1), (2-2), and (2-3), cavities buried in insulation systems of electrical apparatus generally subject to higher electric field than the surrounding insulating material. When the enhanced electric field in a cavity is raised with the increase of applied voltage and beyond a threshold, i.e., the minimum breakdown strength of the gaseous medium within the cavity, a PD can occur. It is worth noting that PDs will not take place instantly although the electric field within the cavity goes above its minimum breakdown strength. The presence of an initiating free electron within the cavity is another prerequisite condition for the occurrence of PDs. The initiating electron is required to appear in a favorable position in the cavity such that it can be accelerated in the electric field to obtain adequate energy for its collision with gas molecules. More electrons can be liberated from gas molecules in such manner, leaving the gas molecules ionized. This process is termed gas ionization and enables the formation of an electron avalanche. Electron avalanche is an essential step involved in the development of PDs occurring in the cavity as it can ultimately produce a conduction channel between the electrodes.

The source of an initiating free electron can be cosmic radiation, field emission, or by detrapping of electrons deposited at the cavity walls due to previous PD activity [40]–[42]. As such, the emergence of this 'first' electron is a stochastic process, which results in a statistical time lag, t_L, as shown in Figure 2-2 [43], between the instant of the minimum breakdown voltage V_{min} (equivalent to minimum breakdown strength) that has been reached and the start of an electron avalanche. t_L is typically in the order of ms and so by no means negligible in discharge process [41]. In Figure 2-2, the voltage across the cavity may exceed

V_{min} by an overvoltage ΔU due to the statistical time lag, and thus, the breakdown voltage V_I of the cavity is the sum of V_{min} and ΔU. Note that the discharge process is strongly affected by this overvoltage ΔU [42]. In short, the occurrence of cavity-induced PDs should satisfy the following two conditions: minimum breakdown strength or voltage of gaseous medium and a free electron with sufficient distance to the cavity wall.

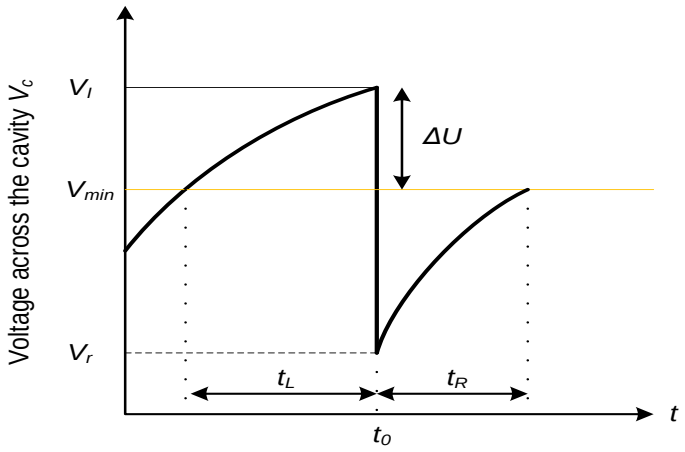

Figure 2-2 Voltage across a cavity in the dielectrics [43]

2.3.2 Recurrence of PDs

A well-known capacitive '*abc*' model, as shown in Figure 2-3 [44], is often used as an analogue circuit to investigate the fundamental quantities related to internal discharges. The components of the electrical insulation with a cavity (Figure 2-3 (a)) are equivalently replaced by some partial capacitances (Figure 2-3 (b)). C_c represents the capacitance of the cavity, C_b represents the capacitance of the components that start or end at the cavity wall, and the rest of the electric insulation is represented by C_a. V_d denotes the applied voltage across the electrical insulation.

If V_d is an alternating-current (AC) voltage, recurrent PDs can occur in the cavity as the capacitance C_c is repetitively charged and discharged when the breakdown voltage of the cavity reaches. That is, PD can repeat when electrical insulation is under continuous AC electric field beyond the minimum breakdown strength of the cavity. The repetition rate is a

critical parameter pertinent to the evaluation of the negative consequences of PDs on electrical insulation. To have a proper understanding of the recurrence of PDs, more details are explicated with the aid of Figure 2-4 [39].

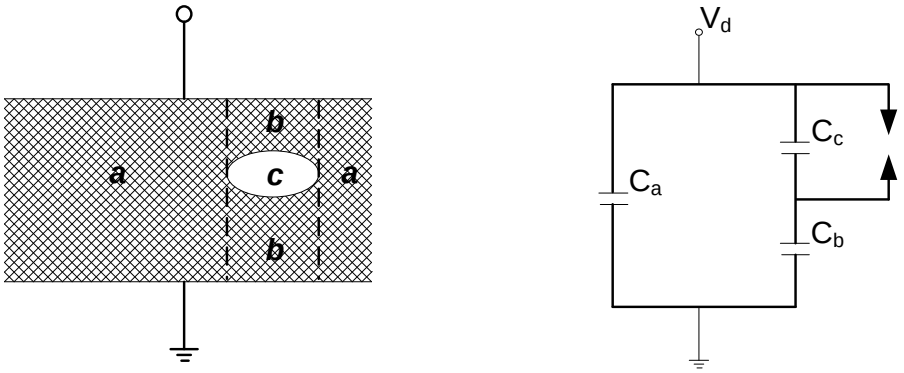

Figure 2-3 (a) Scheme of an insulation system comprising a cavity, (b) Analogue circuit [44]

In Figure 2-4, the red sinusoid V_d is an AC voltage of power frequency applied across the dielectrics. The blue dash sinusoid V_c is the voltage which would appear across the discharge-free cavity. V_{ac} is the actual voltage applied across the cavity. The black lines $V_I^+$and V_I^- represent the breakdown voltages (or inception voltages) of the cavity in the positive and negative half cycles of V_d, respectively. If the voltage V_c goes beyond V_I^+ in the positive half cycle, a PD occurs in the cavity. V_c then drops after the discharge to a residual voltage V_r^+, a voltage that cannot sustain PDs. A recovery time t_R and a statistical time lag t_L, as shown in Figure 2-2, are required for V_c to increase to V_I^+ again, and thus, a second PD occurs afterwards. This process repeats several times until the voltage V_d over the dielectrics decreases (from positive to negative half cycle) and drops to V_I^- before a new PD occurs in the negative half cycle. Recurrent PDs also can occur in this negative half cycle when applied voltage over the cavity reaches breakdown voltage. Generally, current impulses of cavity-induced PDs, as shown in Figure 2-4, are concentrated at the ascending and descending regions of the AC voltage applied across the dielectrics. Meanwhile, the recurrence of PDs is distributed at each positive and negative half cycle of the AC voltage. ΔV^+ and ΔV^- are the voltage drops between breakdown voltage and residual voltage in both polarities. If ΔV^+ and ΔV^- are equal and constant, all PD pulses will have equal

magnitudes [39], [45]. However, it is not the most frequently occurring case in practice. The repetition rate of PDs varies due to the statistical time lag required for the formation of an electron avalanche. As such, the recurrence of PDs is also a stochastic process. PD pattern recognition can benefit from this stochastic nature in the distribution of PDs as well as the associated PD magnitudes and phases [45]–[48].

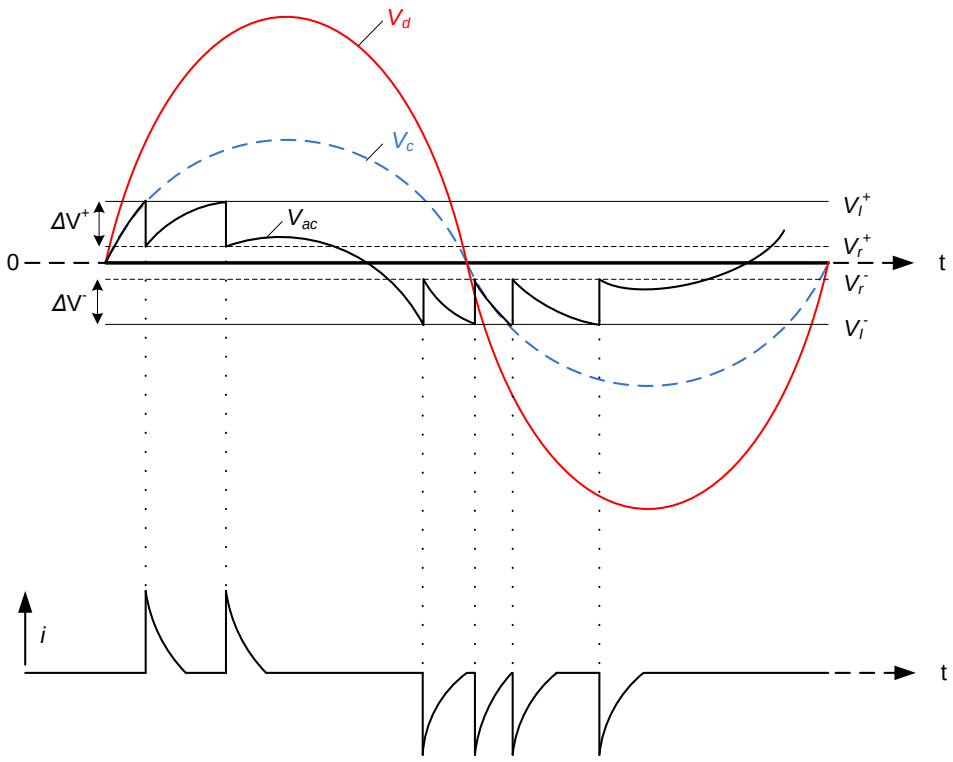

Figure 2-4 Recurrence of partial discharges [39]

2.4 Partial Discharge Forms

The term 'Partial Discharge' includes a wide range of discharge phenomena. To distinguish different PD sources, PDs can be simply categorized into surface discharge, corona discharge, electrical tree, and internal discharges [39], [44].

2.4.1 Surface discharge

Surface discharge is a form of external discharge in the sense that gas breakdown takes place

between a stress component and a dielectric surface. Research reveals that the roughness and cleanliness of the dielectric surface have impacts on surface discharge behaviors [39]. Surface discharge can occur at bushings, cable splices and terminations, the overhang of windings, and where a discharge coming from the outside but contacting the insulation surface. The locations in which various surface discharges may occur are illustrated in Figure 2-5 [39].

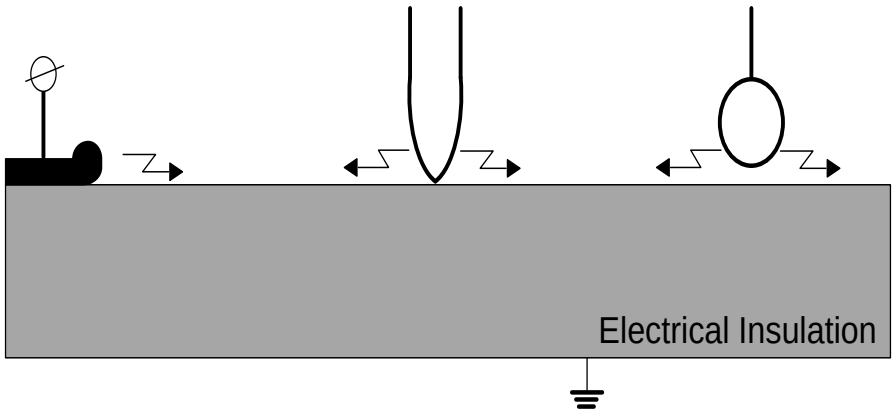

Figure 2-5 Surface discharge [39]

2.4.2 Corona discharge

Corona discharge is a 'partial discharge' that occur at sharp edges of stress components, e.g., electrodes, as shown in Figure 2-6 [39]. For a highly stressed electrode, the electric field concentrated at its sharp edge can result in the ionization of the surrounding gas. This ionization is referred to as corona discharge, and it normally produces light, audible noise, and ozone. Corona discharge is undesirable in electrical equipment due to its negative effects, e.g., the electromagnetic noise generated by corona discharge can interfere the associated low-energy-level communication or control circuits of electrical equipment [49]. In high voltage networks, corona discharge is an indicator of insufficient insulation condition and an early warning of potential flashovers [50]. To alleviate the negative effect of corona discharge in HV area, rounding the sharp edges, replacing sharp edges with large radii of curvature, and putting a round piece of wax over sharp edges have been adopted to alter the concentration of electric field at sharp edges [49]. As a result, the ionization of the surrounding gas cannot

occur due to insufficient energy. Although corona discharge is not desirable in certain cases as mentioned above, it has been surprisingly developed for a large number of commercial uses, e.g., ozone generations and electrostatic precipitator.

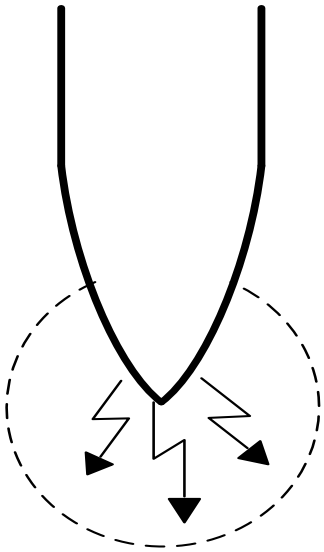

Figure 2-6 Corona discharge [39]

2.4.3 Electrical tree

Discharges in electrical tree is considered a special form of internal discharges [39]. Electrical tree can initiate from the defects in polymeric insulation under high electrical stress. Discernable damage caused by the development of electrical tree can be found in the form of a small tube or cavity after a long time since the start of electrical tree in the insulation [51]. This time can be hours, days, weeks, or even years, which is highly dictated by the applied voltage. The stem and larger branches of electrical tree, as shown in Figure 2-7 [52], then grow hallow, and thus, the number of partial discharges increases rapidly in these hollow spaces. Under continuous application of high electric stress, the electrical tree ultimately propagates across the insulation and leads to the insulation breakdown. Normally, this breakdown can take place in an extremely short period (seconds or minutes) after considerable discharges are detectable. More details regarding the inception and propagation of electrical tree in polymeric insulation can be seen from [41].

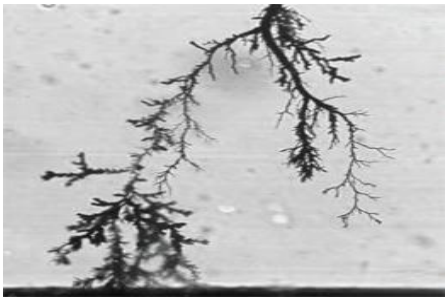

Figure 2-7 Electrical tree grown in epoxy resin [52]

2.4.4 Internal discharge

Inclusions of low dielectric strength in the dielectrics are the major sources of internal discharges. Various inclusions, including cavities (or voids in some references), dirt, paper, textile fiber, and other foreign particles, are introduced in the dielectrics during various manufacturing processes [39], [53]. Usually, gas-filled cavities are the most common inclusion within solid or liquid dielectrics. Behaviors of cavity-induced discharges are governed by the shape and the size of cavities, the type of gas, the gas pressure in cavities, and temperature. Figure 2-8 [39] delineates the internal discharges caused by cavities within the electrical insulation. Cavity-induced discharges are not expected in electrical apparatuses, particularly in those with solid insulation. These discharges can cause irreversible insulation deterioration and may ultimately lead to insulation breakdown. Electrical tree, as introduced above, may also be initiated in polymeric insulation due to partial discharges taking place in the gas-filled cavities [51].

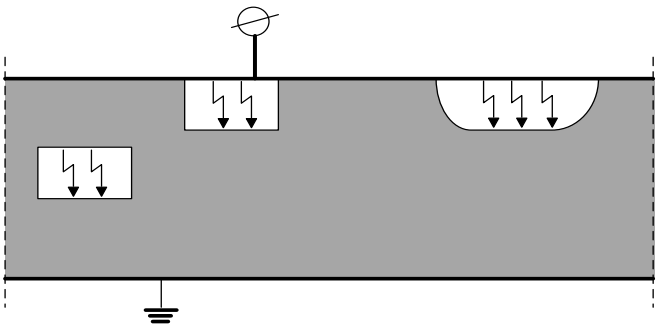

Figure 2-8 Internal discharges caused by cavities in the electrical insulation [39]

2.5 Partial Discharge Characteristics

Partial discharge characteristics include PD waveform, the rise-time time of PD pulse, pulse width, pulse magnitudes, PD repetition rate (PD numbers), and to name a few. From a practical point of view, these characteristics are essential for the analysis of PD phenomena in PD detection, location, and pattern recognition.

The rising time of a PD pulse is in the order of *ns* and the pulse duration, generally, is very short and much less than $1\mu s$ (in nanosecond range) [38], [54]. In the case of long cables, the pulse duration may be extended above $1\mu s$ due to shape distortions [55]. A typical PD pulse is illustrated in Figure 2-9, in which T_r (rise from 10% to 90% of the peak) represents the rising time of the PD pulse, and T_{FWHM} (full width at half peak) represents the pulse width [56]. The rising time of a detected PD signal is highly dependent on the rising time of this PD signal at the site of its origin and its subsequent degradation along the propagating path from the site of origin to PD detectors. This degradation is particularly dominant in PD signals that take place in electrical apparatus that exhibits transmission line behaviors, e.g., power cables, transformers and rotating machines [45]. Investigation in [57] also reveals that the possible change in the overvoltage (ΔU, mentioned above) can lead to considerable variations in the rising time of PD signals. Lower overvoltage values can cause longer delays in the formation of PD signals as well as the increase of the corresponding pulse width. From practical point of view, the degradation of the rising time of PD signals can be applied in the techniques for PD location, e.g., the rising time of PD pulses combined with the transfer function of power cables to locate the defect within the insulation systems has been proposed in [56].

A typical PD waveform is non-symmetrical, as shown in Figure 2-9. Due to this non-symmetry, PD signals have a non-Gaussian nature, which is different from the most common sources of noise, e.g., white noise and RFI, in PD detection. In statistics, the non-Gaussianity can be calculated based on the 3rd and 4th moment, i.e., the measures of skewness and kurtosis, of a series of data. As such, the skewness and kurtosis of PD signals have been investigated to minimize the effect of noise on PD detection [48], [53], [58].

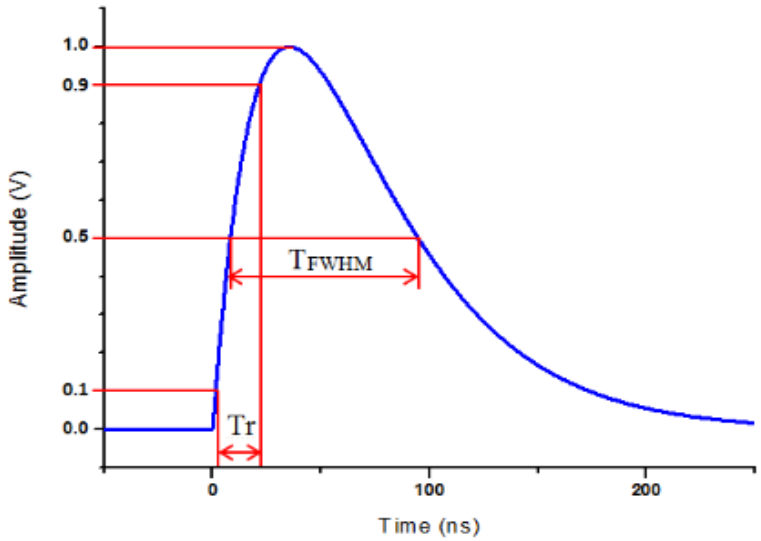

Figure 2-9 Rising time T_r and pulse width T_{FWHM} of a typical PD pulse [56]

The magnitude of PD signals is usually small, and measured in pC. In practice, the magnitudes of detected PD signals undergo substantial fluctuations due to the value of overvoltage ΔU. These fluctuations are, generally, accompanied by significant variations of the phases of the PD pulse positions with respect to the applied AC voltage. Statistical analysis of variations of PD magnitudes and the associated phases have been intensively performed, and a phase-resolved PD (PRPD) analysis has been proposed afterwards to classify the types of defects that cause the discharges. The magnitudes and phases were first used for PD pattern recognition in [59], and later the number of the discharges in one or more power cycles was introduced into this technique in [60]. As such, PRPD is often referred to as '$\varphi - q - n$' in some references (φ denotes the phase, q denotes the magnitude, and n denotes the number). An example for the use of PRPD in PD pattern recognition is illustrated in Figure 2-10 [61]. Significant differences can be seen from the statistical distributions of magnitudes (amplitudes) and phases of cavity-induced discharges, corona discharges and surface discharges.

The repetition rate, or the number, of PDs in one or more power cycles is highly correlated to the statistical time lag as mentioned before. In general, this parameter is not only involved in PD pattern recognition, but used to evaluate the severity caused by PDs on the insulation systems of power apparatus, together with PD magnitudes.

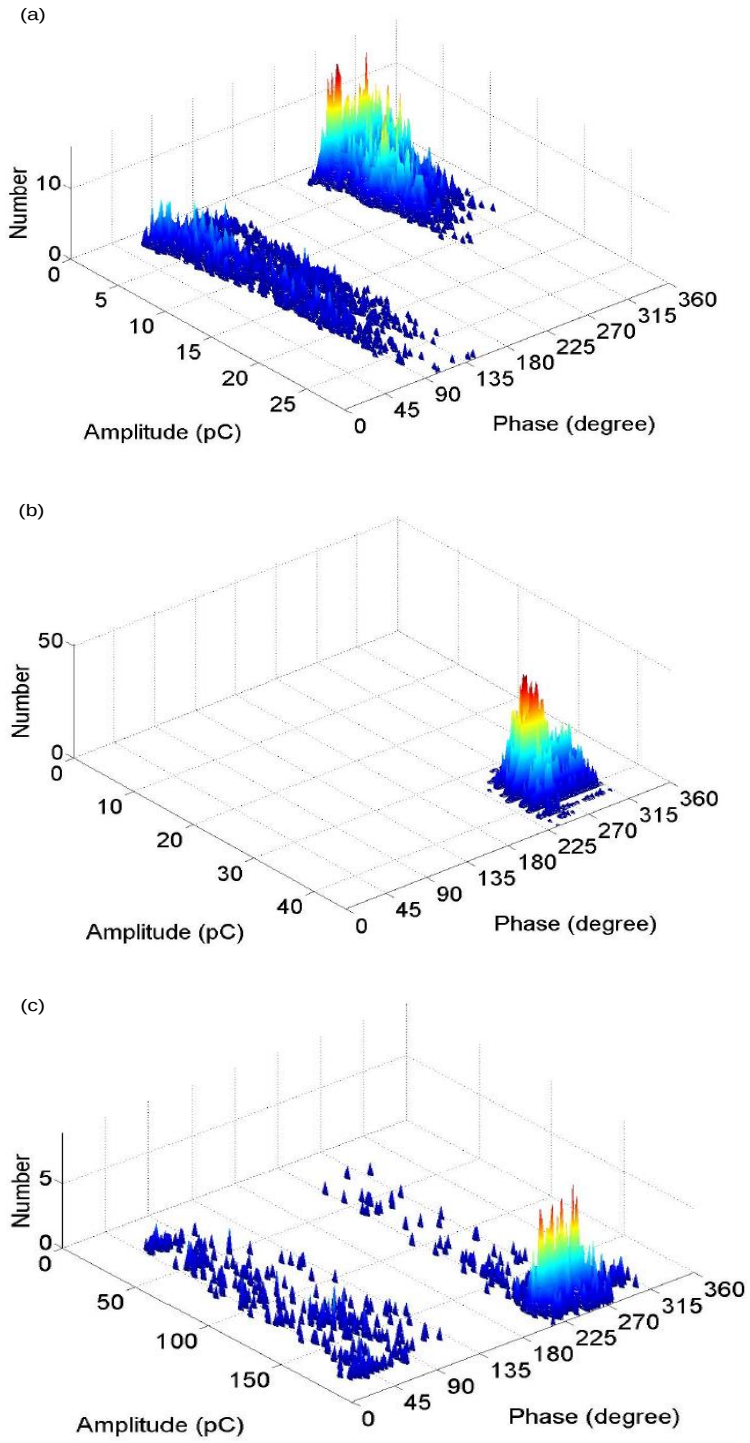

Figure 2-10 Phase resolved PD analysis. (a) 3D diagram of PD in a cavity, (b) 3D diagram of PD due to corona, (c) 3D diagram of PD at surface [61]

Based on the discussion above, the statistical behaviors of PD characteristics can be analyzed to provide useful information of PD signals. Some of PD characteristics as aforementioned have been adopted for PD detection, location and pattern recognition. Further improvement of

these techniques for PD detection, location, and pattern recognition may be achieved from the study of more statistical behaviors of PD signals.

2.6 Power Cable Theory

Power cables constitute an essential component in power transmission and distribution networks. The ever-increasing demand to replace conventional overhead lines by underground cables for power transmission and distribution is accompanied by the accelerated urbanization. The history of the application of underground cables for electrical power delivery can date back over one century [62]. The technique of underground cables has developed steadily since World War II when several varieties of synthetic rubber and polyethylene were introduced in cables. To date, underground cables play a significant role in transmission and distribution applications, particularly in urban and congested areas, and areas where overhead lines are inappropriate due to practical limitations or risks involved [62], [63]. It is necessary to mention that underground cables are always more expensive when compared to equivalent overhead lines [63]. As such, their applications continue to be focused on specific areas as mentioned above. Submarine cables have been widely used to transport electrical power across water bodies. Recent years have seen a rapid development of high voltage direct current (HVDC) cables due to the expansion of renewable energy, e.g., offshore wind power, into current power supply and the need for inter-continent bulk power transmission. The figures regarding the dramatic increase of HVDC cables are shown in Figure 2-11[64]. HVDC cables are in preference to transmit the massive amount of electrical power for long distances due to their substantially lower power loss as compared to high voltage alternating current (HVAC) cables.

The design of power cables for practical uses is dictated by the factors such as circumstance of application, voltage rating, thermal rating, and environmental concerns [65]. Various applications include underground and submarine cables. In IEC 60038, cables are categorized into low voltage (LV≤1kV), medium voltage (MV≤35kV), high voltage (HV≤230kV), and extra high voltage (EHV>245kV) based on the operating voltage level. Insulation requirement of cables is dependent on the operating voltage level. Ampacity of the cable determines its

thermal rating. Environmental concerns include recycling of cable insulating materials, health issues raised by hazardous substances, and magnetic induction [53]. In Figure 2-12 [66], an extruded power cable is used as an example to illustrate the principle components of power cables. These components include conductor, conductor screen, insulation, insulation screen, metal sheath, and Jacket. It is necessary to mention that the cable example shown in Figure 2-12 is an extruded HVDC cable. However, the components in extruded HVAC and HVDC cables have no major differences except the thickness of insulation. The insulation for HVAC cables is thicker than that for HVDC cables under the same voltage level.

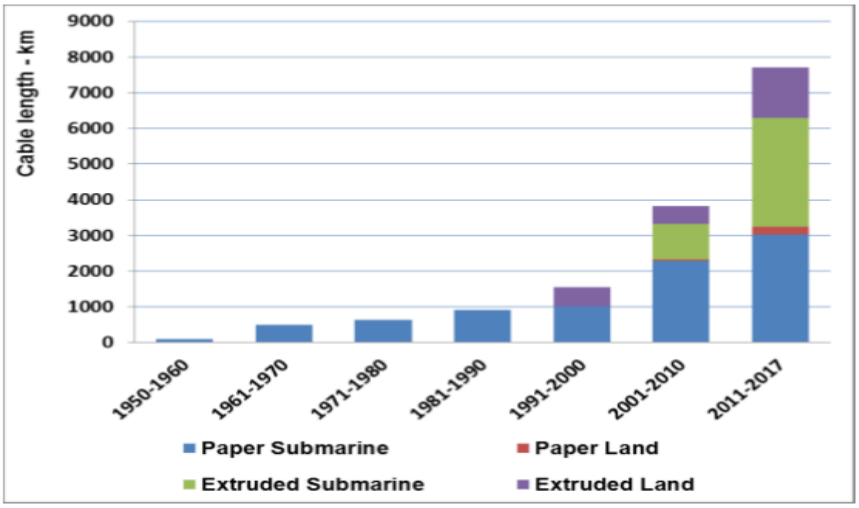

Figure 2-11 Rapid development of HVDC cables in power transmission [64]

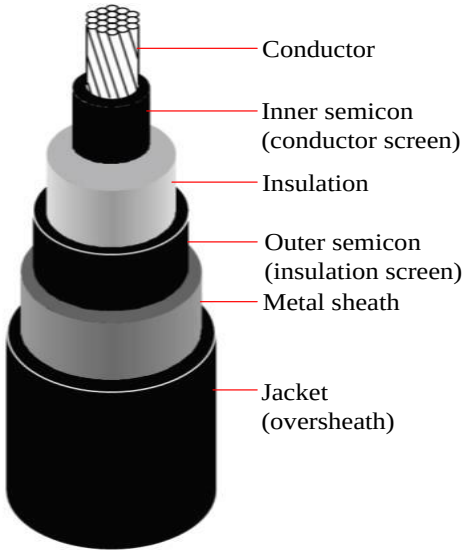

Figure 2-12 Schematic diagram showing the principle components of an extruded power cable [66]

2.6.1 Conductor

Copper and aluminum are the most frequently used materials for conductors of power cables. Copper is a more efficient conductor as compared to aluminum due to its lower electrical resistance. Also, conductors made from copper can have smaller cross sections than those made from aluminum with the same cable rating. However, aluminum is lighter than copper, and thus, can have an advantage of enabling longer lengths to be safely handled [67]. Longer length means less jointing required. Aluminum generally has a lower price than copper on the metal commodity markets. Although aluminum has lower conductivity and mechanical strength than copper, aluminum conductors are more preferable for power cables due to the economic benefits. Usually, a power cable can have 1 or 3 conductors, which depends on its applications. The conductors are designed in strands or solid as shown in Figure 2-13 [68]. Stranded conductors are flexible and can mitigate the negative effects, e.g., skin effect, and proximity effect. Solid conductors are good for soldering and compression jointing, but with low flexibility.

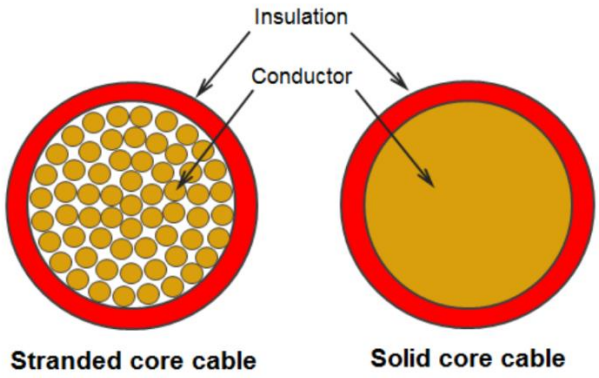

Figure 2-13 Stranded and solid conductors

2.6.2 Conductor Screen

Conductor screen normally consists of a semi-conducting layer functioning as the interface between the conductor and insulation. It is required for cables rated over 2kV, i.e., MV and HV cables [69]. The external surface of the conductor may not be smooth, particularly for

stranded conductors. As such, conductor screen can enable a uniformly distributed electrical field around the surface. For this layer to function properly, it needs to be fully bonded into the insulation to ensure that no air voids exist between the layer and the insulation.

2.6.3 Insulation

Insulation plays a significant role in power cables. As its name implies, it insulates the conductor at voltage from the outer screens which are at ground potential. The insulation is provided to withstand electric field under rated and transient operating conditions. The past century has seen a dramatic increase in the complexity of materials available for use as insulations of power cables. Figure 2-14 delineates the classifications of power cables in terms of the insulating material used. Oil-impregnated paper-insulated high voltage cables were applied for commercial uses by 1895 [62]. Polymeric materials, such as rubber and polyethylene (PE), were introduced into cable insulation during World War II. MV and LV voltage cable technology moved from mass impregnated paper insulated to extruded cross-linked polyethylene (XLPE) insulations since the 1960s [62], [64]. Ethylene propylene rubber (EPR) has been seen on market since 1980s [53]. Oil filled cable was first installed in the 1950s [62]. Gas insulated line (GIL) is an emerging technology that provides an alternative to fluid filled or XLPE cables [63].

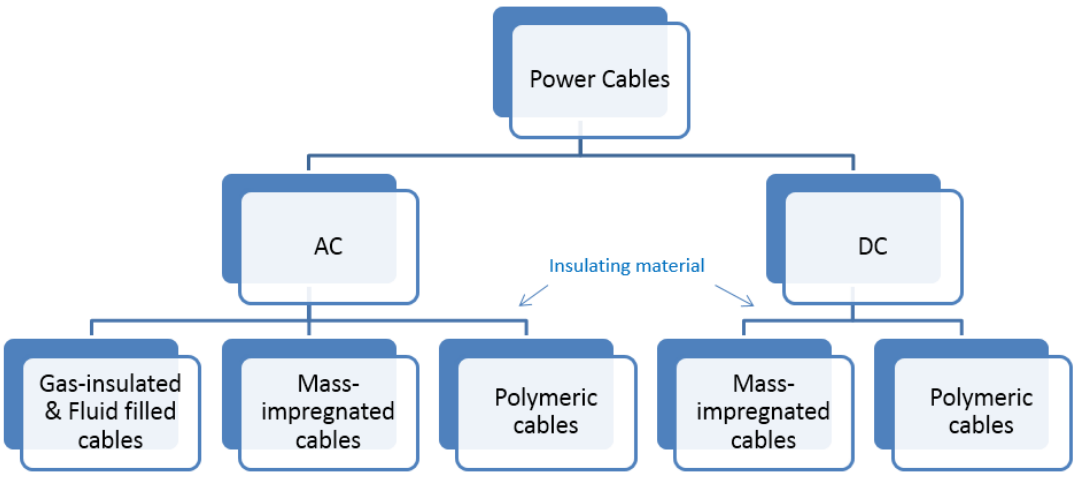

Figure 2-14 Cable classification based on the insulating material

To date, XLPE and EPR are the two major insulating materials applied in AC power cables. XLPE and mass-impregnated paper are the two main insulating materials used in HVDC power cables. The conductor temperature of XLPE insulated power cables during the normal operation can reach 90℃ as its cross-linking structure helps prevent the polymer from melting or separating at elevated temperatures. Although XLPE insulated cables have higher dielectric loss than PE, they have better aging characteristics and resistance to water treeing. EPR is a copolymer of ethylene and propylene, and also has a cross linking structure. It is more flexible than PE and XLPE, but it has higher dielectric loss than both. The temperature tolerance of EPR is better than XLPE, which can be safely operated at the temperature up to 140℃ [70]. Mass-impregnated paper is not suitable for AC cables with voltages above 33kV due to its high technical stress levels, but it is now being used for DC cables at higher voltages [63].

2.6.4 Insulation Screen

Insulation screen is a layer of semi-conductive material used in MV and HV power cables. It has the similar function as that of the conductor screen, providing a smooth transition from insulation to the grounded metal sheath.

2.6.5 Metal Sheath

Metal sheath of power cables is normally made up of concentric copper wire, laminated aluminum foil, or extruded lead alloy [67]. The main function of metal sheath is to confine electric field inside of power cables. Metal sheath needs to be grounded at least at one cable end. As such, it provides a path for the fault and leakage currents to flow away. Also, metal sheath protects the cable from moisture and other chemicals present in the environment or soil. Extruded polymeric insulations, such as XLPE and EPR, should not get exposed to moisture as the presence of a strong electric field with moisture can cause severe deterioration and ultimate failure of the insulations.

2.6.6 Jacket

Jackets or sheaths are commonly used to protect underlying layers of the cable from physical abuse, sunlight, flame or chemical corrosion [69]. There are two main materials used for cable jackets. One is poly-vinyl chloride (PVC), the other is PE. PVC is primarily used for power cables with wire armouring or lead sheaths as it is softer than PE. Although PVC has fire retardant properties, the toxic or corrosive fumes released are prohibited by many users [67]. The need of fire-retardant properties of jackets is becoming more prevalent, particularly for the cables installed in tunnels and confined spaces. The material used for jackets is expected to reduce the spread of fires and release no harmful and corrosive fumes to protect personnel and equipment. Halogen-free fire retardant (HFFR) materials are applied for this purpose. However, this type of jackets has inferior mechanical properties than PVC or PE and the cost is high.

2.7 Partial Discharge Detection

PD testing is widely accepted as a diagnostic tool for the evaluation of insulation integrity. The main purpose of this testing is to determine if, or the degree to which, the insulation system of electrical apparatus is deteriorating while in service. During apparatus operation, the electrical insulation system can be degraded due to the combination of the thermal, electrical, mechanical and environmental stresses [71]. To effectively perform PD testing, various techniques for PD detection have been developed for many years. Partial discharges occurring in electrical apparatus can produce sound, light, chemical reactions, and electrical pulse currents [44], [72]. As such, PD activities can be detected by the techniques based on the aforementioned acoustic, optical, chemical and electrical phenomena. All these techniques used to ascertain PD activities can be divided into two categories, i.e., electrical and non-electrical PD detection.

2.7.1 Non-electrical PD detection

Non-electrical PD detections, as the name suggests, are techniques that measure PD signals

occurring in electrical apparatus based on the aforementioned non-electrical phenomena, i.e., sound, light, and chemical reactions, pertinent to PD activities.

2.7.1.1 Acoustic Detection

Partial discharge activities can cause collisions between molecules of the insulating materials resulting in the generation of acoustic emission (AE) waves. These waves propagate from the discharge-generating sites through the insulation and can be detected by AE sensors. Acoustic technique has remarkable advantages of being non-invasive and immune to electrical interferences [73]. However, AE waves are prone to be distorted during the propagation by a variety of factors, which imposes the difficulty on the detection and interpretation of acoustic signals from PDs. The acoustic technique is more suitable for the detection and location of PDs in oil-filled transformers, gas insulated substations (GIS) and switchgear [3], [72], [74]. Although acoustic technique has been applied in power cables [75], it requires that AE waves are not highly attenuated in this application, that is, AE sensors need to be placed as closely as possible to the origins of PDs.

2.7.1.2 Optical Detection

The optical technique for PD detection has two different measurement strategies, namely, direct and indirect optical detection. Direct optical detection detects the light generated as a result of various ionization, excitation and recombination processes during the discharges. Indirect optical detection evaluates the changes of an optical beam within an optical fiber caused by PD activities.

In direct optical detection, the optical spectrum of light produced by PD activities varies with the types of PD. For a specified PD activity, the optical spectrum of the associated light is dependent on the chemical composition of its adjacent medium and the discharge intensity, e.g., the spectrum of PD in air is dominated by the spectrum of nitrogen [76]. Figure 2-15 and Figure 2-16 illustrate the optical spectrum of the light produced by a breakdown arc in transformer oil and corona discharge, respectively [77]–[79]. Differences between the optical spectra can be seen from the wavelength and discharge intensity. As such, the optical

spectrum of light emitted by PDs need to be taken into account at the selection of an optical sensor that can yield optimal measuring sensitivity. A photomultiplier is generally involved in the direct optical detection due to its high sensitivity of small light quantities [78]. The application of direct optical detection has been investigated in oil-filled transformer and GIS in laboratory measurement [78], [80].

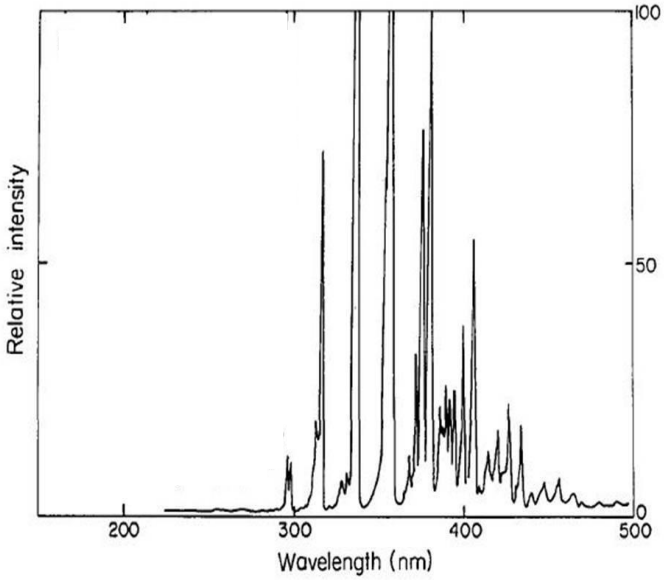

Figure 2-15 The optical spectrum of the light produced by corona discharge [76]

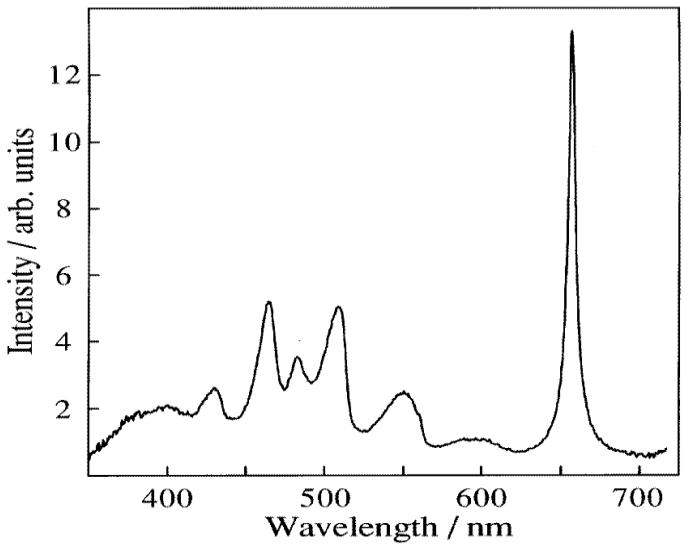

Figure 2-16 The optical spectrum of the light produced by a breakdown arc in transformer oil [74]

In indirect optical detection, PD-induced stress on the fiber optic cable can cause deformation

of the fiber optic structure. This deformation in turn can change the transmission properties, i.e., refractive index and fiber length, of the optical beam traversing the fiber. Subsequent interpretation of the demodulated optical beam can indicate the occurrence of PDs. The investigations of indirect optical detection can be seen in [81]–[83].

Although both direct and indirect optical detection are insensitive to electrical interference and detected PD signals suffer extremely low degradation during transmission, their applications have not been widely used in field measurement. Factors contributing to this may include the insulating materials in which PD occurs need to be optically transparent (not applicable in most cases), invasive nature of the technique, and cost of equipment [74], [78].

2.7.1.3 Chemical Detection

PD activities can result in chemical changes in the insulating materials. These chemical changes have been investigated for the detection of PDs in electrical apparatus. Generally, there are two chemical techniques, namely, Dissolved Gas Analysis (DGA) and High Performance Liquid Chromatography (HPLC), developed for practical uses [84], [85]. DGA is primarily used in liquid-insulated electrical apparatus, e.g., oil-filled transformer. Gases, such as hydrogen and ethane, are generated in the transformer oil due to PDs. Detection of these gases dissolved in the oil can provide information of PDs. Equally, PD activity in gas-insulated electrical apparatus can be detected in a similar manner through the analysis of gas samples [72]. For instance, SOF_2 and SO_2F_2 are produced by PD activity in a SF_6-insualted GIS [74], [86]. HPLC is also applied to detect PDs in power transformers with paper insulation. Glucose and degraded forms of glucose are generated due to PD-induced degradation of paper insulation [74]. As such, PD information can be obtained through monitoring glucose levels in the impregnated oil by HPLC.

Unfortunately, the application of chemical detection of PDs in practice is limited. Nature, location, and intensity of PD activity cannot be ascertained through chemical detection. Also, chemical detection requires the analysis of liquid or gas samples from liquid or gas dielectrics, which is not suitable for on-line PD measurement.

2.7.2 Electrical PD Detection

Electrical PD detections are the most commonly used and effective detection methods as compared to non-electrical PD detections. Also, the issued IEC standard for PD testing is related to electrical PD detections. Electrical techniques for PD detection are based on the capture of a PD pulse at the terminals of a test object. The flexibility of the test object in electrical PD detection is one of the reasons for its prevalence in both academic and industrial applications. The test object can be a simple dielectric test sample for fundamental investigations, or a real HV electrical apparatus for manufacturing quality check. Figure 2-17 delineates a typical PD measuring circuit recommended by the IEC standard [31]. The elements of the measuring circuit are listed as follows,

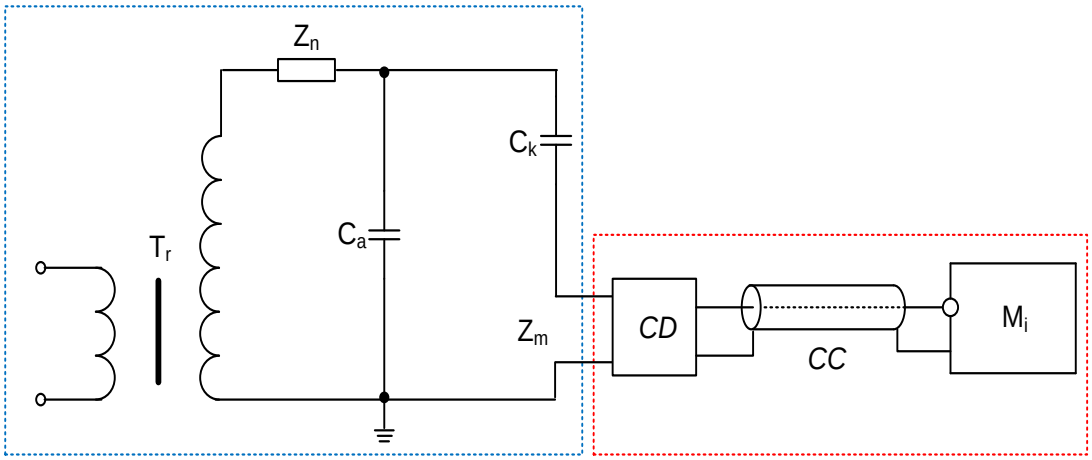

Figure 2-17 Most common PD measuring circuit used in practice

1. A high voltage source, T_r ,which is preferably PD free,
2. A test object, C_a,
3. A noise block filter, Z_n,
4. A coupling capacitor, C_k, which facilitates the pass of the high-frequency current impulses, and generally is the same order of magnitude as test object,
5. An input impedance of measuring system, Z_m, across which voltage impulses are caused by the discharge in test object,
6. A coupling device, *CD*, which is an integral part of the measuring system and test circuit,

7. A connecting cable, *CC*, which transmits the voltage impulses from *CD* to the measuring instrument,
8. A measuring instrument, M_i, which can store and manifest PD pulses in various forms.

PD measuring circuit can be divided into two parts: test circuit (blue rectangle in Figure 2-17) and measuring system (red rectangle in Figure 2-17). It is worth noting the function of coupling device in the measuring system. As an integral part of the measuring system and test circuit, the frequency response of an individual coupling device is not of general interest, but the magnitude and frequency characteristics of the input impedance are of tremendous importance as this impedance interacts with the coupling capacitor and test object and is thus an essential part of the test circuit [31]. Based on this, the components of coupling device are specially designed to achieve the optimum sensitivity. The input impedance varies as a result of different configurations of the coupling device, resulting in different pulse shapes of detected PD signals. Generally, the coupling device behaves often as a RC impedance circuit or a RLC impedance circuit, as shown in Figure 2-18 (a) and (b), respectively [21], [44], [87]. The transfer function of RC and RLC impedance circuit can be expressed in the following Laplace forms [21],

$$G_{RC}(s) = \frac{1}{C} \cdot \frac{1}{s + 1/\tau} \tag{2-4}$$

$$G_{RLC}(s) = \frac{1}{C} \cdot \frac{s}{s^2 + s/\tau + \omega_0^2} \tag{2-5}$$

where $\tau = RC$, and $\omega_0 = 1/\sqrt{LC}$. If $i(t)$ is an ideal Dirac current input, the output voltage $v(t)$ behaves as a damped exponential pulse in RC circuit and a damped oscillatory pulse in RLC circuit. The parameter τ is the delay time of both pulses and ω_0 is the oscillating frequency of the damped oscillatory pulse.

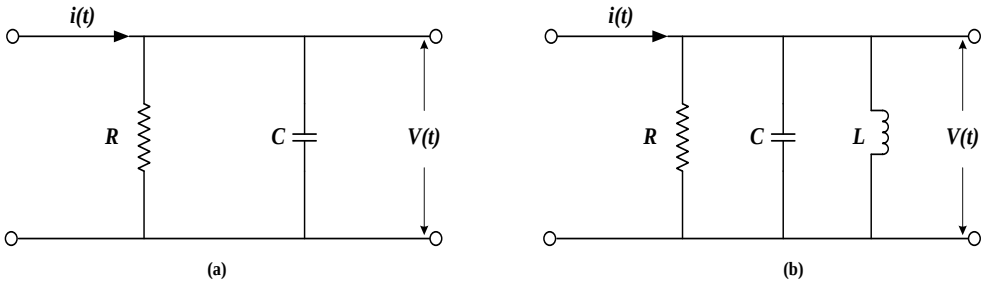

Figure 2-18 Input impedance, (a) equivalent RC circuit, (b) equivalent RLC circuit

The current impulse generated by a PD is not an ideal Dirac current impulse. As mentioned in PD characteristics, a PD pulse normally has a duration in the order of *ns* (see Figure 2-9). As such, the voltage pulse converted from the current impulse of a PD through the coupling device has a finite rising time. Based on this, the output voltage pulse of RC and RLC circuit, termed damped exponential pulse (DEP) and damped oscillatory pulse (DOP), can be simulated through the following mathematical frames (2-6) and (2-7), respectively [88].

$$v_{DEP}(t) = A(e^{-\alpha_1 t} - e^{-\alpha_2 t}) \tag{2-6}$$

$$v_{DOP}(t) = A(e^{-\alpha_1 t}\, cos(\omega_0 t - \varphi) - e^{-\alpha_2 t}\, cos\, \varphi) \tag{2-7}$$

where A denotes the peak value of the pulses, α_1 and α_2 are the rise and decay time constants of the pulses, ω_0 is the angular frequency of the oscillation of DOP, and $\varphi = \tan^{-1}(\omega_0/\alpha_2)$. Figure 2-19 (a) and (b) delineates these two typical PD pulses DEP and DOP respectively using the mathematical frames above. The sampling rate is set to 60MHz, and the values of related parameters used in (2-6) and (2-7) are listed in Table 2-1, where f_0 is the oscillatory frequency of DOP [88]. Generally, DOP signal shown in Figure 2-19 (b) is closer to a real high-frequency PD signal detected from electrical power equipment in practice [88].

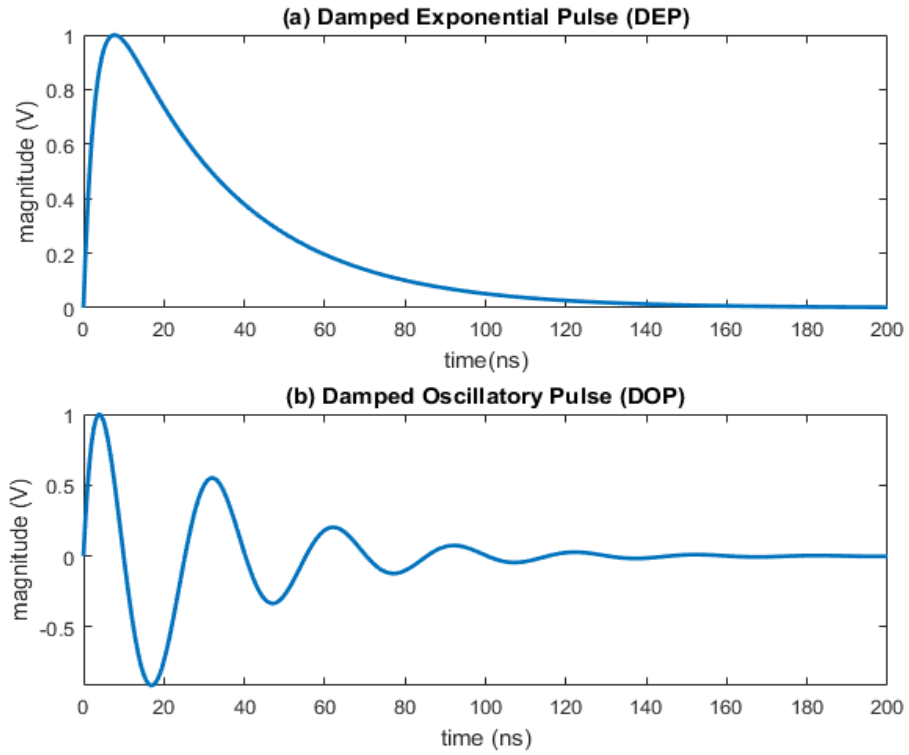

Figure 2-19 Typical PD pulses, (a) damped exponential pulse, (b) damped oscillatory pulse

Table 2-1 Values of parameters used in (2-6) and (2-7)

Parameters	Values
A	1
α_1	$10^6 s^{-1}$
α_2	$10^7 s^{-1}$
f_0	$1MHz$
w_0	$2\pi f_0$
φ	$tan^{-1}(w_0/\alpha_2)$

The measurable frequency components of PD signals are in the range of high kHz, and sometimes up to MHz due to the very short duration of a single PD signal (see Figure 2-9 and Figure 2-19). PD signals are also a type of electromagnetic waves. Based on the high-frequency and electromagnetic natures of PD signals, various PD detectors, such as current transformers (CTs), capacitive couplers, antenna and ultra-high frequency (UHF) sensors, have been successfully adopted in electrical PD detections. For a specified electrical apparatus, the selection of an appropriate PD detector is governed by its electrical property, together with its geometry and surrounding environment. The configurations of measuring circuits using these PD detectors are slightly different from that in Figure 2-17, but the basic principle for PD detection remains the same.

Although electrical techniques are frequently used for PD detection in practice, they still have their weaknesses. The measuring circuit, as shown in Figure 2-17, is highly sensitive to noise. In a severe noisy condition, PD signals generated in the test object can be totally obscured by the interferences. It is obvious that up to now numerous methods to minimize the effect of noise on PD signals have been and still are a topic for research and development. As the main research objectives of this project, PD signal denoising techniques will be fully reviewed in Chapter 3. In addition, electrical PD detection suffers serious attenuation for the propagation of a PD signal from its origins to PD detectors. As a result, the characteristics of detected PD signals are highly dependent on the geometries of electrical apparatus.

2.7.3 PD Detection in Power Cables

The significance of PD on the life of insulation systems of power cables has long been recognized. As a major and effective tool, PD detection in power cables should have the

capability of measuring a wide range of PD characteristics. In practice, off-line and on-line measurement are the two general approaches for PD detection in power cables. An off-line PD measurement requires a separate voltage source to apply on the cable that has been removed from service, while an on-line PD measurement is carried out during normal operation of power cables.

2.7.3.1 General Test Setup

An off-line PD measurement is generally deployed as shown in Figure 2-20 [89]. The cable under test needs to be taken out of service from the power network at both ends and properly isolated for safety requirements. An external voltage source is applied to energize the power cable at the near end (i.e., close to voltage source), whilst the remote end (i.e., far from voltage source) is left open. The capacitive or inductive coupling devices are used to pick up PD signals and connected to PD data processing unit. Differences can be seen from the placement of capacitive and inductive sensors in the cable for PD signal capture (e.g., label 1, 2, 3, and 4 indicate various positions of sensor installation). In general, PD measurement in HV power cables most often is performed off-line in the factory or site at regular time intervals for routine maintenance [4], [90].

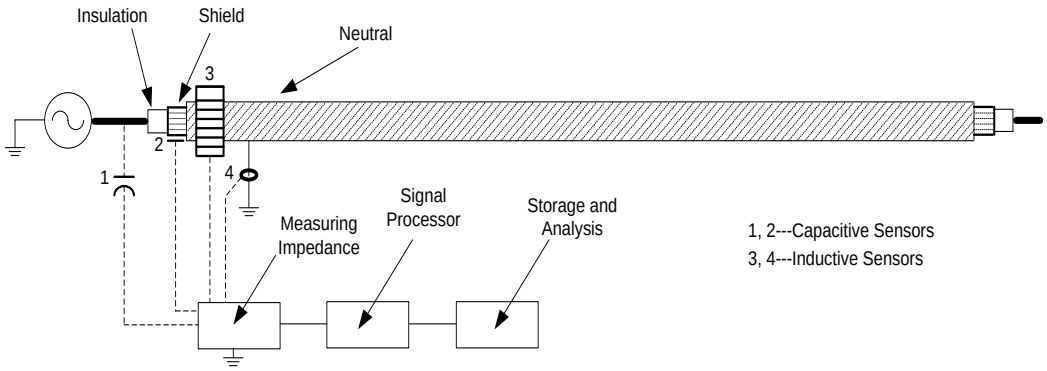

Figure 2-20 Typical setup for an off-line PD measurement in power cables

For off-line PD measurement, the cable needs to be removed from online service, and thus, an external power supply is required to energize the cable under test (as shown in Figure 2-20). The most commonly used are 50/60Hz AC voltage, very low frequency (VLF) voltage, and

oscillating voltage wave (OVW) or damped AC voltage (DAC) [4], [89], [91]–[96].

- The 50/60Hz AC voltage measuring circuit, as shown in Figure 2-21, is composed of a 50kV set-up transformer, a damping resistor with high resistance to reduce the short-circuit current if breakdown occurs in test object, and a coupling capacitor in series with a measuring impedance used to couple out PD signals [94]. This measuring circuit is almost the same as that shown in Figure 2-17. During PD measurement, the 50/60Hz AC voltage is applied to the cable system ranging from 1 to 2 times operating voltage during a limited time interval [93]. Sequences of PD pulses, together with magnitude and phase information of each pulse can be obtained through this test [94]. Using 50/60Hz ac voltage for an off-line PD measurement has several advantages, e.g., the testing voltage waveform is the same as the operating voltage waveform, and this testing method is applicable to all types of failures related to cable insulation [92]. However, it is limited to apply on long-distance cable at one time.

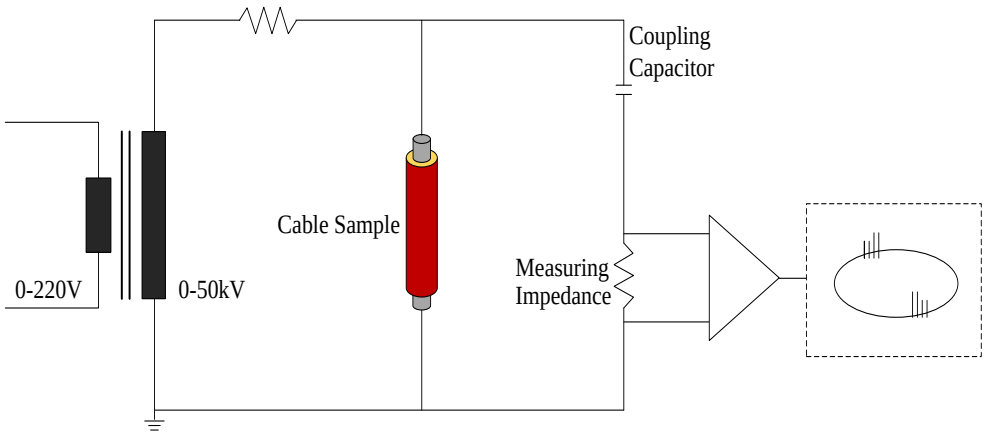

Figure 2-21 50/60Hz AC voltage measuring circuit

- Figure 2-22 [94] illustrates the general setup for the VLF measuring circuit. Very low frequency voltage uses AC signals with a frequency in the range from 0.01Hz up to 1Hz to test all the types of power cables [92], [94], [97]. The most commonly used frequency for off-line PD measurement is 0.1Hz. Low current is required to energize the capacitive components in the dielectrics of the cable, since the test is conducted at low frequency [98]. The VLF setup consists of HVAC power supply at 0.1Hz, HV divider, cable sample,

and a coupling capacitor in series with a detection impedance. The magnitudes of PD pulses and the phase position of each PD pulse can be recorded during test. However, the results of PD measurement using VLF voltage, e.g., PD magnitudes and inception voltage, may be quite different to those obtained from tests under power frequency voltages [98], [99]. Also, a higher electrical stress is required for VLF test as compared to tests performed with power frequency voltages [89], [99]. Under high electrical stress, the initiation of new weak points of the cable during test may occur and hence lead to final electrical breakdown under service stress [93].

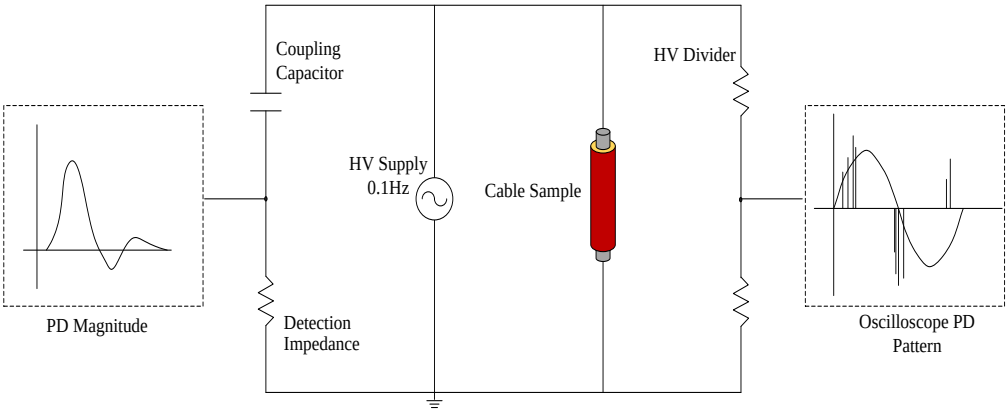

Figure 2-22 The VLF measuring circuit

- Damped alternating current voltage (DAC), also called oscillating voltage waves (OVW), is applied for off-line PD measurement in after-laying tests of new cables and in diagnostics of old cables [4], [91]–[96]. DAC uses damped alternating current at frequencies between 20Hz and 500Hz [92], [96]. It is an effective method to all types of cable systems [4]. A DAC (OVW) measuring circuit is illustrated in Figure 2-23 [94]. The testing circuit mainly has a HV DC power supply, a specially designed HV switch, an air-core inductor, HV divider, PD coupling capacitors, detection device, and cable sample. For an off-line PD measurement using DAC as its external voltage source, the cable sample is charged with a dc voltage over a few seconds to a predefined voltage level. Then a solid-state switch (HV switch *S* in Figure 2-23) with an extremely short closure time (< 1us) is closed after the DC supply is disconnected [94]. Following this

step, an oscillating voltage (DAC voltage) wave is generated by a resonant circuit, formed by the inductance of the air-core inductor and the capacitance of the cable sample, and applied to the cable sample for PD detection [91], [95]. The resonant frequency of the damped oscillation varies from tens of hertz to a few kilohertz, which is dependent on the parameters of the resonant circuit [89]. It can be calculated based on the following equation,

$$f = 1/(2\pi \cdot \sqrt{L \cdot C}) \tag{2-8}$$

where L is the inductance of the air-core inductor and C represents the capacitance of the cable sample. Testing systems using DAC have low power demand as the cable under test is charged with a dc power supply [95], [99]. The frequency of the oscillating voltage wave is as close as possible to power frequencies, and hence the testing results under DAC are generally comparable to those using power frequency voltages [92], [95]. Different from VLF test, PD under DAC test can be ignited at electrical stresses approaching to the operating voltage levels, at which new defects in the cable are not induced by test processes [99].

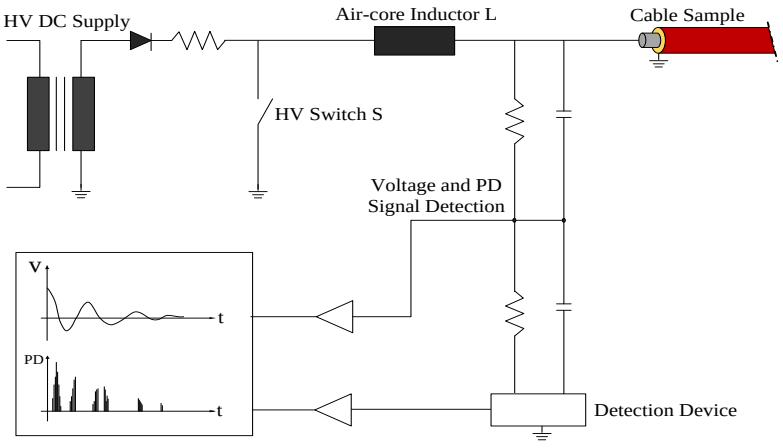

Figure 2-23 The OVW measuring circuit [94]

Different from off-line PD measurement systems, the setup of an on-line PD measurement is illustrated in Figure 2-24 [89]. The cable under PD measurement is remained in service with both ends connected to power network. It means the excitation voltage for on-line PD measurement is the operating voltage of the cable system rather than an external voltage

source. The coupling devices used to capture PD signals for on-line PD test are generally types 2, 3, and 4 shown in Figure 2-20.

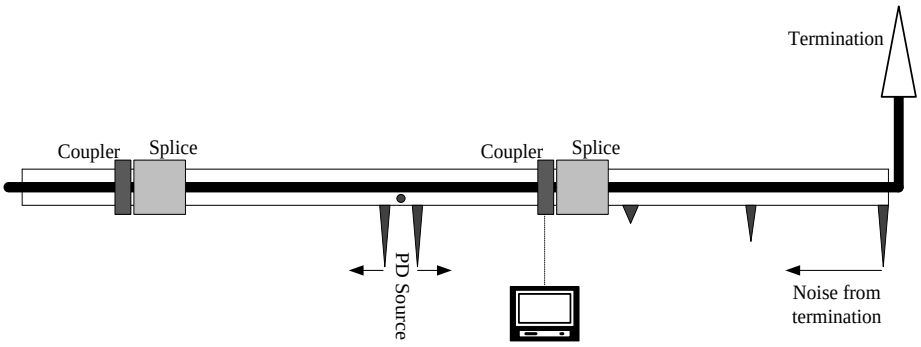

Figure 2-24 Typical setup for an on-line PD measurement in power cables [89]

It has in the past been assumed that off-line PD measurement is non-destructive to power cables under test. According to Boggs & Densley [34], and Ahmed & Srinivas [100], it is a misleading assumption because off-line PD measurements using elevated voltage stresses may accelerate the degradation of cable insulation, and thus reduce the service life of power cables. Off-line PD measurements only indicate the integrity of cable insulation at a certain time period of the routine maintenance. For power utilities it is desirable that a PD diagnostic tool can provide continuous information showing a progressive deterioration process in power cables for insulation evaluation [90], [101]. The cables under on-line measurements are stressed by the operating voltage of power network and risk of cable failure can be estimated with the trends of PD signals, e.g., the trends of PD magnitude and repetition rate can reveal the severity of the insulation degradation [90], [101], [102]. With support of the information, power utilities can make their decisions if any maintenance action is required to avoid the unplanned outages or minimize the downtime of cable systems. Therefore, on-line partial discharge measurements have received intensive investigations in recent years. Techniques now available for on-line PD measurement, ranging from simple-to-use and handheld PD screening test units to continuous (real-time) PD monitoring system, are widely deployed for detection and location of PD activities in power cables and their accessories while in service [102], [103]. The detection of PD signals through on-line PD monitoring system can provide an early warning against insulation failures [102]. However, the effectiveness of PD capture

for on-line detection is determined by the sensitivity and frequency response (bandwidth) of PD sensors as well as the noise level [102].

2.7.3.2 PD Detection Methods

In general, acoustic and electrical techniques for PD detection are often used in power cables. The pros and cons of acoustic technique have been analyzed before. For its application in power cables, the challenge is the degradation of AE waves, which requires the placement of PD detectors close to the origins of PDs. It is not the case in practical PD measurements as PD detectors are normally placed at the terminations or joints of power cables. Based on this, electrical PD detection is more prevalent for power cables than acoustic one. Capacitive and inductive coupling sensors are the most commonly used sensors for electrical PD detection in power cables (see Figure 2-20) [102].

A capacitive coupling sensor deployed in HV power cables for PD measurement is illustrated in Figure 2-25 [73]. A 40mm tin tape functioning as a capacitive sensor is wrapped around the cable at which a section of the cable metal sheath is removed. The capacitive sensor is attached on the top of outer semiconducting layer of the cable. The placement of the capacitive sensor does not impair the insulation system of the cable, since the outer semiconducting layer which works as power frequency ground does not disturb the internal field distribution of the cable [73], [104], [105]. PDs can be coupled out through a capacitive sensor and a resistive measuring impedance connected across the metal sheath interruption, as shown in Figure 2-25 [73], [106]. The use of capacitive sensors requires high frequency for PD detection, and thus, is qualified for short-distance detection [107]. From a practical point of view, it is essential to place capacitive sensors in an appropriate position. Capacitive sensors are good in the sense of sensitivity, but sensitivity check still needs to be performed by injecting calibration pulses into the system. Note that the sheath interruption can negatively affect the short circuit current handling ability, particularly for screen interruption method where a measuring impedance is connected across the interruption [73]. For on-line applications, the measuring impedance is required to place across an inductor or a voltage protection to enhance the short circuit current handling ability [73]. Based on this, the use of

capacitive sensors is not non-invasive to the cables. Also, the cable under test needs to be removed from service for the placement of capacitive sensors. Thus, awareness of the application of capacitive sensors for on-line PD measurement in power cables should be raised.

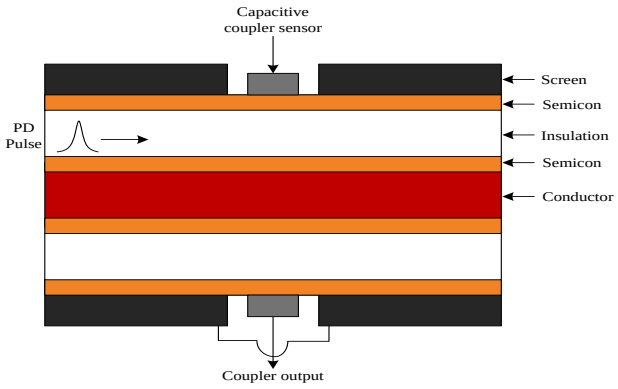

Figure 2-25 Diagram of a capacitive sensor [73]

A high frequency current transformer (HFCT) sensor, also called radio frequency current transducer (RFCT), consists of an induction coil with a ferromagnetic and split core. HFCT sensors are often clamped around the cable earth wire or the shielded cable itself to detect PD signal, as shown in Figure 2-26 [108]. When a PD occurs in a cable, HFCT sensors can pick up the PD-induced current that flows through the metallic sheath or earth wire of the cable. In turn, an induced voltage can be measured over the input impedance of the HFCTs. The use of HFCT sensors for on-line PD measurement does not require the galvanic connection with the MV or HV conductors. As a result, HFCT sensors can be placed safely around the cable for on-line PD detection without the interruption of power supply or disconnection from power network [101], [106]. To comply with the requirements of sensitivity for sensor options, a wideband frequency response (in the range from 200kHz up to 30MHz) of HFCT sensors is recommended for their applications in on-line PD measurement systems [109]. This frequency response means that HFCT sensors have the ability to detect low-frequency PD signals that propagate certain distance along the cable. HFCT sensors have a measurement range of up to 5km in XLPE cables and approximately 2.5km in paper insulated lead covered (PILC) cables [102]. The reason for this discrepancy in measurement distance between XLPE

and PILC cables is signal attenuation in mass-impregnated PILC cables is higher than that in XLPE cables [110]. Due to relatively high sensitivity, wide frequency response, and simple placement, HFCT sensors are prevalent in the applications for on-line PD detection in MV and HV cables [33].

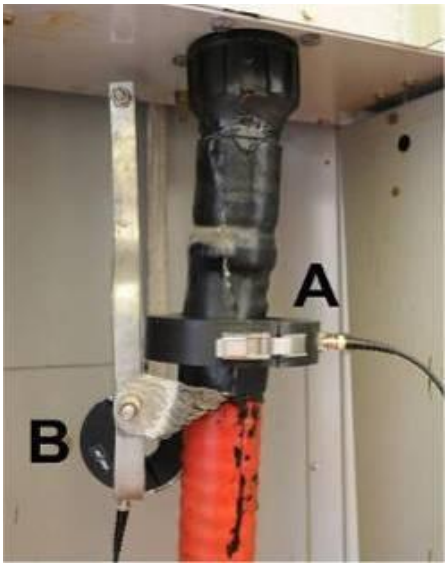

Figure 2-26 HFCT sensors attached to a 3 core 11kV cable, A: Cable with metallic sheath brought back through, B: Cable sheath connection to ground [108]

Rogowski coil is another type of inductive sensors applied for PD detection in power cables. Figure 2-27 depicts a schematic diagram of a Rogowski coil [111]. Due to relatively small inductance, the Rogowski coil can be used to measure the transient current of pulses with considerably short time duration in the order of nanoseconds [111]. These air-cored, inductive sensors do not have a problem like a ferromagnetic-core sensor that material saturation occurs under heavy load currents [102]. However, Rogowski coils have very low sensitivity and narrow bandwidth in the detection of PD signal as compared to capacitive and HFCT sensors [102], [112]. It indicates that only PD signals with large magnitudes that occur in the cable can be detected by the use of the Rogowski coil.

Apart from the sensors introduced above, coaxial cable sensors [113], [114] and directional coupling sensors [115]–[117] are also deployed as PD sensors for PD detection in power cables.

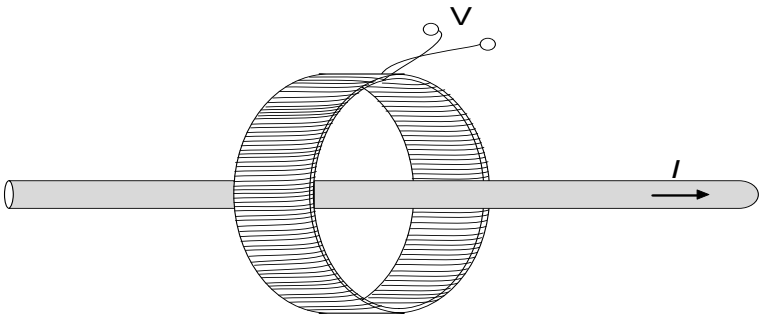

Figure 2-27 A schematic diagram of the Rogowski coil as a current sensor [111]

2.8 Challenges in PD Measurement for Power Cables

As shown in Figure 1-1, challenges, including signal attenuation, noise, PD sensor placement and location, are still existing for on-line PD measurement in power cables. In relation to the research objectives in this thesis, noise will be widely discussed in this section.

Noise is considered the major challenge for the extraction of PD signals that take place in power cables while in service. A PD signal is generally very small, while environmental noise under field conditions can be very large [34]. As a result, PD signals, particularly for those with small magnitude, can be overwhelmed by noise in on-line measurement. This imposes difficulty on the decision makers of power utilities as wrong decisions can be made based on the evaluation of insulation degradation through on-line PD measurement under this condition. To improve the detection sensitivity of on-line PD measurement, numerous signal processing techniques have been investigated to mitigate the effect of noise on PD signals. All such techniques are based on either the knowledge of the characteristics of PD signals or the nature of noise.

In general, noise that contaminates PD signals through on-line PD measurement can be broadly categorized as follows based on its nature [12], [23], [118]–[123].

1. Discrete spectral interference (DSI), also termed radio frequency interference (RFI), which mainly comes from communication and radio transmissions.
2. Pulse-type noise, including repetitive pulses that result from power electronics and random pulses that originate from switching operations or lighting or external discharges of HV apparatus.

3. White noise, sources of which can be ambient noise, amplifier noise, and thermal noise generated by the detection system.

2.8.1 Discrete Frequency Interferences

As mentioned above, DSI mainly comes from communication and radio transmissions. It is a narrow-band noise and has a sinusoidal nature. In reality, radio transmissions include amplitude modulation (AM) radio and frequency modulation (FM) radio, the frequency bands of which are 144-30000kHz and 880-108MHz, respectively. Mobile communication signals have frequency bands 890-1880MHz. With the sampling rate of tens of megahertz, the effect of FM and mobile communication signals on PD measurement are considered to be limited and only AM radio signals are considered the source of DSI for PD measurement [119], [124]. Generally, medium waves with the band of 530-1700kHz are the frequently used AM radio signals for commercial broadcasting. Based on this, DSI can be simulated by a combination of a series of amplitude modulated signals using the following formula [119],

$$dsi(t) = \sum_{i=1}^{5} (c + m \cdot sin(2\pi f_m t)) \cdot sin(2\pi f_i t) \tag{2-9}$$

where c is the amplitude of the carrier wave, m is the amplitude of modulating signal, f_m is the frequency of modulating signal, f_i is the frequency of the carrier wave. Figure 2-28 and Figure 2-29 illustrate a simulated DSI with 2048 sampling points in MATLAB using (2-9) and its spectrum, respectively. It is obvious in Figure 2-29 that the spectrum of DSI is centered around dominant frequencies. The sampling rate is in accord with that used in PD signal simulation in Figure 2-19. The values for the parameters in (2-9) are listed in Table 2-2 [119].

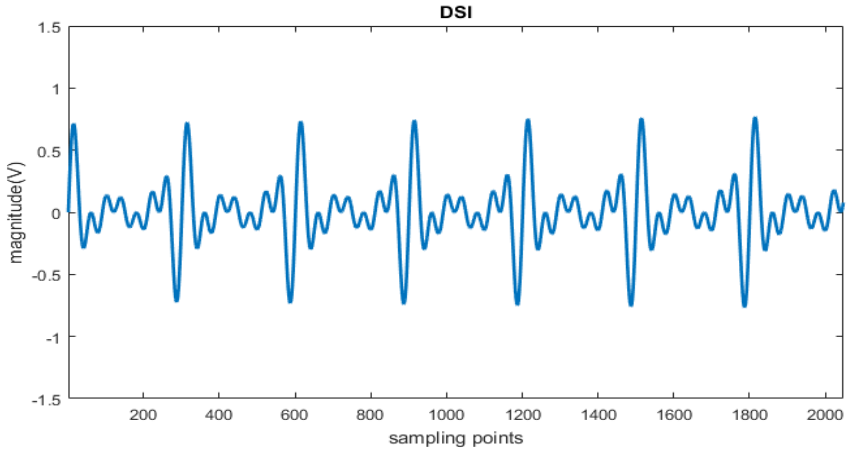

Figure 2-28 Simulated DSI with 2048 sampling points

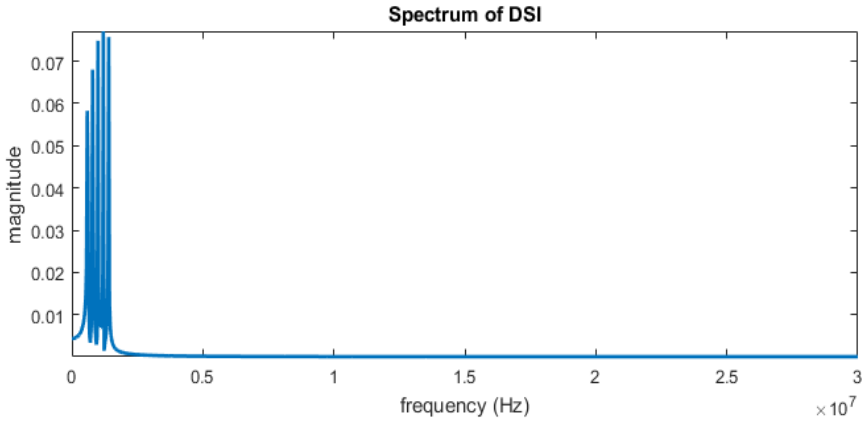

Figure 2-29 Spectrum of DSI

Table 2-2 Parameters used in (2-9) for DSI simulation

Parameters	c	m	f_m	f_1	f_2	f_3	f_4	f_5
Values	1	0.4	1kHz	600kHz	800kHz	1000kHz	1200kHz	1400kHz

2.8.2 Pulse-type Noise

Pulse-type noise, especially random pulse-type noise, is one of the most common interferences and causes significant problems for on-line PD measurement, since they have many characteristics that are quite similar to PD signals both in time and frequency domain. As aforementioned, switching operation is one of the main source of pulse-type noise, which is shown in Figure 2-30 [125]. The pulse-type noise shown in Figure 2-30 is detected from

on-site measurement conducted in [125]. From the shape of this noise caused by switching operations, it shows similarity with PD signals, and DOP-type PD signals (see Figure 2-19 (b)) in particular. Due to this similarity, this pulse-type noise could remain even if denoising strategy has been applied to the original noisy PD signals. As a result, it could be misinterpreted as PD signals occurred in the cable under monitoring, reducing the credibility of PD diagnostics. To minimize the negative effect of pulse-type noise, advanced techniques have to be adopted to extract PD signals. It is necessary to mention that this noise source in PD measurement is not covered in this thesis.

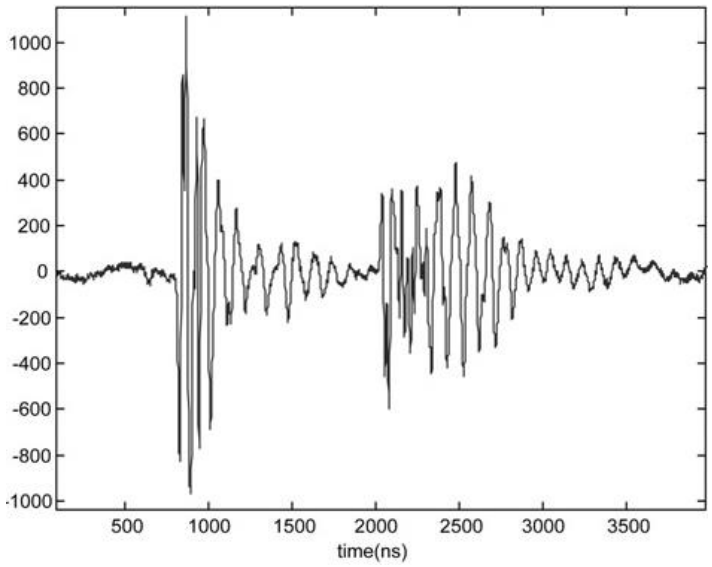

Figure 2-30 Typical pulse shape from local switchgear events [125].

2.8.3 White Noise

White noise, stemming from ambient noise, amplifier noise, or thermal noise from detection systems, is simply formed by random number with zero mean and varied standard deviation. Based on this, it can be simulated by a Gaussian model in MATLAB [12]. Figure 2-31 and Figure 2-32 delineate white noise (2048 sampling points) simulated at the sampling rate equal to 60MHz and its spectrum, respectively. Figure 2-32 demonstrates that the spectrum of white noise spreads over all the frequency domain. Figure 2-33 shows the Gaussianity of white

noise. This unique characteristic has been investigated to separate PD signals from white noise, which will be discussed later in this thesis.

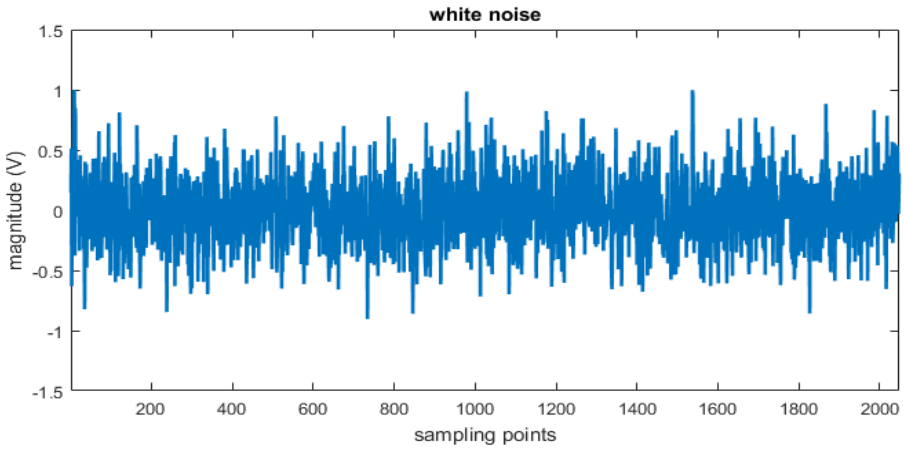

Figure 2-31 Simulated white noise with 2048 sampling points

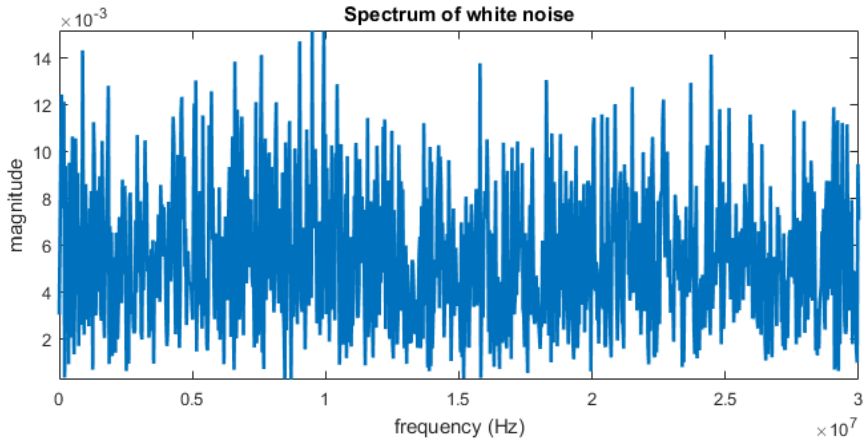

Figure 2-32 Spectrum of white noise

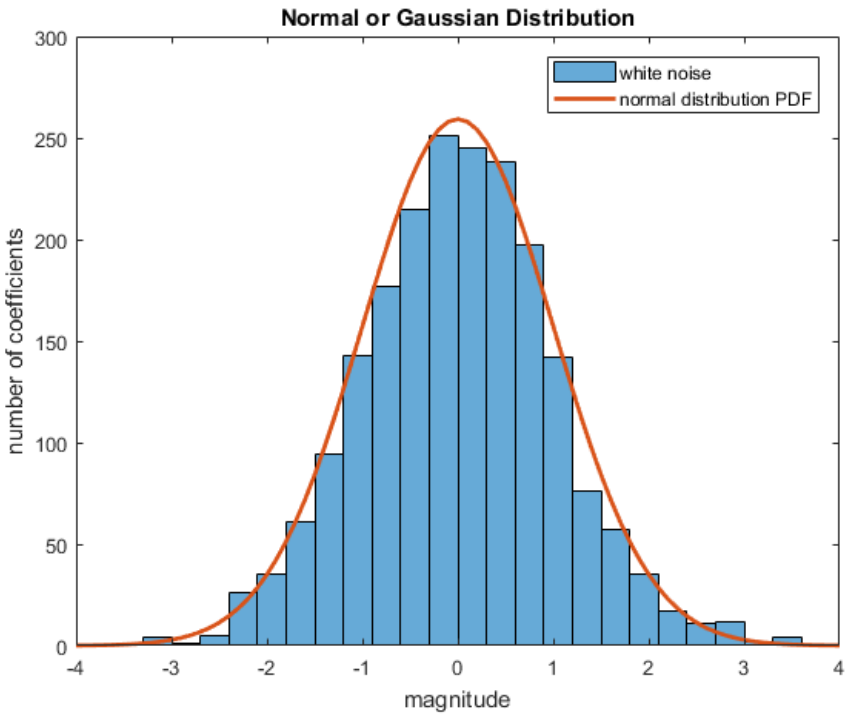

Figure 2-33 Gaussianity of white noise

As discussed above, these three types of noise are the primary sources that negatively affect the sensitivity of PD detection in power cables. The differences between noise and PD signals in both frequency and time domain can be investigated to develop related denoising strategies for the suppression of different noise. Also, the unique statistical characteristics of PD signals or noise, e.g., Gaussianity, can be integrated into signal processing techniques, and thus, PD signals can be effectively extracted. Details regarding these techniques will be explicated in Chapter 3.

2.9 Conclusion

In this chapter, PD as a very common phenomenon that occurs in high voltage electrical apparatus has been introduced through its definition and the forms of which they express in practice. To have a better understanding of PD phenomena, the underlying mechanism has been explained with details using a development of cavity-induced PD in dielectrics. It is necessary to know that the occurrence of PD requires the minimum breakdown strength or voltage as well as an initiating free electron. The emergence of this ‘first’ electron is a

stochastic process, which leads to different characteristics of PDs, e.g., magnitude and phase. For a 50Hz AC voltage, the change of voltage polarity every 10ms results in the recurrence of PD. Characteristics, including rise time of PD pulse, PD magnitude and phase, and PD repetition rate, have been presented with details as they play a significant role in PD detection, location, and pattern recognition. In relation to the research objectives of this thesis, power cable theory has been introduced through a general structure of power cables. It has been accepted that PD can cause progressive deterioration of insulation systems of electrical apparatus, and thus, it is of tremendous importance to deploy PD detectors to couple out PD signals from electrical apparatus. Past decades have seen the development of various detection methods based on the phenomena generated by PDs, such as light, sound, chemical reactions, and electrical impulses. Each of them has its own strengths and weaknesses for practical applications. In power cables, two general approaches, e.g., off-line and on-line measurement, have been discussed. Based on practical concerns, capacitive and inductive PD sensors have been analyzed for their applications in PD detection in power cables. It is an unavoidable problem that PD measurement is most often accompanied with noise. Noise can reduce the credibility of PD measurement as a diagnostic tool, particularly for on-line PD measurement. As such, the types of noise, including discrete spectral interferences, pulse-type noise, and white noise, that can contaminate PD signals in PD measurement have been discussed. The differences between noise and PD signals in both time and frequency domain can be investigated, and thus, the sensitivity of PD detection can benefit from the development of related signal processing techniques based on these differences. Also, exclusive statistical characteristics of PD signals or noise can be integrated into denoising techniques to effectively separate PD signals from noise. It is worth noticing that the simulated PD signals are ideal signals used in investigations as the attenuation of PD signals during its propagation in real situations are not taken into account. The differences aforementioned in statistical characteristics between PD signals and noise requires further investigations through real PD signals for demonstration. As a result, both simulated PD signals and real PD signals will be applied for this investigation in this thesis.

3 Partial Discharge Denoising

3.1 Introduction

A signal is defined as a function that conveys information regarding the state, the characteristics, the composition, and to name a few [30]. In other words, a signal is a means to convey information. Based on this, noise can be defined as a signal that conveys unwanted information. Noise is a very common phenomena that frequently takes place in the real world, interfering with the communication, measurement, or processing of an information-bearing signal. Figure 3-1 illustrates a general process of the development of a noisy signal in practice. It is inevitable for a signal of interest to suffer distortion during its transmission from its origin to a signal receiver or detector. Meanwhile, noise from surrounding environment is added to the distorted signal during its propagation. At the signal receiver or detector, the signal is further corrupted by the internal noise generated in the receiver or detector. The combination of the signal of interest and various types of noise ultimately presents as a noisy signal at the output of the signal receiver or detector. Numerous signal processing techniques, such as model-based and non-parametric signal processing, have been developed to recover a signal of interest from its associated noisy signal.

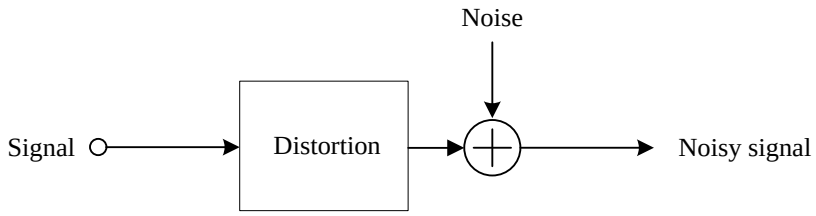

Figure 3-1 Illustration of a general process of the formation of a noisy signal

In relation to PD measurement in power cables, PD signals are normally attenuated or distorted due to the dielectric loss or the semi-conducting layer [8]. Environmental noise, i.e., discrete spectral interference (DSI), pulse-type noise, and white noise, can be coupled out with distorted PD signals by PD detectors. The magnitude of PD signals is generally very small in pC range [126]. The sensitivity of PD detection is therefore significantly reduced due

to noise. To increase the sensitivity, digital filters, matched filters, wavelet transform, and empirical mode decomposition are the frequently used signal processing techniques to suppress noise that affects PD measurement in past decades. It is necessary to mention that none of the aforementioned techniques is an omnipotent technique that can suppress all the types of noise.

To have a better understanding of these techniques, details regarding their investigations in PD denoising will be presented in the following subsections. Section 3.2 introduces digital filters with their applications and limits in PD measurement. Matched filters are presented in Section 3.3, which require the prior knowledge of PD waveforms. The effectiveness is highly dependent on the match of PD waveform and filter impulse response. Wavelet-based technique, the most prevalent method applied in the field of PD measurement, is discussed in Section 3.4. The advances of wavelet transform in PD denoising are presented as compared to traditional signal processing techniques, such as Fourier Transform (FT) and Short-time FT (STFT). Although wavelet-based denoising has been frequently used, it is often criticized for its non-adaptivity in the use of predefined mother wavelets. As such, Section 3.5 introduces an adaptive and data driven signal decomposition method, called empirical mode decomposition (EMD), and its associated denoising technique. Section 3.6 summarizes the application of all these techniques in PD measurement.

3.2 Digital Filters

Digital filtering is a very mature technique and has been widely applied in the fields, e.g., communication, speech enhancement, radar, and to name a few. As a result, a variety of structures and algorithms are available to design and realize digital filters. In relation to PD measurement, the spectral characteristics of periodic interferences, e.g., DSI, and PD signals are very different. DSI has a narrow-band spectrum while PD signals have a broad-band spectrum. Based on this difference, digital filters are primarily designed to reject DSI [118]. A filtering algorithm based on Fast Fourier Transform (FFT) was first proposed in [127] to reduce DSI in PD measurement in 1988. Nagesh and Gururaj [118] fully investigated the feasibility of digital filters and introduced a new digital filtering method based on a cascaded

2nd order infinite impulse response (IIR) lattice notch filter for rejection of DSI in 1993. Du [124] proposed an alternative technique using an adaptive notch filter to suppress DSI in 1994. Kopf and Feser [16] designed a digital finite impulse response (FIR) filter with a fixed order to reject DSI in 1995. Differences can be seen from these digital filtering applications, but the basic principle remains the same, i.e., a configurable digital filter is used as a multi-stopband filter to reduce the effect of DSI on PD detection, and thus, increase the signal to noise ratio of the measurement [16]. For a digital filter, it can be configured adaptively or non-adaptively. In a non-adaptive configuration, the filter coefficients of the digital filter are fixed regardless of the input signal. In contrast, the filter coefficients of an adaptive filter can automatically update based on the input data to reduce DSI. It is obvious that manual intervention is required for non-adaptive digital filters when the frequency ranges of periodic interferences change [15]. However, this is not the case for adaptive digital filters. Structure of a digital filter with an adaptive configuration is illustrated in Figure 3-2 [30]. In this system, $y(n)$ is the input noisy signal, $s(n)$ is the measured noise and processed by an adaptive digital filter to make it equal to the noise that residences in $y(n)$. The estimated noise $\hat{s}(n)$ is subtracted from $y(n)$ to cancel out the noise.

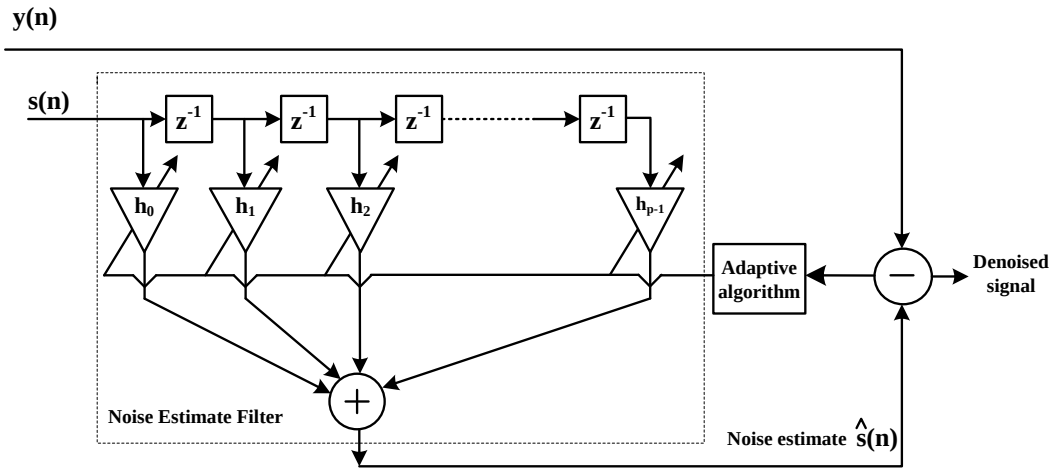

Figure 3-2 Configuration of an adaptive digital filter [30]

In digital filtering techniques, filters transform a detected PD signal from its time domain to frequency domain using FFT. In the frequency domain, the frequency ranges of DSI are set to zero, thus rejecting the DSI from the noisy signal. This denoising technique is based on the

knowledge of the signal bandwidth of PD signals and DSI. The application of digital filters is applicable to suppress narrow-band noise, but not suitable for the rejection of noise that has broad-band frequency as PD signals, e.g., pulse-type noise and white noise. Also, the involvement of FFT in signal transformation is one of the main reasons for the ineffectiveness of digital filtering techniques. It is known that FFT analyzes the signal in an infinite interval using trigonometric functions for signal decomposition. As a result, localized information (time instant) of the frequency components cannot be obtained through FFT (Further details regarding the drawbacks of Fourier Transformation in PD denoising will be introduced in Section 3.4). PD signals generally are transient, irregular, and non-periodic. A number of important pieces of information carried by PD signals cannot be unveiled by FFT. Based on this, the application of digital filters is limited for on-site PD measurement.

3.3 Matched Filters

The concept of Matched filter was first introduced in the 1940s for the use in radar detection. Matched Filters have become a standard technique for signal detection in the presence of noise [17]. For the application of Matched filters in the area of PD signal extraction, it was initially investigated in [87] and further discussed in [128].

Matched filters are a basic signal-processing tool for the extraction of a known waveform from a signal that has been interfered by noises. Figure 3-3 uses a typical digital filter as an example to explain the underlying principle of matched filters. Provided that $s(t)$ is the input signal and $n(t)$ is the noise that contaminates $s(t)$, the filter $h(t)$ that maximizes the signal-to-noise ratio (SNR) of the output signal $y(t)$ at sample time T_s is given by the filter $h(t) = s(T_s - t)$. It becomes clear by recognizing that filtering a waveform, which is referred to as $s(t)$ in Figure 3-3, is equivalent to calculating the cross correlation of the waveform and the reversed and time delayed filter impulse response [17]. If the reversed and time delayed impulse response equals the waveform itself (as shown in Figure 3-4), which is the case for a matched filter, then the filter output signal equals the autocorrelation function of the waveform. Details regarding the proof for the required filter response of a matched filter can be seen in [129].

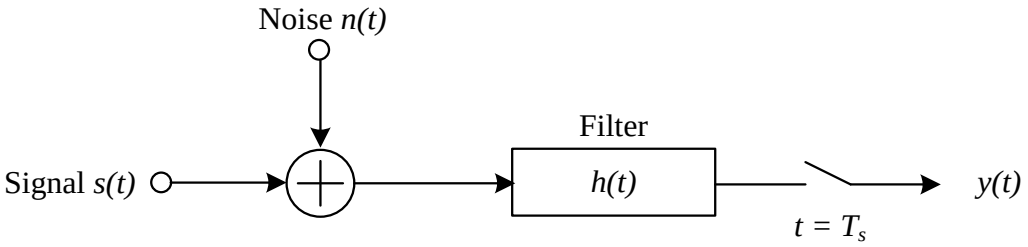

Figure 3-3 Basic filter

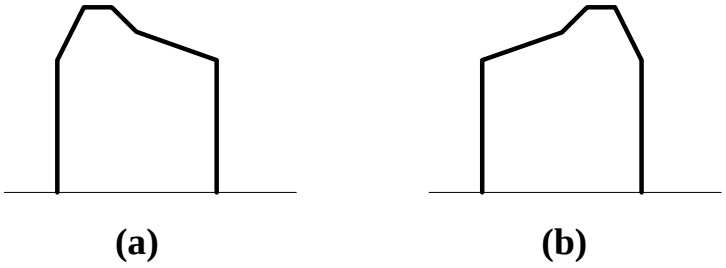

Figure 3-4 (a) Waveform of transmitted signal (b) The required filter impulse response of a matched filter

In [17], a semi-automated PD detection method was proposed using matched filter banks, which is shown in Figure 3-5. A range of matched filters with different impulse responses was employed in this matched filter banks. This configuration aims to improve the accuracy of the application of matched filters in PD denoising. The variation of impulse responses of matched filters is dependent on the PD pulse shapes, which can be derived from a PD pulse propagation model based on power cables. In this model, pulse width of PD waveform, frequency dependent dispersion, and phase response of detection circuit are taken into account. Wagenaars *et al.* [18] in 2011 proposed an adaptive algorithm to update the impulse responses of existing matched filters based on the detected PD pulses. Although this method introduced a new template for matched filter banks, adaptive process still depends on the initial PD pulse propagation model. It needs to be aware that the PD pulse propagation model mentioned above varies with the transmission parameters of power cables.

In theory, perfect match between the desired PD pulse shape and the matched filter can produce maximum effect. However, factors, e.g., variation in the PD pulse shapes due to dispersion and attenuation, can reduce the effectiveness of matched filters for their

applications in on-site PD measurement [130].

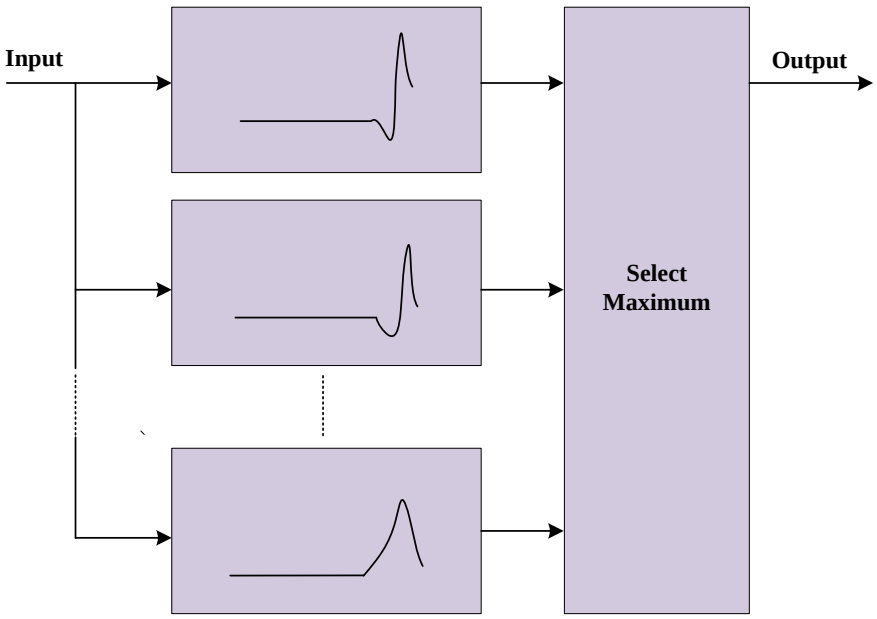

Figure 3-5 Example of a partial discharge matched filter bank [17]

3.4 Wavelet-based Technique

Signal representation plays a critical role in signal processing. A good signal representation is desirable as it is beneficial for a series of signal processing procedures, such as analysis, noise filtering, and compression [131]. Most of the signals are generally time-domain signals in practice. The underlying idea of signal representation is certain aspect of a signal that is not readily available in its time domain can be highlighted, and thus, the information of interest carried by this signal can be easily obtained. Fourier transform (FT), as mentioned before, is such a representation that frequency information of time-domain signals can be represented through it. To represent local characters of frequency information, FT is applied to windowed signals. This is referred to as Short Time Fourier Transform (STFT). Although STFT is the first approach that can simultaneously represent the time and frequency information of a signal, the time resolution is determined by the length of the window. In recent decades, the prevalence of wavelet transform reveals that it is a more robust technique that can simultaneously represent the time and frequency information of a signal than STFT. Based on

this, it is necessary to have a brief introduction of FT and STFT before wavelet transform.

3.4.1 Fourier Transform (FT)

The history of FT can date back to early 19th century, a French mathematician J. Fourier discovered that a periodic function can be represented as an infinite sum of sine and cosine functions [20], [132]. This discovery has been expanded to non-periodic functions, periodic discrete and non-periodic discrete functions many years later. In 1965, the invention of a new algorithm called Fast Fourier Transform (FFT) broadened the popularity of FT as a tool in signal processing.

As mentioned before, a signal can be decomposed by FT through a series of sine and cosine functions at an infinite interval. Given a time-domain signal $x(t)$, its frequency representation $X(w)$ through FT can be calculated based on the following equation,

$$X(w) = \int_{-\infty}^{\infty} x(t) \cdot e^{-j\omega t}\, dt \tag{3-1}$$

Equally, if $X(w)$ is known, an inverse FT can be applied to obtain $x(t)$ using the equation below,

$$x(t) = \int_{-\infty}^{\infty} X(w) \cdot e^{j\omega t}\, dw \tag{3-2}$$

The definition of FT in (3-1) further indicates that the transformation is implemented through an infinite time interval. The time information of any frequency component in $X(w)$ is not available. In other words, no matter a frequency component occurs in any time instant t_1 or t_2, it has the same effect on $X(w)$. It is the reason that FT is not suitable for the analysis of signals that have time varying frequency. Figure 3-6 delineates the frequency representations of two different time-domain signals. It can be seen that difference between (a) and (c) in the time domain has no effect on their frequency representations in (b) and (d). A few oscillations around the frequency components in (d) is due to the sudden change between frequencies. Based on this, only frequency information of a signal can be easily obtained through FT. If concerns of the local characters of frequency components raised, i.e., requirement of a time-frequency representation of a signal, FT is not a suitable tool for signal representation.

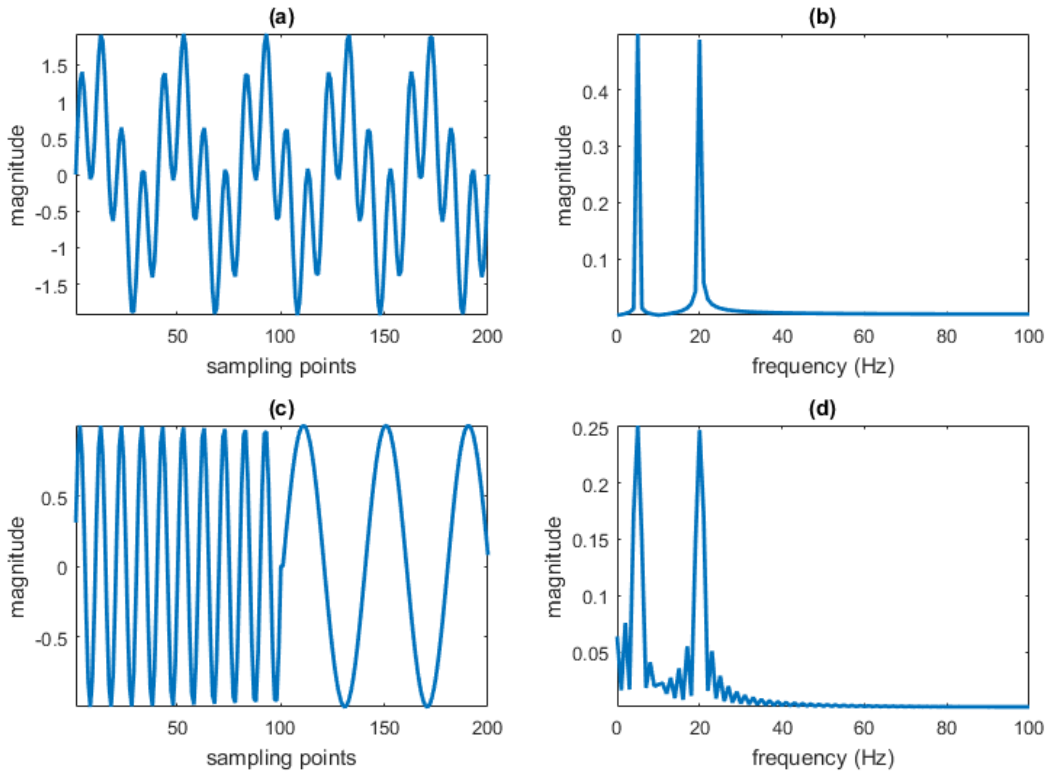

Figure 3-6 (a) A time-domain signal with the summation of a 5Hz and 20Hz sine waves, (c) A time-varying signal, (b) and (d) are the frequency representations of (a) and (c), respectively

3.4.2 Short-Time Fourier Transform (STFT)

STFT, introduced by Dennis Gabor in 1940s, is a revised version of FT with a minor difference. However, this difference in STFT enables it to represent a signal in both time and frequency domain. In STFT, a time-domain signal $x(t)$ is portioned into many small enough segments such that these segments can be assumed stationary [20], [133]. A fixed window function is adopted for this purpose and its width is required to be equal to the signal segments where its stationarity is valid. The window function is applied to the beginning of the signal $x(t)$, and then shifted to the end of $x(t)$ with equal intervals. Figure 3-7 illustrates the procedures of the involvement of a window function in STFT. As assumed to be stationary, FT can be applied to these windowed signals to obtain their frequency information. The result of STFT of the signal $x(t)$ can be calculated by the given expression as follows,

$$STFT_f(\tau, w) = \int_{-\infty}^{\infty} x(t) \cdot w^*(t-\tau) \cdot e^{-jwt} dt \qquad (3\text{-}3)$$

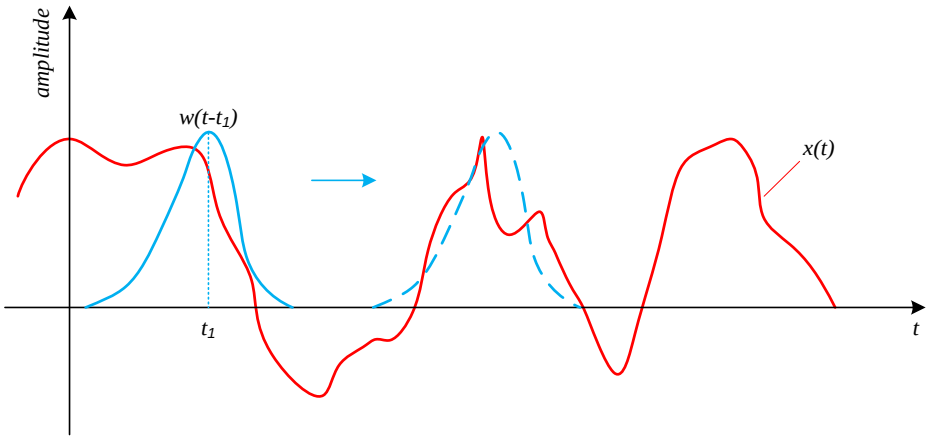

Figure 3-7 Illustration of the application of a window function in STFT

(3-3) indicates that STFT of $x(t)$ is a measure of similarities between the signal and the selected window function, and decomposes the signal into a set of frequency bands at any given time as well, e.g., t_1 in Figure 3-7. STFT thus can represent a signal simultaneously in time and frequency domain.

Although STFT can map a signal in the time-frequency plane, it has its inherent resolution problem due to the Heisenberg Uncertainty Principle [134]–[136]. That is, it is impossible to obtain precise localization in both time and frequency domain. In STFT, time and frequency resolution are dictated by the width of the selected window. With a pre-defined window width, the STFT gives a fixed resolution in both time and frequency over the entire period of signal analysis. Narrow windows can improve time resolution at the cost of frequency resolution, while wide windows can improve frequency resolution at the cost of time resolution. For aperiodic and transient signals, they cannot be easily characterized by the STFT using continuous, periodic basis functions [22].

3.4.3 Wavelet Transform

3.4.3.1 Wavelet Theory

Wavelet transform (WT), as its name implies, is a wavelet-based expansion or decomposition of signals. It is developed as an alternative to STFT for the analysis of non-stationary features of a signal and hence it can map the signal in the time-scale plane. Different from Short-time Fourier

Transform (STFT), applies the Fourier Transform to a fixed window of a signal, the wavelet used in WT can be scaled to analyse the localized features of a signal at various time or frequency resolutions. The scaling operation of wavelet transform is visually depicted in Figure 3-8 [137].

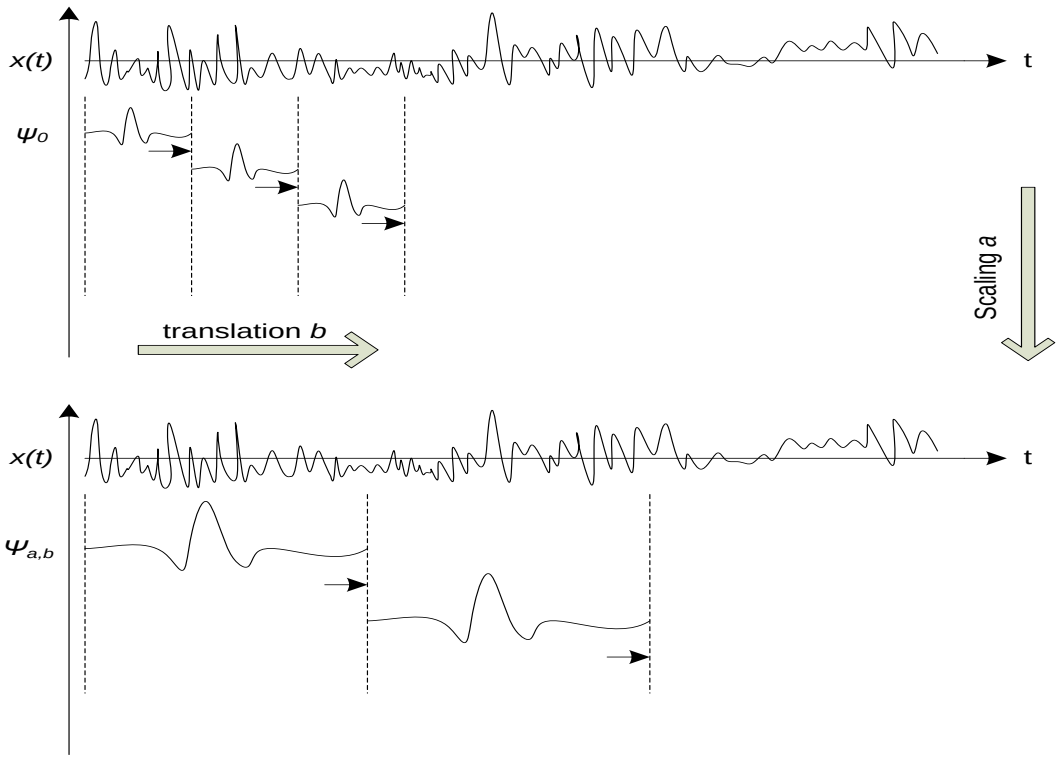

Figure 3-8 The scaling and translating operation of wavelet transform. *x(t)* is the signal under analysis, ψ_0 is mother wavelet, $\psi_{a,b}$ is the daughter wavelet scaled by *a* and translated by *b* [137]

The process of a wavelet-based expansion or decomposition can be explicated with the amid of Figure 3-8. The performance of WT is started from a selection of a prototype, or mother wavelet, ψ_0, which is then dilated or shrunk with an unaltered wavelet shape. At some scale, ψ_0 is scaled by *a* and translated by *b* to be a daughter wavelet, $\psi_{a,b}$. The relationship between ψ_0 and $\psi_{a,b}$ can be expressed as [131]

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi_0\left(\frac{t-b}{a}\right) \tag{3-4}$$

where $1/\sqrt{a}$ is a normalized factor. The daughter wavelet $\psi_{a,b}$ is then shifted along the entire signal $x(t)$, and correlated with $x(t)$. The procedure of such shift and correlation is iteratively executed through all possible scales and around each temporal translation, obtaining a matrix of correlation coefficients that indicate the similarity between $x(t)$ and the daughter wavelets.

The concept of scale replaces the concept of frequency in wavelet spectral analysis. Typically,

scale is a duration of wavelet and for many wavelets is related to the reciprocal of frequency [137]. This relationship between scale and frequency can be easily seen from Figure 3-9. In Figure 3-9 (a), b is assigned to be 0, and a is assigned to be 1, 2, and 4 respectively to obtain 3 different daughter wavelets based on (3-4). Note that $a = 1$ and $b = 0$ indicate that the mother wavelet is not scaled and translated. The corresponding FT of these daughter wavelets are illustrated in Figure 3-9 (b). It can be seen that the central frequency of the daughter wavelets decreases with the increase of scale (from 1 to 4). It also demonstrates the underlying mechanism of wavelet transform that can give time-frequency localizations of a signal through scaling operations.

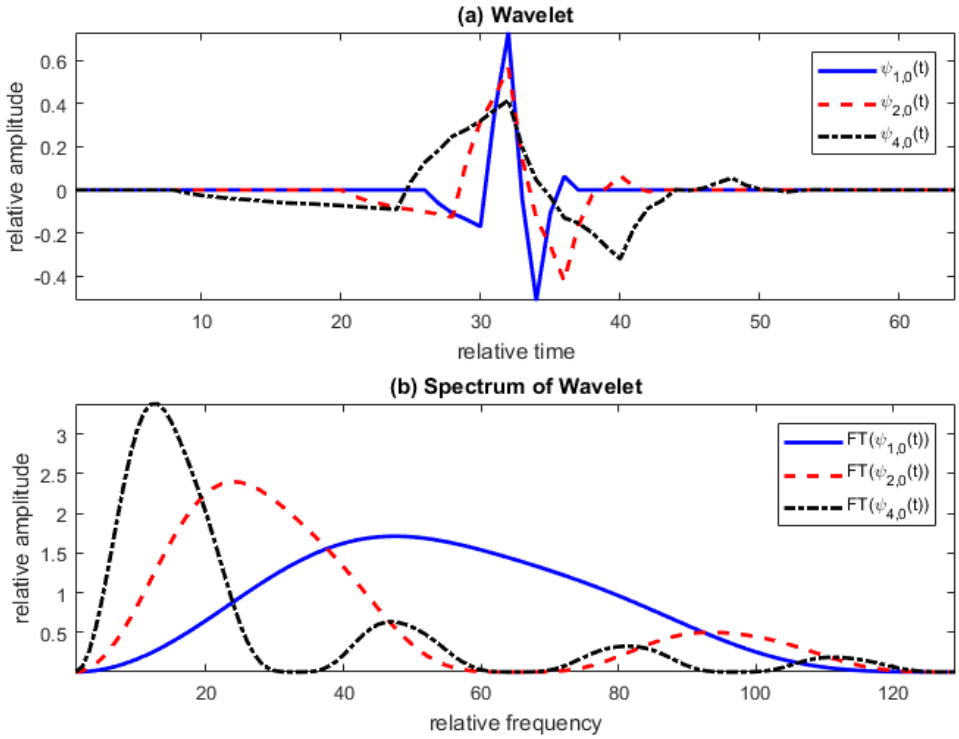

Figure 3-9 Scale is the reverse of frequency. (a) daughter wavelets with scale a = 1,2, and 4 respectively, and translation b = 0; (b) the resulting spectrum of these daughter wavelets in (a)

Generally, wavelet transform can be achieved through the application of continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The CWT of a signal, $x(t)$, is a linear operation defined as [137]

$$W_x^{\psi}(a,b) = \int_{-\infty}^{\infty} x(t) \cdot {\psi_0}^*\left(\frac{t-b}{a}\right)dt \tag{3-5}$$

where $W_x^{\psi}(a,b)$ is the matrix of correlation values (wavelet coefficients) and asterisk (*)

denotes the conjugate of the scaled and translated mother wavelet. The CWT of a signal, as mentioned above, calculates the wavelet coefficients at all possible scales and around each temporal translation. This fully calculation using CWT, however, can lead to a remarkable redundancy in signal representation. It is desirable that wavelet-based techniques applied for signal compression and denoising can provide the economy in signal representation [137]. With the employment of orthogonal basis functions, DWT is introduced to reduce the redundancy in signal representation and enable the possibility of perfect signal reconstruction using inverse DWT (IDWT). In DWT, scale a and translation b are replaced by 2^{-j} and $2^{-j} \cdot i$ respectively. Substituting these new parameters into (3-4) and (3-5) yields

$$\psi_{j,i}(t) = 2^{j/2}\psi_0(2^j \cdot t - i) \tag{3-6}$$

$$x(t) = \sum_i \sum_j W_{j,i} \cdot \psi_{j,i}(t) \tag{3-7}$$

Note that (3-7) is actually the definition of IDWT, $\{W_{j,i}\}$ is the matrix of correlation values, which is also called the DWT of $x(t)$. j and i in (3-6) and (3-7) denote the scale index and the translation index respectively. The wavelet ψ_0 used in DWT has a stricter constraint than that in CWT: it should be orthogonal or biorthogonal. This constraint can guarantee the transformed signal $x(t)$ to be sufficiently represented by the coefficients $\{W_{j,i}\}$ with no redundancy, and thus, the information embedded in $\{W_{j,i}\}$ is adequate to reconstruct $x(t)$ with good accuracy [137]. Based on this, DWT is preferable in the application of signal denoising.

Figure 3-10 shows the implementation of DWT in signal expansion, which is in agreement with the analysis above. A signal is convolved by a low-pass filter $\boldsymbol{h}$ and a high-pass filter $\boldsymbol{g}$, followed by a downsampling operation by 2, generating approximation and detail coefficients, respectively. The downsampling operation constitutes a critical step in DWT. Generally, it is a process that removes the odd-numbered entries of the data set and generates a new data set with half length. Figure 3-11 illustrates the downsampling operation in DWT, it demonstrates that the length of the data set is halved repeatedly. Approximation coefficients, as the output of the low-pass filter, is then iteratively filtered by the low- and high-pass filters until the predefined scale J reaches. Note that the maximum decomposition scale J_{max} is defined as $log_2 N$, where N is the length of the input signal. The results of DWT, as shown in Figure 3-10, is composed of one set of

approximation coefficients and J-set of detail coefficients. The reconstruction of the input signal, i.e., inverse DWT (IDWT), is a reverse operation as shown in Figure 3-10. Instead of downsampling in DWT, upsampling is involved in IDWT. Figure 3-12 delineates the processes of IDWT for signal reconstruction. To obtain perfect signal reconstruction, the low- and high-pass filters are designed as quadrature mirror filters (QMFs).

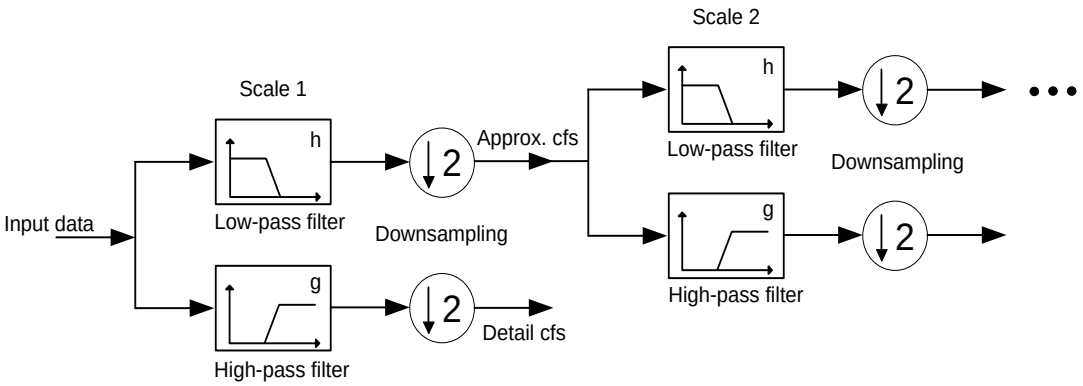

Figure 3-10 The implementation of DWT in signal decomposition

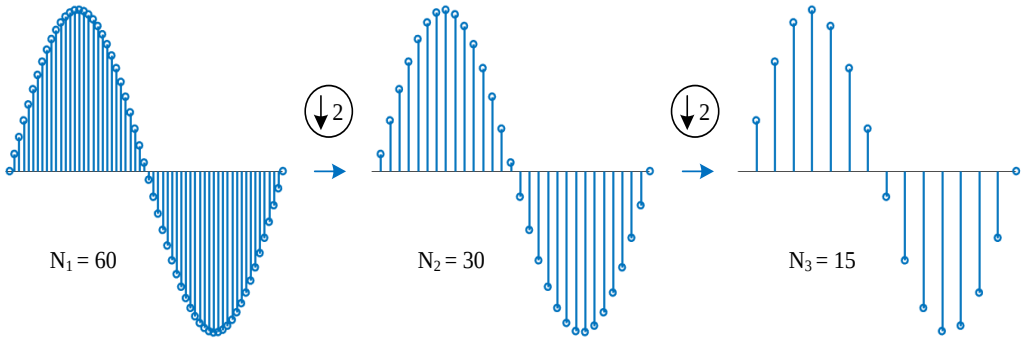

Figure 3-11 Illustration of downsampling operation in DWT

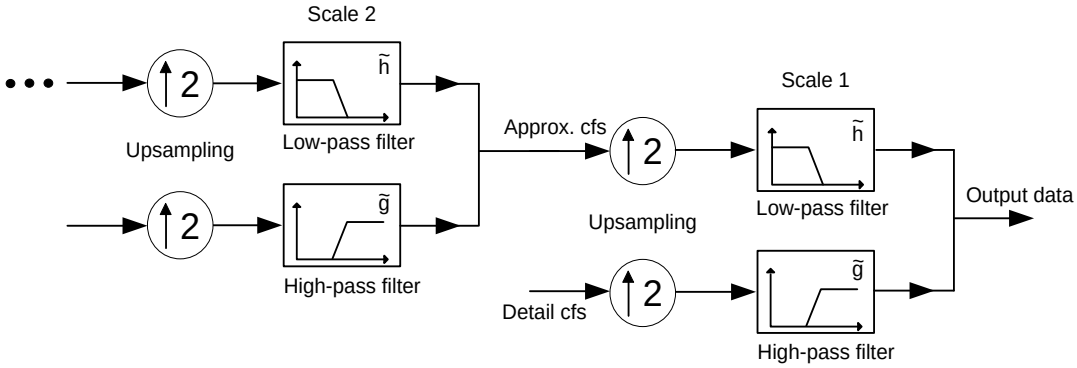

Figure 3-12 The implementation of IDWT in signal reconstruction

As shown in Figure 3-9 (b), the scaled wavelet can be interpreted as band-pass filters in the frequency domain. WT is therefore equivalent to filtering a signal in different subbands by the filter pairs, e.g., the low-pass filter h and the high-pass filter g, and each subband represents the signal in a different resolution. Ideally, these filter pairs halve the frequency band with the increase of scale. Let f_s be the sampling frequency of the input signal, the frequency band, $G_1(w)$, of the output of the high-pass filter is $f_s/4 - f_s/2$, while the frequency band, $H_1(w)$, of the output of the low-pass filter is $0 - f_s/4$. For next scale, $H_1(w)$ is further split into $G_2(w)$ and $H_2(w)$, which are $f_s/8 - f_s/4$ and $0 - f_s/8$, respectively. The frequency band is iteratively halved in the subsequent decomposition in the same manner until the predefined scale reaches. For a J-scale DWT, the distribution of the corresponding frequency bands is shown in Figure 3-13. It can be seen that the frequency band of low-pass filter is $0 - f_s/2^{J+1}$ and the frequency band of high-pass filters is $f_s/2^{J+1} - f_s/2$ for a J-scale decomposition.

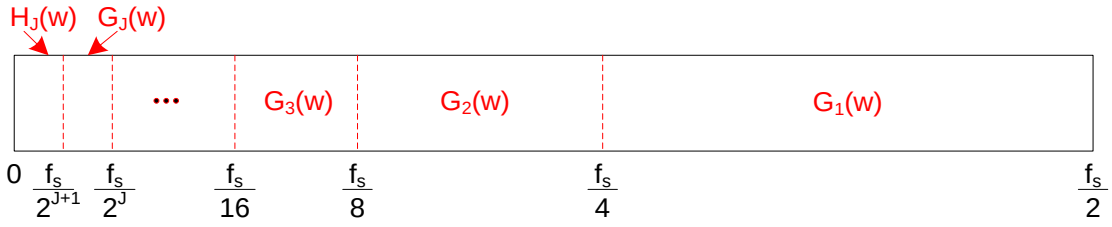

Figure 3-13 The frequency bands of filters at each decomposition scale

3.4.3.2 Wavelet-based Denoising

The wavelet-based denoising theory is dependent on the fundamental idea that the energy of a signal is often concentrated in only a few coefficients while the energy of noise is widely spread among all the coefficients in the wavelet domain [22], [138], [139]. General procedures for the wavelet-based denoising of a signal are as follows,

1) Select a suitable wavelet from a pre-defined wavelet library.
2) Apply DWT to decompose a noisy signal s with a selected wavelet to a predefined scale J, and obtain approximation coefficients a_J at the final scale J and detail coefficients d_j at decomposition scale j, where $j = 1, 2, ..., J$.

3) Estimate the threshold through a noise estimation technique and apply this threshold to detail coefficients, d_j, at decomposition scale j using hard or soft thresholding function.

4) Apply IDWT to the approximation coefficients a_J and the thresholded detail coefficients d_j' to reconstruct the denoised signal s'.

Based on the noise estimation technique proposed in [22], [140], the scale-dependent threshold used in this thesis , if not specified, is estimated by

$$thr_j = \frac{MAD|d_j|}{0.6745}\sqrt{2\log(n_j)} \tag{3-8}$$

where $MAD|\cdot|$ is the median absolute deviation of the detail coefficients d_j at decomposition scale j, and n_j is the length of d_j. It is worth mentioning that the threshold estimated by (3-8) is effective in the suppression of white noise, while is not effective in the reduction of pulse-type noise.

For the thresholding scheme, two thresholding functions are normally applied in wavelet-based denoising, i.e., hard thresholding and soft thresholding. The mathematical expressions for these two functions are given by

Hard thresholding:

$$d'_{j,i} = \begin{cases} d_{j,i} & if\ |d_{j,i}| > thr_j \\ 0 & otherwise \end{cases} \tag{3-9}$$

Soft thresholding:

$$d'_{j,i} = \begin{cases} sgn(d_{j,i})(|d_{j,i}| - thr_j) & if\ |d_{j,i}| > thr_j \\ 0 & if\ |d_{j,i}| \le thr_j \end{cases} \tag{3-10}$$

where $i = 1,2,\dots,n_j$.

Difference of these two thresholding functions in (3-9) and (3-10) can also be seen from Figure 3-14. Note that soft thresholding is used in this thesis, if not specified.

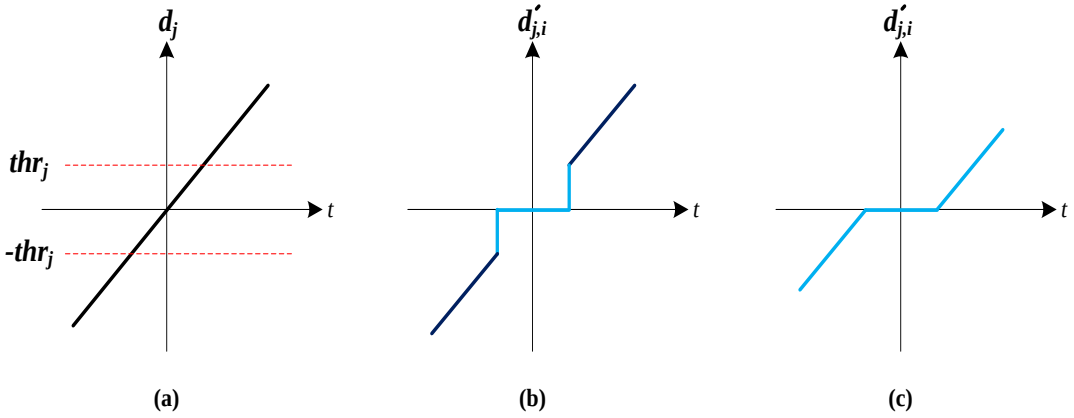

Figure 3-14 (a) the original detail coefficients d_j, (b) hard thresholding, (c) soft thresholding

3.4.3.3 Issues in the Implementation of Wavelet-based Denoising

From the analysis of wavelet theory and wavelet-based denoising, the implementation of wavelet-based technique in signal denoising is highly correlated to the issues, such as mother wavelet selection, decomposition scale determination, and noise estimations. These three aspects determine the effectiveness of wavelet-based denoising. During the past decades, numerous techniques have been proposed to optimize the application of wavelet-based denoising through these aspects. In relation to PD denoising, improvement of wavelet-based technique can be obtained through further analysis of the statistical characteristics of PD signals and the associated noise. As the research objectives in this thesis, the choice of mother wavelet, selection of an appropriate decomposition scale, and effective noise estimation will be fully discussed with newly proposed algorithms in Chapter 4 - 6, respectively.

3.5 Empirical Mode Decomposition (EMD)

3.5.1 EMD Fundamentals

Empirical mode decomposition (EMD), introduced by Huang *et al.* [141] in 1998, is a nonlinear technique for non-stationary signal analysis and representation [141]–[143]. Since the emergence of EMD, it has been intensively investigated in non-stationary signal processing due to its major advantage for no requirements of base functions. The base functions of the EMD are derived from the signal itself, which is the fundamental difference

between WT and EMD [144]. Accordingly, EMD is known as an adaptive and fully data-driven method for signal decomposition. EMD is merely an algorithm without admitting an analytical formulation that allows for theoretical analysis and performance evaluation, although it has been proven remarkably effective in signal decomposition [142]. The lack of theoretical frame may impose difficulties on the interpretation of the transformed signals in some sense.

The essence of the EMD is that any signal is assumed to have a large number of simple intrinsic modes of oscillation. In turn, a signal can be effectively decomposed into a series of intrinsic mode functions (IMFs). The IMFs are extracted through a sifting process in EMD, and each IMF has a distinct time scale [138], [141]. To be successfully decomposed into IMFs, a signal must have at least two extrema, i.e., one maximum and one minimum. For an IMF, it has the following properties: 1) zero mean, that is, the mean value of the envelop defined by the local maxima and the envelope defined by the local minima is zero, 2) all the maxima and all the minima will be positive and negative, respectively, which guarantees that a single zero-crossing point exists in any pair of two adjacent maxima and minima [139]. The upper and lower envelopes are produced by the use of a cubic spline interpolation to connect the local maxima and minima, respectively. Figure 3-15 illustrates the formation of these two envelopes using a cubic spline interpolation in EMD. The red dash line is the upper envelope and the blue dash line is the lower envelope in Figure 3-15.

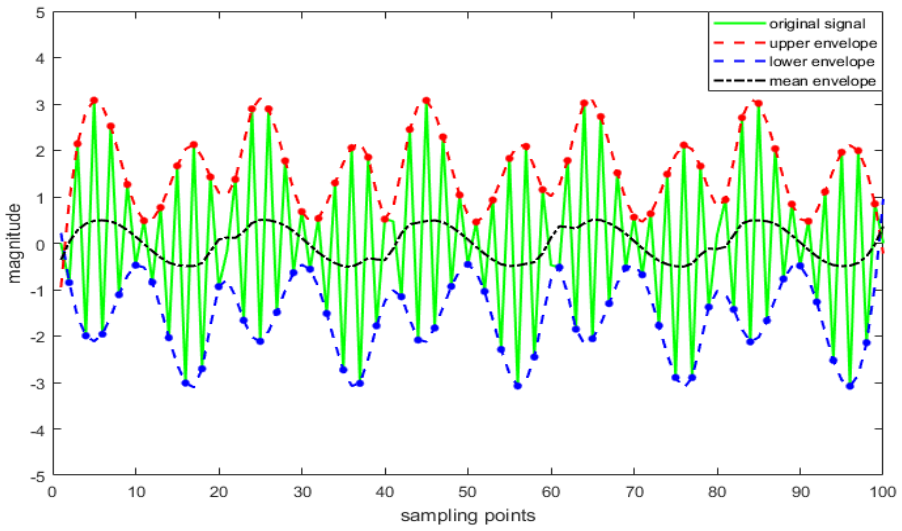

Figure 3-15 The formation of upper and lower envelopes using a cubic spline line in EMD

A sifting process used to extract the IMFs from a signal serves two purposes: to eliminate riding waves and to make the wave-profiles more symmetric [141]. The procedures of this process to extract an IMF is detailed as follows [98], [138], [141], [145]–[151],

1. Locate the extrema of a signal $x(t)$, $x_{max}(t)$ and $x_{min}(t)$ (red and blue dots in Figure 3-15),
2. Apply a cubic spline interpolation to produce the upper and lower envelopes, $e_{max}(t)$ and $e_{min}(t)$ (red and blue dash line in Figure 3-15),
3. Calculate the mean of the two envelopes, $m_1 = (e_{max}(t) + e_{min}(t))/2$ (black dash line in Figure 3-15),
4. Subtract the mean from the original signal $x(t)$ to obtain the first potential IMF, $h_1 = x(t) - m_1$,
5. Check if h_1 is zero-mean signal with the number of maxima and minima differing at most by one, if yes, h_1 is remained as the first IMF,
6. Repeat the process by replacing $x(t)$ with h_1, if it is not the case in step 5
7. A stopping criterion for the sifting process to extract an IMF is used to guarantee that the IMF components retain enough physical sense of both amplitude and frequency modulation. The standard deviation (SD) of the signal, calculated by the following equation (3-11), is often used as the stopping criteria. The sifting process stops when the SD reaches 0.2,

$$SD = \sum_{t=0}^{T} \left(\frac{|h_{k-1}(t) - h_k(t)|^2}{h_{k-1}^2(t)}\right) \tag{3-11}$$

8. When the stopping criterion is satisfied, the remaining component of the signal is known as the residual. For residual, it is used to extract IMFs from the original signal repeatedly. This extraction of IMFs will not stop until the residual r_n in equation (3-12) is less than a predetermined small value or the residual r_n becomes a monotonic function from which no more IMF can be extracted. The signal can then be expressed in equation (3-12) after *n* iterations.

$$x(t) = \sum_{k=1}^{n} h_k(t) + r_n(t) \tag{3-12}$$

Figure 3-16 is the flow chart that can clearly illustrate the procedures of the sifting process used to extract an IMF from an arbitrary signal [98], [143]. The objective of the EMD is to use these procedures in the original signal to extract the first IMF and then to iterate on the residual components of the signal to extract more IMFs. It allows the identification of various oscillatory modes within the signal [143]. An example of EMD applied to a noisy signal is delineated in Figure 3-17. As discussed above, the signal with 2048 sampling points is decomposed into 9 IMFs and the residual.

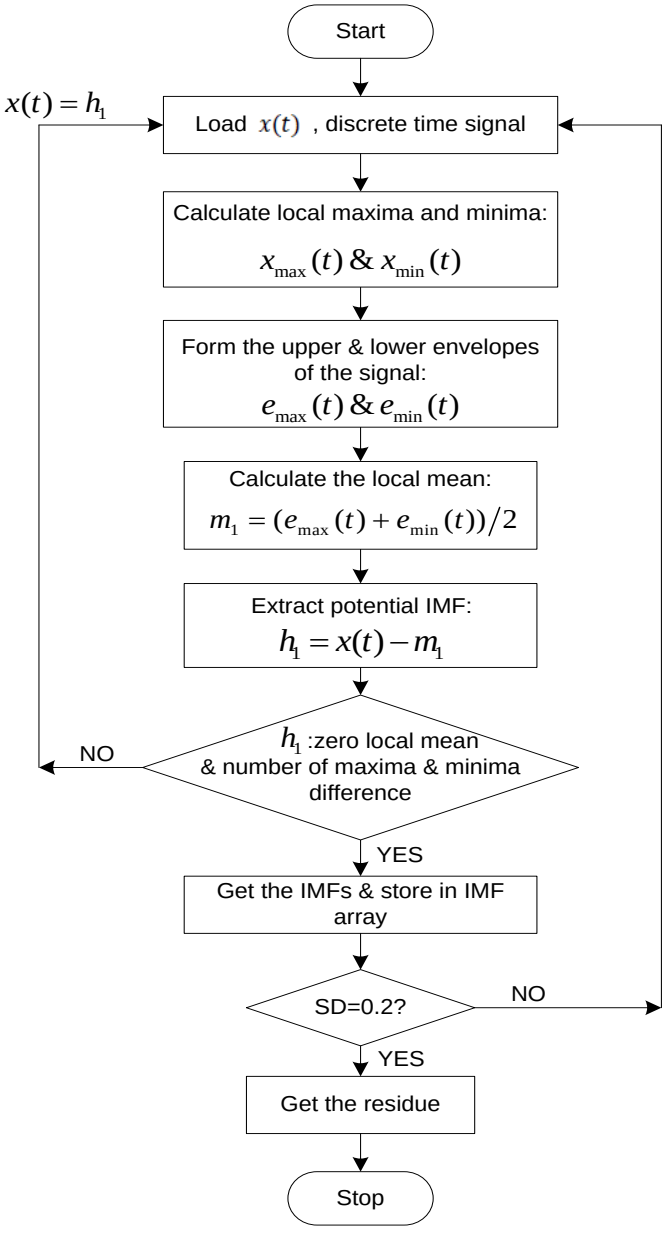

Figure 3-16 Flow chart of the sifting process in the EMD

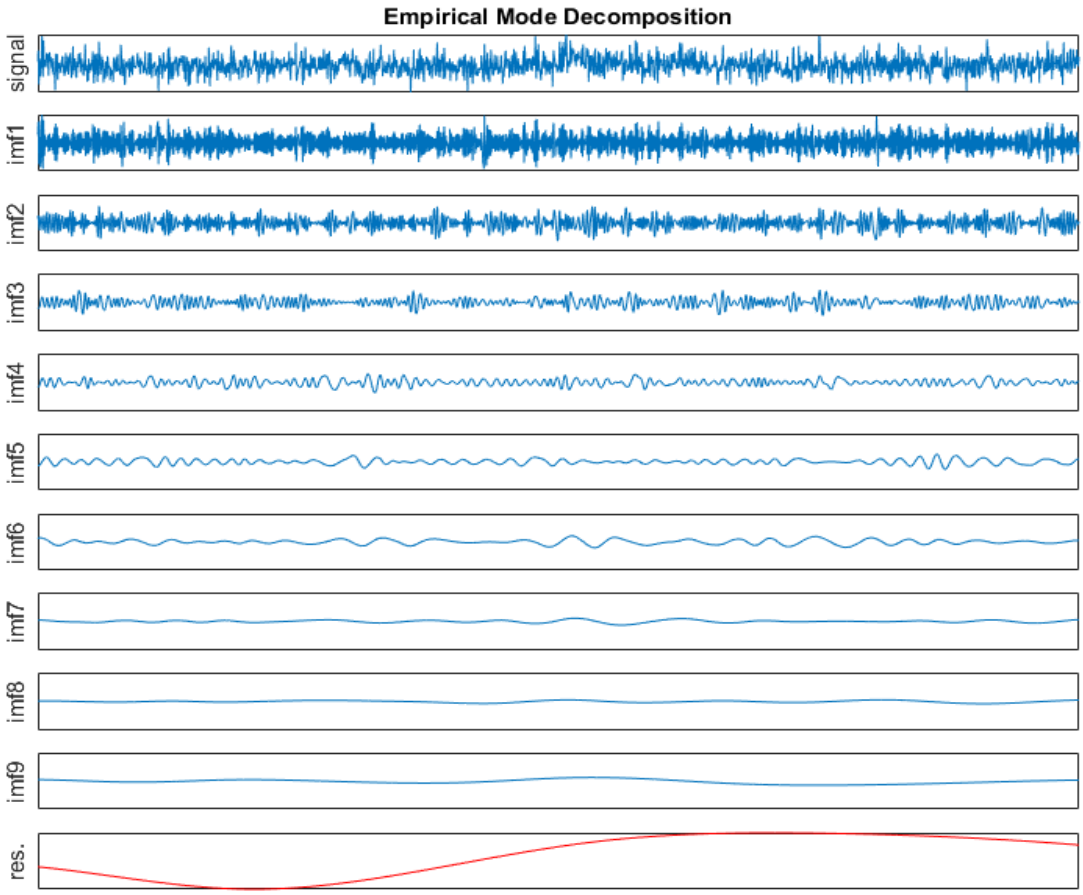

Figure 3-17 EMD of a noisy signal with 2048 sampling points

3.5.2 EMD-based Denoising

3.5.2.1 EMD Traditional Denoising

EMD-based denoising was initially investigated to understand if a specific IMF extracted from a signal contained useful information or was primarily noise [139]. Flandrin *et al.* [152] and Wu *et al.* [153] developed the significance IMF test procedures using a signal contaminated only by fractional Gaussian noise and white Gaussian noise, respectively. It is necessary to mention that fractional Gaussian noise in discrete time corresponds to a time series $\{y_H[n], n = \cdots, -1,0,1, \ldots\}$ indexed by a real-valued parameter $0 < H < 1$, where H refers to Hurst exponent of this time series [152]. It is known that white noise is the special case of this time series when $H = 1/2$ [152]. The significance test procedure is proposed based on the statistical analysis of modes resulting from the decomposition of these noisy

signals [139]. In both investigations, noise that interferes the signal of interest is known, i.e., fractional Gaussian noise or white Gaussian noise, and thus, the energy of the IMFs extracted from noise-only signals can be obtained through EMD. Equally, the EMD can be performed on the noisy signal to obtain the energy of its IMFs. Comparison of the IMF energy between the noise and the associated noisy signal can indicate the presence of useful information [139]. In a denoising scenario, the signal of interest can be reconstructed by the IMFs that only carry useful information and those IMFs that contain primarily noise are discarded. To differentiate from other EMD-based denoising methods, this method discussed above is termed EMD traditional denoising (EMD-TR), i.e., IMFs dominated by noise are set to zeros. In short, EMD-TR is mainly based on a noise model such that the difference of energy distribution of each IMF of the noise and the associated noisy signal can be highlighted. Figure 3-18 shows this difference based on a signal only corrupted by white noise. The original noisy signal and the noise only are shown on the top-right in Figure 3-18, along with their energy distributions of IMFs. It can be seen that the energy of IMFs differs from the 5^{th} IMF, which indicates that noise is mainly carried by IMF1-4 and useful information of the signal is primarily concentrated on IMF5-10. The noise-only signal is generally not the case in the real world for the application of EMD. As such, the finding of the mode boundary that can distinguish if the IMFs carry signal information or not becomes difficult. As one of the research objectives in this thesis, this will be discussed with a novel method for relative mode selection (RMS) in Chapter 8.

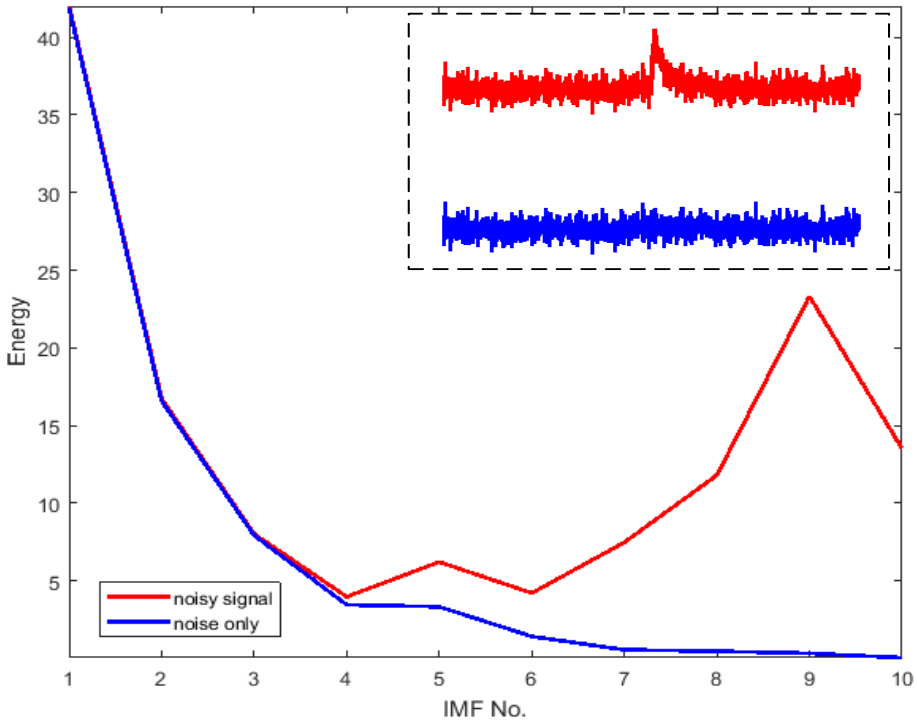

Figure 3-18 Energy distribution of IMFs from white noise and a signal corrupted by white noise (top-right: original noisy signal and noise only)

3.5.2.2 EMD Direct Thresholding

Signal reconstruction using direct summation of the IMFs that contain signal information in EMD-TR is not strictly true, since noise and signal of interest are generally mixed within each IMF. An alternative EMD-based denoising method was inspired by noise estimation techniques adopted in wavelet-based denoising. A noise threshold is estimated using (3-13) and then directly applied to each IMF through soft or hard thresholding function (see (3-15) and (3-16)) in [144], [154], [155].

$$thr_j = \frac{MAD|h_j|}{0.6745} \cdot \sqrt{2logn} \tag{3-13}$$

$$\sigma_j = \frac{MAD|h_j|}{0.6745} \tag{3-14}$$

where h_j denotes the j^{th} IMF, n denotes the length of original signal, and σ_j is the estimation of noise level of the j^{th} IMF.

Hard thresholding is defined by

$$h_j' = \begin{cases} h_j(i) & if\ |h_j(i)| > thr_j \\ 0 & otherwise \end{cases} \quad (3\text{-}15)$$

Soft thresholding is given by

$$h_j' = \begin{cases} sgn(h_j(i))(|h_j(i)| - thr_j) & if\ |h_j(i)| > thr_j \\ 0 & if\ |h_j(i)| \le thr_j \end{cases} \quad (3\text{-}16)$$

where i denotes the i^{th} element in the j^{th} IMF, and $i = 1, 2, ..., n$.

This alternative method is termed EMD direct thresholding (EMD-DT) in this thesis. However, the direct application of wavelet-like thresholding to IMFs can cause discontinuity in the reconstructed signals. Figure 3-19 illustrates this unwanted result of EMD-DT. The drawback of EMD-DT is obviously shown in Figure 3-19 (d), i.e., the reconstructed signal suffers severe signal discontinuity. In Figure 3-19 (c), it is the denoising results of EMD-TR, which is in agreement with the statement above that noise is remained in the reconstructed signal and can be seen through the oscillations at both sides of the main pulse.

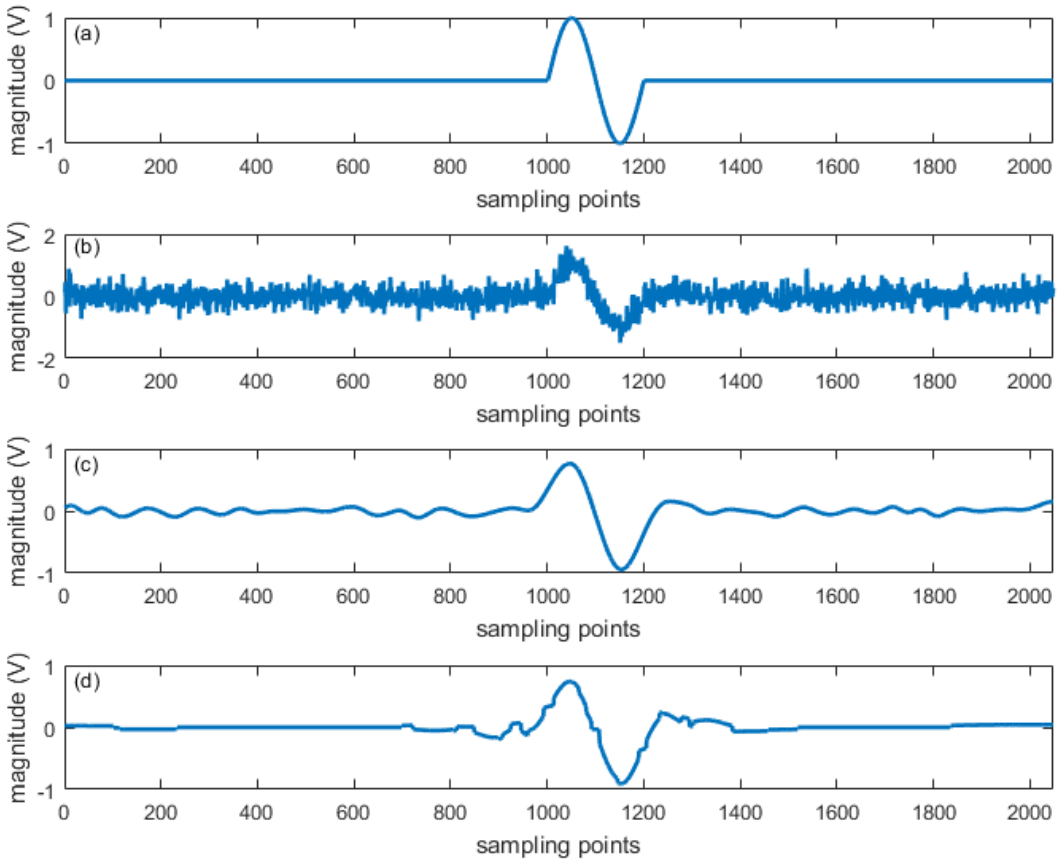

Figure 3-19 EMD-based denoising results, (a) original signal, (b) noisy signal, (c) EMD-TR, (d) EMD-DT

3.5.2.3 EMD Interval Thresholding

To overcome the drawback of EMD-TR, an improved method, termed EMD interval thresholding (EMD-IT), was proposed in [139]. This interval thresholding was developed based on the characteristics of an IMF, i.e., zero mean and the number difference between zero-crossings and extrema must be zero or at most by one. The definition of EMD-IT is that if the absolute value of an extrema is higher than the threshold, the part of an IMF between its two adjacent zero-crossings will remain, or this part will be discarded. The understanding of the definition of EMD-IT can benefit from Figure 3-20, which illustrates an IMF with highlighted maxima (red dots), minima (blue dots), and zero-crossings (black dots). The zero-crossings positioned at the time instances z_{i-1} and z_i, maxima positioned at the time instance m_i, and minima positioned at the time instance n_i are marked as well. Given that this IMF is $h_j(t)$, the part of $h_j(t)$ from $h_j(z_{i-1})$ to $h_j(z_i)$ (red circle in Figure 3-20) will be remained if $\left|h_j(m_i)\right|$ is higher than the threshold, or this part will be discarded if $\left|h_j(m_i)\right|$ is less than the threshold. Similar as wavelet thresholding and EMD-DT, the IMFs can be thresholded through hard or soft thresholding functions. It is worth noting the thresholding functions mentioned here are slightly different from (3-15) and (3-16). The thresholding functions used in EMD-IT are given by equation (12) and (13) in [139]. Difference between EMD-DT and EMD-IT can be seen from the results of the thresholded IMF in Figure 3-21, where the threshold is set to 0.05 and hard thresholding function is applied. Figure 3-21 (a) illustrates the thresholded IMF using EMD-DT and EMD-IT. Figure 3-21 (b) - (d) focus on the extrema of the IMF higher than the threshold and show the difference of the application of thresholding on the IMF between EMD-DT and EMD-IT. The results clearly show that the discontinuity in thresholded IMF caused by EMD-DT is not the case in EMD-IT. The noisy signal, as shown in Figure 3-19 (b), is still used as an example to compare the denoising results among EMD-TR, EMD-DT, and EMD-IT. Results are illustrated in Figure 3-22. From the direct observation in Figure 3-22, EMD-IT can obtain a smoother denoised signal than the others.

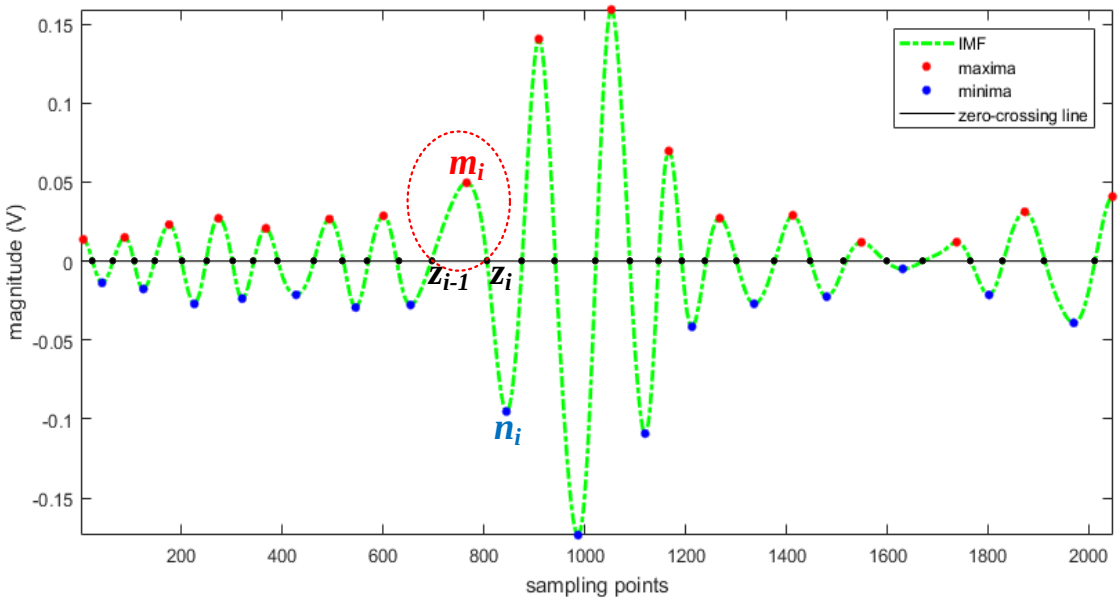

Figure 3-20 An IMF with highlighted maxima, minima, and zero-crossings

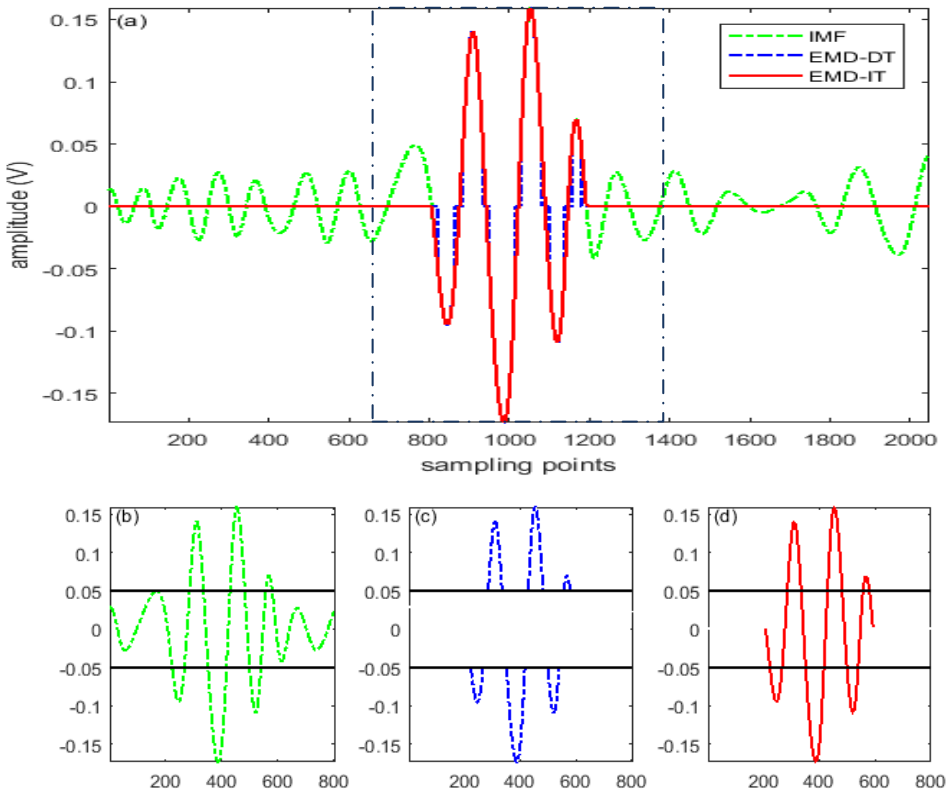

Figure 3-21 (a) Difference between EMD-DT and EMD-IT, (b) – (d) Separation of dash-square area in (a) to highlight the difference

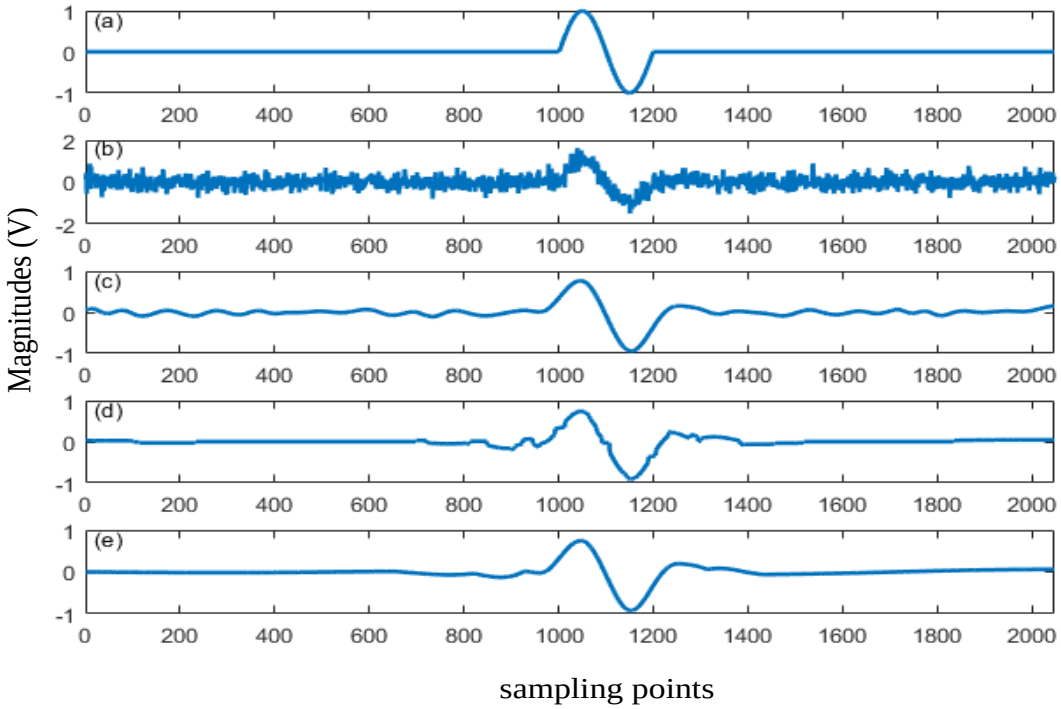

Figure 3-22 EMD-based denoising results, (a) original signal, (b) noisy signal, (c) EMD-TR, (d) EMD-DT, (e) EMD-IT

3.6 Conclusion

This chapter reviewed the strengths and weaknesses of the most commonly used denoising techniques, i.e., digital filters, matched filters, wavelet-based denoising, and EMD-based denoising, in PD denoising. Digital filters are more suitable for noise that has a major difference with PD signals in frequency domain. Matched filters require a prior knowledge of the shape of PD signals, which determines their effectiveness in noise reduction. Wavelet-based denoising presents its advances in PD denoising due to its appealing attribute, i.e., the transformed PD signals can be simultaneously represented in both time and frequency domain. However, its effectiveness is dictated by the wavelet used, the decomposition scale selected, and the threshold estimated. Based on this, the motivation of this research is to propose new algorithms based on the study of characteristics of both PD signals and noise, and thus, to improve the performance of wavelet-based denoising in PD detection. EMD-based denoising has an attractive attribute, i.e., a signal can be decomposed by EMD without the requirement of any predefined base function, which is a concern commonly

expressed by critics for wavelet-based technique. As a result, another area of interest in this research is to investigate the effectiveness of EMD-based denoising in PD detection through the relative mode selected and noise estimation of IMFs. In turn, a suitable PD denoising strategy can be recommended through performance comparisons between wavelet-based and EMD-based denoising.

4 Wavelet Selection Schemes for Wavelet-based PD Denoising

4.1 Introduction

Signal denoising can benefit from a good signal representation. Generally, a signal representation is defined by a function $\Re: H \to S$ that maps a Hilbert space H into a space of sequences [131]. For a given signal, $x \in H$, its representation $\Re(x)$ is a sequence in space S and defined by

$$\Re(x) = (s_1, s_2, s_3, \dots) \tag{4-1}$$

where s_n, $n = 1,2,3,\dots$, is a pair $(\alpha_n, g_{\gamma n})$, α_n represents a coefficient and $g_{\gamma n}$ represents a waveform. For a signal decomposition, it is the critical step to obtain the sequence of waveforms $(g_{\gamma n})_{n \in Z}$ and their associated coefficients $(\alpha_n)_{n \in Z}$. A suitable sequence of waveforms can successfully reveal signal properties through the corresponding coefficients. Numerous methods in signal decomposition have been investigated for this purpose. Basis, e.g., mother wavelet in wavelet transform, is such a method widely used as an effective tool for signal decomposition.

A basis is defined as a set of linearly independent elements $(\phi_\lambda)_{\lambda \in \Gamma}$ that span the Hilbert space H. Linear independence indicates that the elements in basis set are orthogonal, i.e., any one function in this set cannot be expressed by the combination of the others. This orthogonal property of base functions implies the set used for signal representation is minimal. For an orthogonal basis set, its base functions should satisfy [131], [156]

$$\int_{-\infty}^{\infty} \phi_i(t)\, \bar{\phi}_j(t) dt = c_i \cdot \delta(i-j) = \begin{cases} c_i, i = j \\ 0, i \neq j \end{cases}, \forall i, j \in \Gamma \tag{4-2}$$

where if $c_i = 1$ for all i, the $\phi_i(t)$s are said to be normalized. In this case, the representation of signal x is exact and the reconstruction is given by [131]

$$x = \sum_{\lambda \in \Gamma} \langle x, \phi_\lambda \rangle \cdot \phi_\lambda \tag{4-3}$$

where the inner product $\langle x, \phi_\lambda \rangle = \int_{-\infty}^{\infty} x(t) \cdot \bar{\phi}_\lambda(t) dt$ is interpreted as the projection of the signal x in the base functions ϕ_λ.

In wavelet transform, ψ_0 is the mother wavelet that corresponds to the basis set as defined above, and thus, its dilated or compressed versions, i.e., daughter wavelets $\psi_{j,i}$ in (3-6), corresponds to the base functions. As shown in Figure 4-1, wavelet selection is one of three important steps necessary for a successful wavelet-based denoising. As a result, different denoised signals can be obtained by using different wavelet base functions in practice [157]. In relation to wavelet-based PD denoising, the simulated or detected PD signals are projected into $\psi_{j,i}$ to obtain the corresponding coefficients $W_{j,i}$. As such, it is of paramount importance to select a suitable ψ_0 for wavelet-based denoising as ψ_0 can be translated and scaled to represent the PD signal of interest as effectively as possible. Based on this, the investigation of an appropriate wavelet for wavelet-based PD denoising has been implemented in [22], [88].

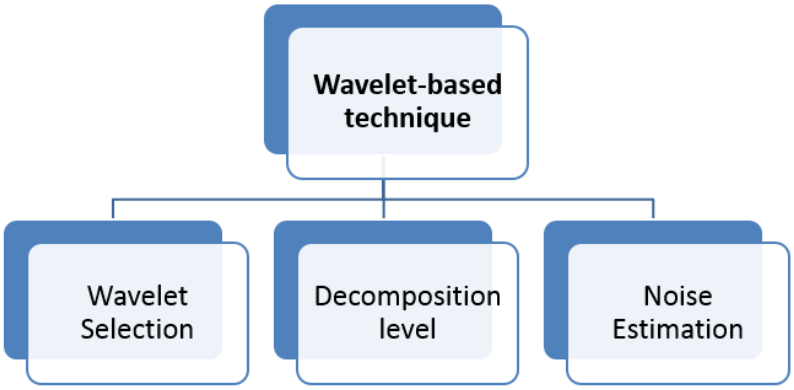

Figure 4-1 Three distinct steps necessary for a successful wavelet-based denoising

A wavelet selection scheme (WSS) was introduced in [22] based on the correlation coefficient between a known PD signal and wavelet waveform. This scheme is termed correlation-based wavelet selection scheme (CBWSS). The optimal wavelet is desired to generate the highest wavelet coefficients in wavelet analysis of PD signals, and thus, the essence of PD signal of interest can be effectively preserved. Based on this, the wavelet that can maximize the correlation coefficient is selected as the most appropriate wavelet for PD denoising. This approach for best wavelet selection, however, has an inherent limitation, it requires prior knowledge of PD waveforms. The waveform of PD signals depends on: the type and location of PD sources, propagating medium and path, and the detecting circuit. The variability of PD

waveforms impedes the application of CBWSS for online PD monitoring systems. Also, it is not a scale-dependent wavelet selection scheme. The selected wavelet is applied to a PD singal during the entire singal decomposition and reconstrucntion. As a result, the denoised PD signal may not be as good as expected. The most significant limit of CBWSS is that a PD signal is normally corrupted by the noise in the environment, which can lead to the selected wavelet being a match of the noise rather than the PD signal, especially when the signal to noise ratio (SNR) is very low. In an attempt to overcome the limitation mentioned above in CBWSS, a scale-dependent energy-based wavelet selection scheme (EBWSS) was presented in [88]. The wavelet that can maximize the energy ratio of approximation coefficients at each decomposition scale is selected as the best wavelet. It has been demonstrated to outperform CBWSS in PD denoising [88]. In EBWSS, two typical PD waveforms, DEP and DOP, as shown in Figure 2-19, were used to demonstrate the energy criterion for the optimal wavelet selection. With further analysis in details of EBWSS, it has been found that the criterion is not strictly true for DOP signals, particularly when the decomposition scale increases. One of the research objectives in this thesis is therefore to provide an automated and data-driven wavelet selection scheme for the choice of an appropriate wavelet in the context of PD denoising.

It is necessary to mention that the noise involved in the investigation of WSS in [22], [88] is white noise. The underlying reason is that white noise is difficult to be identified and suppressed as both white noise and PD signals have wide-band frequencies [158]. Also, white noise is a very common noise source for PD contamination in off-line and on-line PD measurement. As such, the new wavelet selection schemes proposed in this chapter start from white noise reduction, and then extends to investigate if it is still effectvie when DSI is involved.

Further study of statistical behaviors of both PD signals and noise may shed light on the improvement of PD denoising techniques. As mentioned in Section 2.7, Gaussianity is one of the prominent characteristics of whtie noise. However, PD signals are generally irregular, aperiodic, and transient, which lead to the non-Gaussianity of them. In statistics, the fourth moment of a data set, i.e., kurtosis, can be calculated to evalute the Gaussianity. Thus, the first new WSS is proposed based on this difference, i.e., the wavelet that can minimize the

kurtosis of detail coefficients will be selected as the best wavelet. This new WSS is termed kurtosis-based WSS (KBWSS). Also, white noise is a random and disordered noise source. The analysis of PD characteristics in Section 2.4 indicates that PD signals have less randomness and disorder than white noise. Based on this, the second new WSS is inspired by the concept of Shannon Entropy [159], and the associated information cost functions (ICF) in information theory [160]–[162]. An ICF can select the best wavelet to expand a signal in wavelet domain. Wavelet entropy, derived from Shannon Entropy, can measure the randomness of the wavelet coefficients at each decomposition scale. The smaller the wavelet entropy, the lower the randomness of the wavelet coefficients. As such, the second new selection scheme is proposed with the combination of ICF and wavelet entropy, and termed wavelet entropy-based WSS (WEBWSS). Simulated PD signals, i.e., DEP and DOP, are used to demonstrate the performance of these novel wavelet selection schemes. Results show that WEBWSS is a promising WSS to improve the effectiveness of PD denoising, while KBWSS cannot provide any improvement as compared to EBWSS. Further demonstration of WEBWSS will be performed through PD signals obtained through laboratory experiment using test samples with artificial defects and on-site PD measurements in Chapter 8.

4.2 Correlation-based Wavelet Selection Scheme (CBWSS)

In signal processing, correlation is a measure of association between two signals, and most commonly used is the linear correlation coefficient. For two signals, x_i and y_i, $i = 1,2,\dots,N$, the normalized correlation coefficient γ is given by [22]

$$\gamma = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \cdot \sqrt{\sum_i (y_i - \bar{y})^2}}, \tag{4-4}$$

where $\bar{x}$ is the mean of x_i and $\bar{y}$ is the mean of y_i. The value of γ is in the range of -1 to 1. It takes on a value close to 1 indicating x_i and y_i are positively correlated, and a value close to -1 denoting they are negatively correlated. A value of γ near zero means x_i and y_i are uncorrelated.

Correlation in CBWSS [22] is used as a measure of the similarity between a pure PD signal and a wavelet. Note that this similarity is referred to as their shapes. The more similar their shapes, the higher the correlation coefficient is. As such, the wavelet that has the highest correlation

coefficient with the shape of a PD signal is selected to maximize the wavelet coefficients through wavelet analysis.

The general process for the choice of appropriate wavelet using CBWSS is described as follows:

a. Analyse the detected PD signal to generate a 'typical' PD pulse,
b. Set up a wavelet library, consisting of the wavelets that have similar characteristics to the PD pulse,
c. Normalize the PD pulse and each wavelet retrieved from the wavelet library,
d. Calculate the correlation coefficient, γ, between the PD pulse and each wavelet,
e. Select the wavelet that has the maximum correlation coefficient with the PD pulse, it will be applied for the following wavelet-based denoising

Figure 4-2 also illustrates a flow chart of the general process of CBWSS, where $s(n)$ is the PD signal and ψ_i is the wavelet.

As aforementioned, the CBWSS approach is limited by noise and is scale-independent. Also, a heuristic method was introduced in [88] to obtain better correlation results. Resampling both the PD signal and wavelet function in time domain is applied to align their peaks as well as their first zero-crossing points after the peaks. This heuristic method is adopted in this chapter for comparisons of denoising results among EBWSS, KBWSS, and WEBWSS methods.

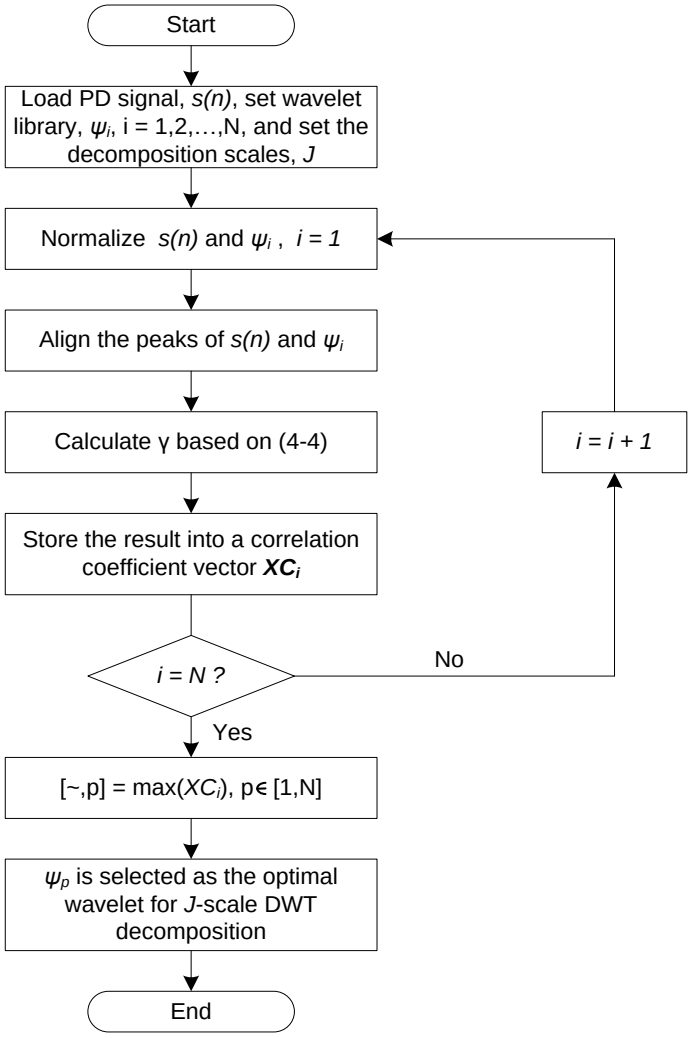

Figure 4-2 Flow chart of the general process of CBWSS

4.3 Energy-based Wavelet Selection Scheme (EBWSS)

EBWSS was proposed by Li [88], in which the wavelet that can maximize the energy ratio of the approximation coefficients is selected as the best wavelet for PD denoising. For a one-dimensional wavelet decomposition, the energy ratio of approximation a_j at scale j is defines as follows [88]:

$$E_a = \frac{\sum_i {a_{j,i}}^2}{\sum_i {a_{j,i}}^2 + \sum_j \sum_i {d_{j,i}}^2} \tag{4-5}$$

where $i = 1,2, \dots, n_i, n_i$ is the length of approximation coefficients or detail coefficients at scale j, and d_j is the detail coefficients at scale j.

The idea of wavelet energy was introduced in EBWSS. For an orthogonal wavelet, energy preservation is one of the desirable properties of DWT [163]. The equation for energy

preservation is given by

$$\|X\|^2 = \|a\|^2 + \|d\|^2 \tag{4-6}$$

where a and d are the approximation and detail coefficients of the DWT of a signal X. This property is also applied to PD signals using DWT decomposition. A PD signal s can be decomposed into J scales with $J+1$ signals, i.e., $s_1, s_2, \ldots, s_J, s_{J+1}$. Among these signals, $s_1, s_2, \ldots, s_J$ are detail coefficients from scale 1 to scale J, while s_{J+1} is the approximation coefficients at scale J. The energy of a decomposed signal s_k is given by

$$E_k = \sum_i {s_k}^2(i) \tag{4-7}$$

where $k = 1,2,\ldots,J+1$, $i = 1,2,\ldots,n_i$, and n_i is the length of s_i. Then, s can be represented by a normalized energy vector $(e_1, e_2, \ldots, e_J, e_{J+1})$, where e_k is defined as

$$e_k = \frac{E_k}{\|s\|^2} = \frac{E_k}{\sum_{k=1}^{k=J+1} E_k} \tag{4-8}$$

It can be seen that the concept of energy ratio in EBWSS can be interpreted as a normalized energy vector. Figure 4-2 shows the DEP, DOP, and white noise (WN) used in [88] to explain the criterion of EBWSS for wavelet selection. Figure 4-3 (a) and (b) show the DEP signal and its normalized energy vector. Equally, Figure 4-3 (c) and (d) present the DOP signal and its normalized energy vector. Figure 4-3 (e) and (f) delineate WGN and its normalized energy vector. Based on Figure 4-3 (b), (d) and (e), the approximations of the DEP and DOP signals cover the most energy of total coefficients while the details of WGN preserve the most energy of total coefficients [88].

The general process for the choice of an appropriate wavelet using EBWSS is presented with its corresponding flow chart (Figure 4-4) as follows:

a. Given a wavelet library $\{\psi_i : i = 1,2,\ldots,N\}$, select a wavelet from $\{\psi_i\}$, and perform a one-scale DWT decomposition of a noisy PD signal *s(n)*. Obtain its approximation coefficients $a_1^{(i)}$ and detail coefficients $d_1^{(i)}$,

b. Calculate the energy ratio of approximation coefficients $E_{a_1^{(i)}}$ based on (4-5). If $E_{a_1^{(p)}}$ is the maximum of $E_{a_1^{(i)}}$, $1 \le p \le N$, select ψ_p as the optimal wavelet for the first scale,

c. Apply ψ_p to obtain the approximation coefficients $a_1^{(p)}$ and $d_1^{(p)}$,

d. $a_1^{(p)}$ is used as the input signal for next-scale DWT decomposition, and select the optimal

wavelet based on the strategy used in steps a, b, and c,

e. Iterate the steps above until the predefined decomposition scale J reaches. The optimal wavelet for each decomposition scale will be selected.

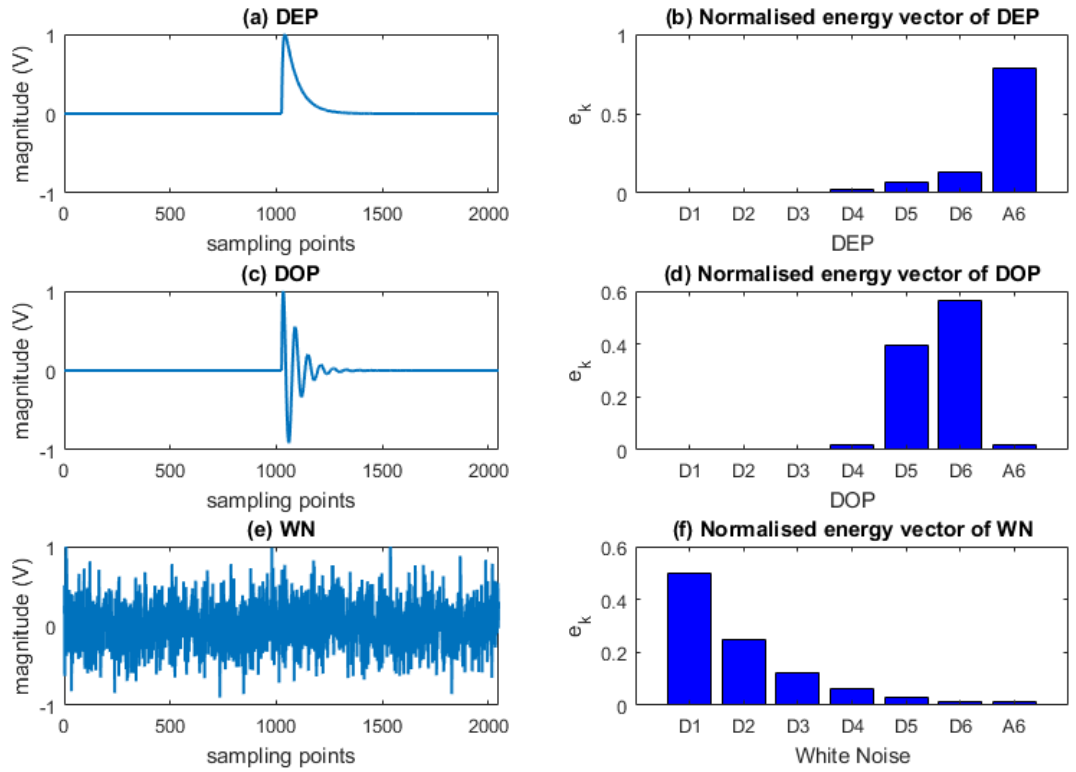

Figure 4-3 Representations of (a): DEP, (c): DOP, and (e): White noise by normalized energy vectors (b), (d) and (f) respectively

Results in [88] show that EBWSS outperforms CBWSS for the best wavelet selection. However, EBWSS is not as robust as expected. It selects the wavelet that can maximize the energy ratio of approximation coefficients. It is not strictly true for DOP signals, particularly when the decomposition scale increases. It can be seen from the normalized vector of DOP in Figure 4-3 (d), the energy of PD signal with a 6-scale decomposition is preserved on the detail coefficients rather than approximation coefficients. When more scales are required, e.g., 7 scales, the EBWSS is still trying to select the appropriate wavelet by maximizing the energy ratio of approximation coefficients.

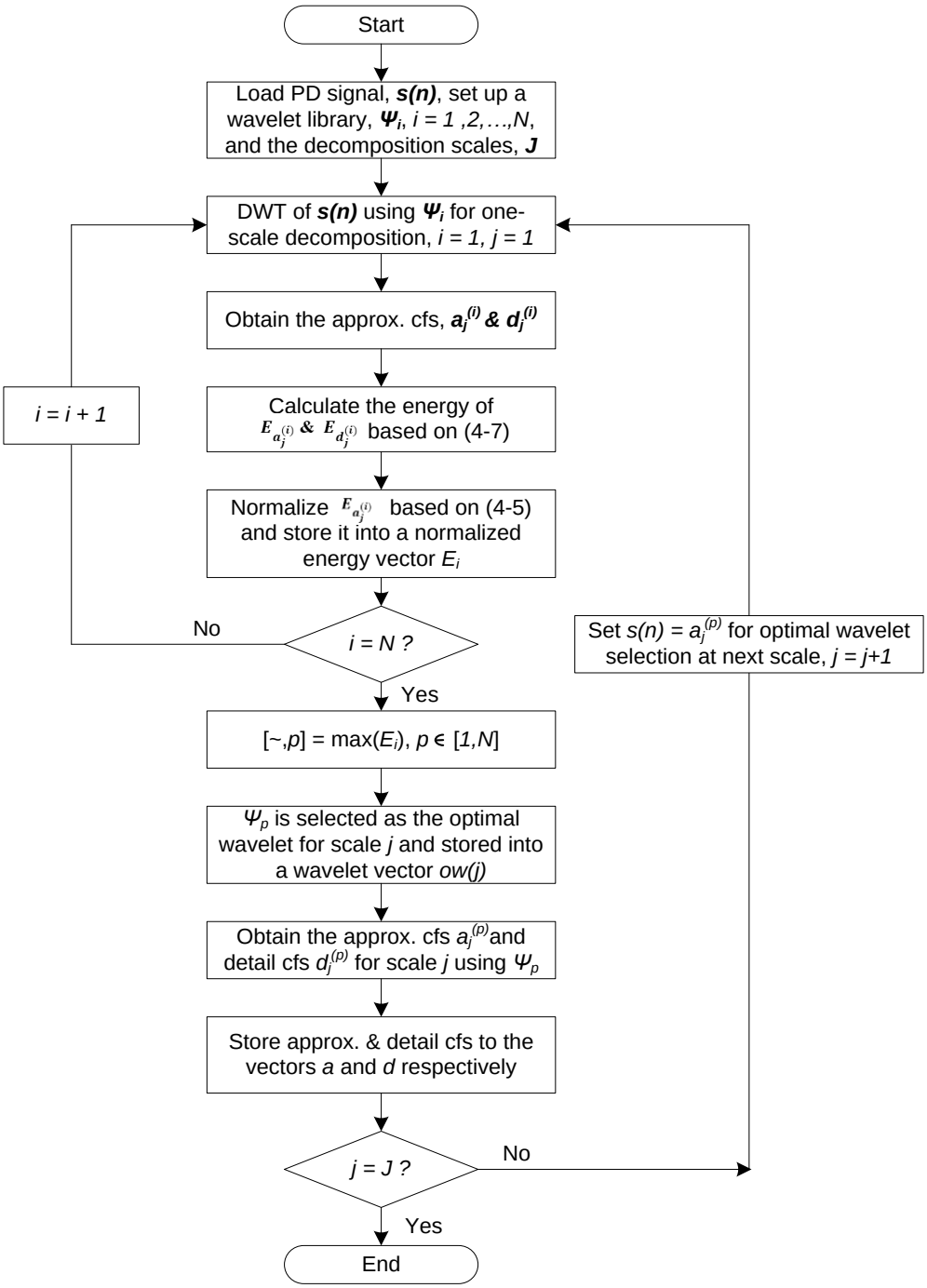

Figure 4-4 Flow chart of the general process of EBWSS

The limitation of EBWSS can be interpreted in the frequency domain. Based on Parseval's theorem, the time and frequency domains are equivalent representations of the signal, and thus, they must have the same energy [164]. As shown in Figure 3-13, the filter pairs of DWT iteratively halve the frequency bands of a signal with the increase of decomposition scales. The spectrum of DEP, DOP, and white noise are illustrated in Figure 4-5 (a), (b), and (c) respectively.

With a 6-scale decomposition, the filter pairs iteratively separate these signals into disjoint frequency bands, $G_1(w), G_2(w), \ldots, G_6(w)$ and $H_6(w)$. From the spectral curve of DOP, it is clear that the magnitudes of frequency in $G_5(w)$ and $G_6(w)$ are larger than those at other frequency bands. It is in agreement with the normalized energy vector of DOP shown in Figure 4-3 (b). With further decomposition, the energy of the signal will be preserved in detail coefficients rather than approximation coefficients.

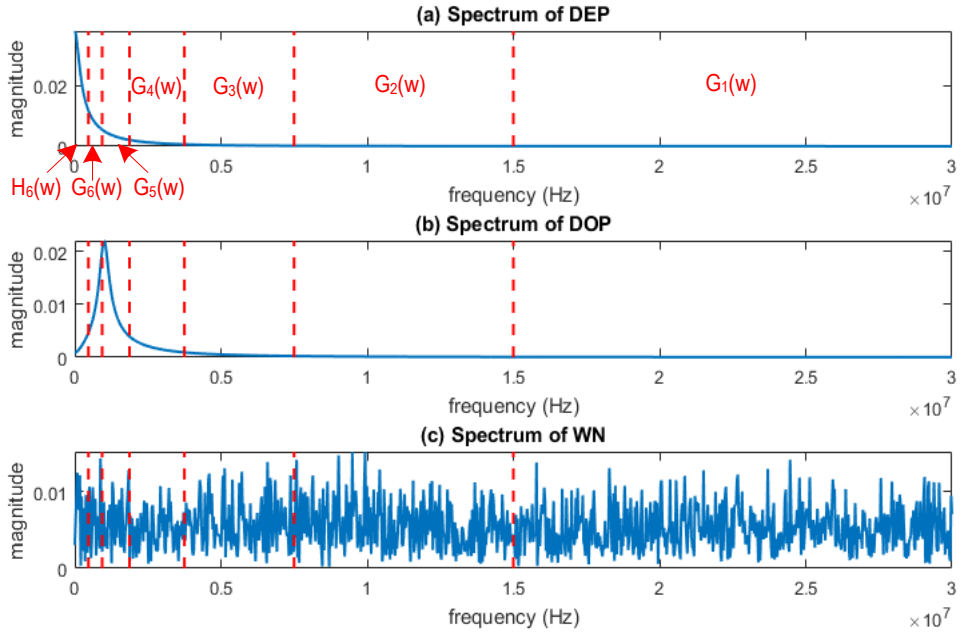

Figure 4-5 Spectrum of (a): DEP, (b): DOP, and (c): WN

4.4 Kurtosis-based Wavelet Selection Scheme (KBWSS)

Kurtosis in statistics is a parameter used to measure the shape of a distribution. It is formally defined as the standardized fourth moment about the mean, and is given by [165]

$$kurt = \frac{E(X-\mu)^4}{(E(X-\mu)^2)^2} = \frac{\mu_4}{\sigma^4} \tag{4-9}$$

where E is the expectation operator of X, μ is the mean of X, μ_4 is the fourth moment about the mean, and ρ is the standard deviation of X. The reference standard is a normal distribution, which has a kurtosis of 3. Generally, excess kurtosis, i.e., $kurt - 3$, is often used, and thus, the reference normal distribution has an excess kurtosis of zero. Note that the kurtosis mentioned below will be referred to as excess kurtosis if not specified.

As discussed in Section 2.7, white noise can be modeled by a Gaussian model, which indicates that white noise obeys the normal distribution. In contrast to white noise, PD signals do not obey the normal distribution as a result of their non-Gaussianity. This difference between white noise and PD signals can be revealed by the kurtosis, as listed in Table 4-1. DEP and DOP are shown in Figure 2-19, and white noise is shown in Figure 2-32. It can be seen that white noise has a kurtosis value extremely close to zero, while both DEP and DOP have a very high kurtosis value.

Table 4-1 Kurtosis of DEP, DOP, and White noise

Signals	***kurt***
DEP	23.5079
DOP	35.3862
$White\ noise$	0.0008

For wavelet-based PD denoising, it is accepted that noise is predominant in detail coefficients of wavelet decomposition. It is the reason that thresholding functions are applied to detail coefficients for noise reduction. Based on this, a new scale-dependent WSS, termed KBWSS, is developed using the kurtosis value of detail coefficients at each decomposition scale. The wavelet that can have the kurtosis value of detail coefficients closer to zeros will be selected as the best wavelet at that scale. A flow chart of the general process of KBWSS for wavelet selection is illustrated in Figure 4-6.

In KBWSS, a wavelet library $\{\psi_i: i = 1,2,\dots,N\}$ is set, and each wavelet in the library is extracted for the first-scale decomposition of a noisy PD signal *s(n)*. The kurtosis value of the detail coefficients is calculated based on (4-9) and stored into a kurtosis vector. When all the wavelets in the library have been applied to *s(n)*, the wavelet ψ_p $(1 \le p \le N)$ that have the kurtosis value closer to zero is selected as an appropriate wavelet at this level. In turn, ψ_p is applied to *s(n)* for first-scale decomposition, generating approximation and detail coefficients, $a_1^{(p)}$and$d_1^{(p)}$, respectively. $a_1^{(p)}$ is then adopted as the input signal for next-scale decomposition. The wavelet selection criterion is iterated until the predefined decomposition scale *J* reaches.

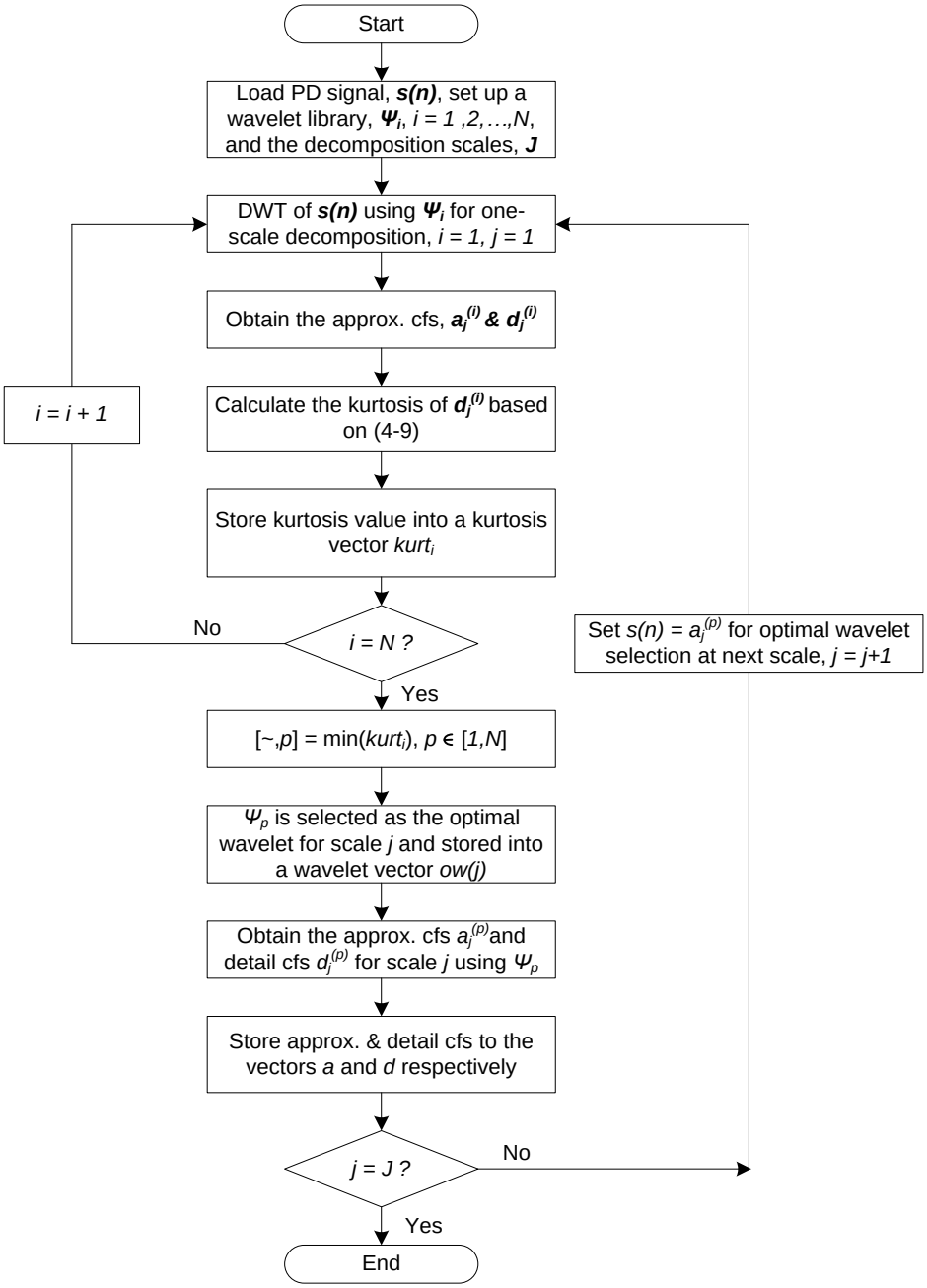

Figure 4-6 Flow chart of the general process of KBWSS

4.5 Wavelet Entropy-based Wavelet Selection Scheme (WEBWSS)

4.5.1 Wavelet Entropy

The concept of wavelet entropy was derived from Shannon entropy and presented in [166].

Suppose $\{W_{j,i}\}$ are the wavelet coefficients obtained through a *J*-scale wavelet transform, in which *j* represents the decomposition scale and $j = 1, 2, \dots, J$, i denotes the i^{th} element in $W_{j,i}$ and $i = 1, 2, \dots, n_j$, n_j is the length of wavelet coefficients at scale *j*. The energy of wavelet coefficients at the decomposition scale *j* can be calculated by

$$E_j = \sum_i |W_{j,i}|^2 \tag{4-10}$$

The distribution of energy probability for wavelet coefficients at scale *j* can be derived by

$$p_i = \frac{|W_{j,i}|^2}{\sum_i |W_{j,i}|^2} = \frac{|W_{j,i}|^2}{E_j} \tag{4-11}$$

with $\sum_i p_i = 1$. Wavelet entropy $WE(j)$ is defined as follows [166]:

$$WE(j) = -\sum_i p_i \ln(p_i) \tag{4-12}$$

where $0 \leq WE(j) \leq ln(n_j)$.

Similar to Shannon entropy, wavelet entropy is applied to measure the degree of disorder of wavelet coefficients or signify the randomness of wavelet coefficients. It is important to note that wavelet entropy is not an information cost function (ICF), since it requires the energy of wavelet coefficients to be normalized as shown in (4-11), and is thus not additive [161], [162]. Substituting (4-11) into (4-12) yields:

$$\begin{aligned} WE(j) &= \sum_i p_i \ln\left(\frac{1}{p_i}\right) \\ &= \sum_i \frac{|C_{j,i}|^2}{E_j} \ln \frac{E_j}{|C_{j,i}|^2} \\ &= \frac{1}{E_j}\left(\sum_i |C_{j,i}|^2 \ln E_j + \sum_i |C_{j,i}|^2 \ln \frac{1}{|C_{j,i}|^2}\right) \\ &= \ln E_j + \frac{1}{E_j}\left(\overbrace{\sum_i |C_{j,i}|^2 \ln \frac{1}{|C_{j,i}|^2}}^{l}\right) \\ &= \ln E_j + {}^{l}/_{E_j}\,. \end{aligned} \tag{4-13}$$

In (4-13), l is an ICF based on the definition in [161]. As such, wavelet entropy is a monotonic-increasing function of l, which means minimizing l over wavelet coefficients minimizes wavelet entropy.

4.5.2 WEBWSS Fundamentals

Given ℓ is an ICF over a sequence x, the best basis used for the representation of x is the basis that can minimize the value of $\ell(\phi x)$, where ϕx is the coefficient vector of x in the orthogonal basis ϕ[160]. Such an ICF, from a practical point of view, can describe concentration or number of coefficients required to sufficiently represent the sequence. As aforementioned, wavelet entropy is not an ICF, but l in (4-13) is monotonically increased with the wavelet entropy. As such, the best wavelet also can be selected when the value of wavelet entropy is minimum. In [161], it was shown that wavelet entropy value is inversely proportional to the energy concentrated in the number of wavelet coefficients. Also, an *N*-sample signal can be viewed as noisy or incoherent relative to the selected wavelet if it does not correlate well with the wavelet, i.e., if its entropy is of the same order of magnitude as $ln\, N - \varepsilon$ with small ε [161]. The denoising can benefit from this incoherence between white noise and the selected wavelet. White noise, as a frequently occurring noise source for PD corruption, has high degree of randomness or disorder, and thus, the entropy value can describe the random characters of noise [167]. Based on this, a smaller value of wavelet entropy indicates that the wavelet used for WT decomposition can preserve more energy of the original signal in fewer number of coefficients and contain less white noise in the wavelet coefficients and, consequently, the wavelet used is closer to the best wavelet as expected. Differences can be seen from Table 4-2 when the wavelet entropy values of decomposed white noise are compared to those of decomposed DEP and DOP and the theoretical maximum entropy values. In Table 4-2, DEP and DOP are simulated PD signals, as shown in Figure 4-3. White noise is simulated by a Gaussian model with 2048 sampling points, equal to the length of DEP and DOP. 6-scale wavelet decomposition has been applied to DEP and white noise using the wavelet 'db3', and to DOP and white noise using the wavelet 'db10'. From the figures listed in Table 4-2, it is clear that the wavelet entropy values of DEP and DOP are less than that of white noise at each scale. Also, the wavelet entropy values of white noise are close to its theoretical maximum entropy value, which means white noise is incoherent to the wavelet used and presents high degree of randomness or disorder at each decomposition scale. In contrast, the wavelet entropy values of both DEP and DOP indicate that they are far less random or disordered at each scale. As such, a new criterion for the best wavelet

selection is therefore proposed, i.e., a wavelet that can have minimum wavelet entropy of the approximation coefficients at each decomposition scale through WT decomposition will be selected for denoising of PD detection. The new method has several promising advantages: it is scale-dependent, automated, and data-driven.

Table 4-2 Wavelet entropy values of approximation coefficients of DEP, DOP and white noise with 6-scale decomposition

	DEP		**WN**		**DOP**		**WN**	
	Entropy value	**Max**	**Entropy value**	**Max**	**Entropy value**	**Max**	**Entropy value**	**Max**
a_1	*3.8526*	6.9334	*6.1321*	6.9334	*3.5901*	6.9402	*6.1391*	6.9402
a_2	*3.1627*	6.2442	*5.5611*	6.2442	*2.9095*	6.2653	*5.4684*	6.2653
a_3	*2.4834*	5.5607	*4.8463*	5.5607	*2.2966*	5.6058	*4.8418*	5.6058
a_4	*1.8301*	4.8828	*4.1916*	4.8828	*1.4736*	4.9767	*4.2325*	4.9767
a_5	*1.2912*	4.2195	*3.2735*	4.2195	*1.4094*	4.4067	*3.6450*	4.4067
a_6	*0.988*	3.5835	*2.377*	3.5835	*1.0035*	3.9120	*3.2127*	3.9120

The general process for the proposed novel wavelet selection scheme is illustrated in a flow chart in Figure 4-7. Given a wavelet library $\{\psi_i: i = 1,2,\dots,N\}$, one wavelet of which is selected for a one-level DWT decomposition of a noisy PD signal *s(n)* each time. Next the wavelet entropy of the generated approximations is calculated based on (6) and (7). The wavelet ψ_p $(1 \le p \le N)$ that minimize the wavelet entropy of approximations will be selected as the best wavelet. The selected ψ_p is then applied for the DWT decomposition of *s(n)* for the first scale, obtaining approximation coefficients ${a_1}^{(p)}$ and detail coefficients ${d_1}^{(p)}$. Finally, ${a_1}^{(p)}$ is used as the input signal for next scale DWT decomposition, using the strategy presented above. When the predefined decomposition scale *J* reaches, the best wavelet for each scale will be successfully selected.

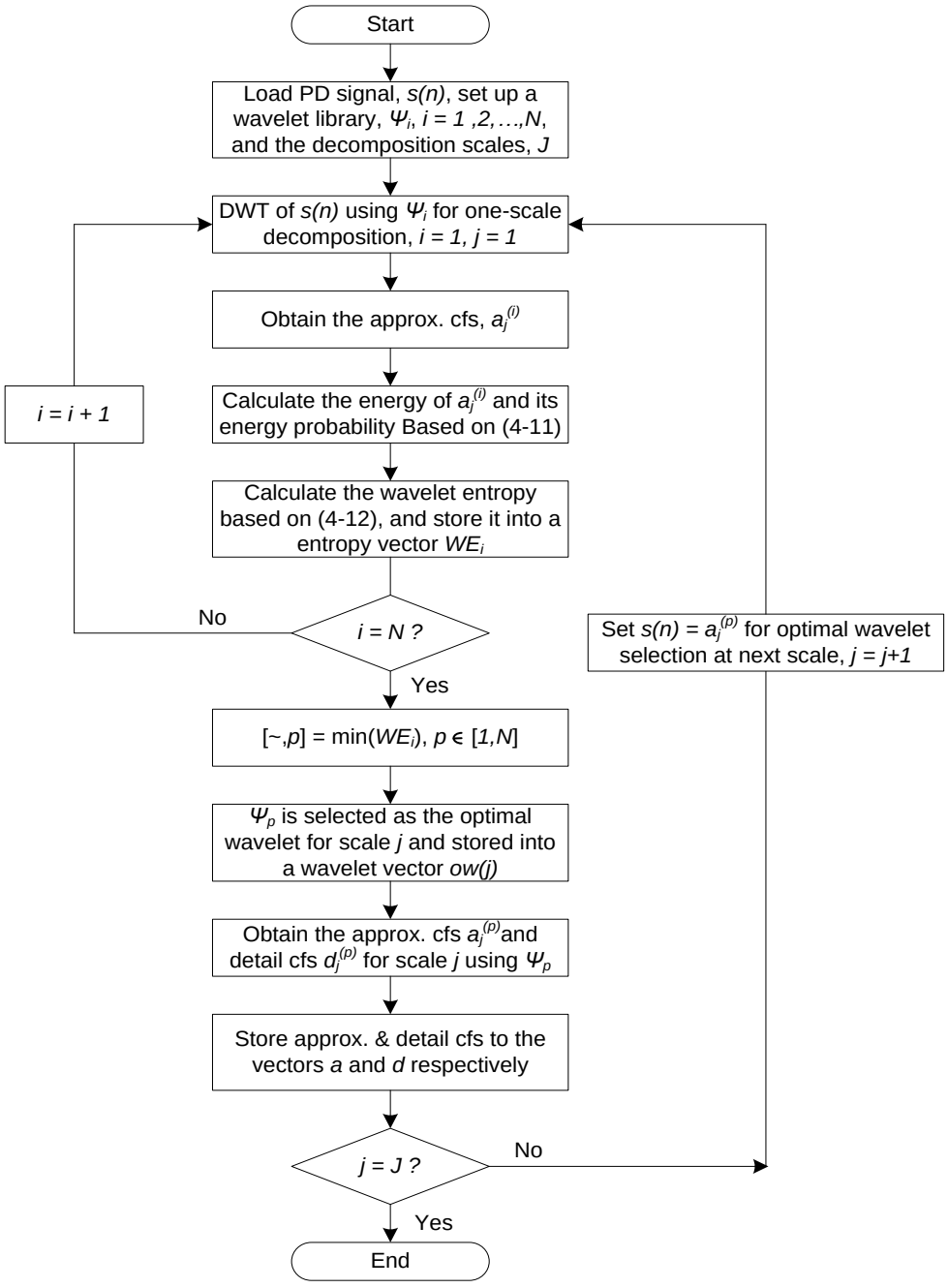

Figure 4-7 Flow chart of the general process of WEBWSS

4.6 Performance Evaluation of New Proposed Selection Schemes

4.6.1 Parameters for Performance Evaluation

Generally, parameters, e.g., magnitude error (ME), mean square error (MSE), signal to noise ratio

(SNR), and cross correlation (XCORR) are adopted to evaluate the performance of a proposed denoising method or algorithm. Given a noisy signal is $y(i), i = 1,2,\ldots,N$, it can be expressed by

$$y(i) = s(i) + n(i) \tag{4-14}$$

where $s(i)$ represents the signal of interest and $n(i)$ represents noise. Based on (4-14), XCORR, ME, MSE and SNR are calculated by the equations as follows,

$$XCORR = \frac{\sum_{i=1}^{N}(s(i) - \bar{s}(i)) \cdot (s'(i) - \bar{s'}(i))}{\sqrt{\sum_{i=1}^{N}(s(i) - \bar{s}(i))^2} \cdot \sqrt{\sum_{i=1}^{N}(s'(i) - \bar{s'}(i))^2}} \tag{4-15}$$

$$ME = \frac{m - m'}{m} \tag{4-16}$$

$$MSE = \frac{\sum_{i=1}^{N}(s(i) - s'(i))^2}{N} \tag{4-17}$$

$$SNR(dB) = 10 \cdot log_{10}\frac{E_s}{E_n} \tag{4-18}$$

where $s'(i)$ in (4-15) denotes the denoised signal, $\bar{s}(i)$ and $\bar{s'}(i)$ are the mean of $s(i)$ and $s'(i)$ respectively. m and m' in (4-16) are the peak value of $s(i)$ and $s'(i)$ respectively. Better denoised results can be obtained with lower ME, MSE, and higher XCORR. The SNR is used to measure the noise level in the signal, E_s and E_n in (4-18) denote the energy of the signal of interest and noise, respectively. Given the signal $s(i)$ is known, the SNR after denoising can be calculated by

$$SNR(dB) = 10 \cdot log_{10}\frac{\sum_{i=1}^{N}(s(i))^2}{\sum_{i=1}^{N}(s(i) - s'(i))^2} \tag{4-19}$$

Equally, the higher SNR in (4-19), the lower level of noise is remained in the denoised $s'(i)$. In this chapter, SNR, ME, MSE, and XCORR are used to compare the denoising results of different wavelet selection schemes.

4.6.2 PD Signals Corrupted by White Noise

As mentioned before, the test of these WSSs starts from PD signals corrupted by white noise. In the analysis of EBWSS, the underlying idea is not strictly true when the decomposition scale increases over 6. To highlight this limit, the decomposition scale is set to 7 for the performance evaluation. The wavelet library used for all selection schemes consists of wavelets in the range of 'db2' to 'db25' from Daubechies family, which is in agreement with

those used in [88], and thus, is convenient for comparisons of denoising results among all WSSs. Two simulated PD signals, s_1 and s_2, which are referred to as DEP and DOP (see Figure 4-3) respectively, as well as their noisy signals ns_1 and ns_2 with SNR = -10 are depicted in Figure 4-8.

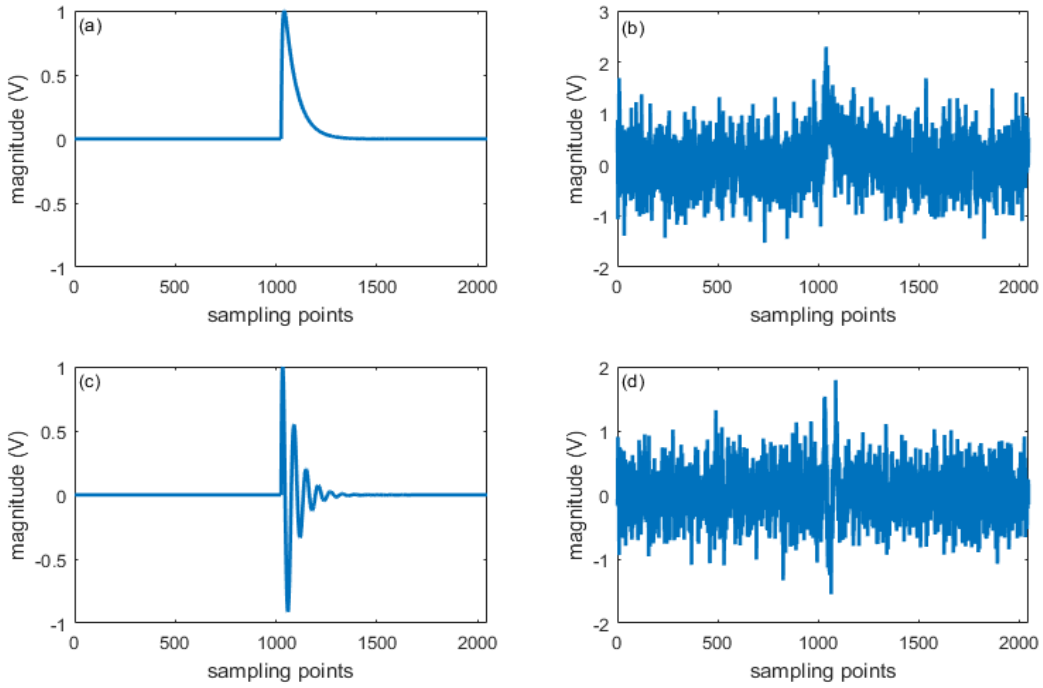

Figure 4-8 Pure and Noisy PD Signals, (a) s_1, (b) ns_1, (c) s_2, (d) ns_2

Various wavelet selection schemes are used to reduce white noise and evaluated by the parameters introduced above. The denoised versions of ns_1 and ns_2 using the four discussed WSSs are delineated in Figure 4-9 and Figure 4-10, respectively. Due to different wavelet selection criteria, Table 4-3 lists the wavelet selected at each decomposition scale for the denoising of ns_1 and ns_2. In CBWSS, fixed wavelets 'db3' and 'db9' are selected for ns_1 and ns_2 at all decomposition scales. In contrast, the wavelet selected at each decomposition scale presents variations in EBWSS, KBWSS, and WEBWSS. In Figure 4-9 and Figure 4-10, the figures in the top right corner are the SNRs after denoising. Based on this parameter, it can be seen that WEBWSS has the best denoising result among all the WSSs for both noisy PD signals ns_1 and ns_2. Also, the SNR indicates that KBWSS does not exhibit any robustness as compared to EBWSS, although it has a higher SNR than CBWSS in ns_1

denoising (see Figure 4-9 (a) and (c)). From a practical point of view, the selection scheme proposed based on kurtosis value of detail coefficients at each decomposition scale is not an appropriate WSS in PD denoising. The underlying reason may be from the accuracy error of kurtosis calculation of white noise in MATLAB. Although white noise is simulated through a Gaussian model, the kurtosis value is not zero. As such, KBWSS will not be further discussed in this thesis.

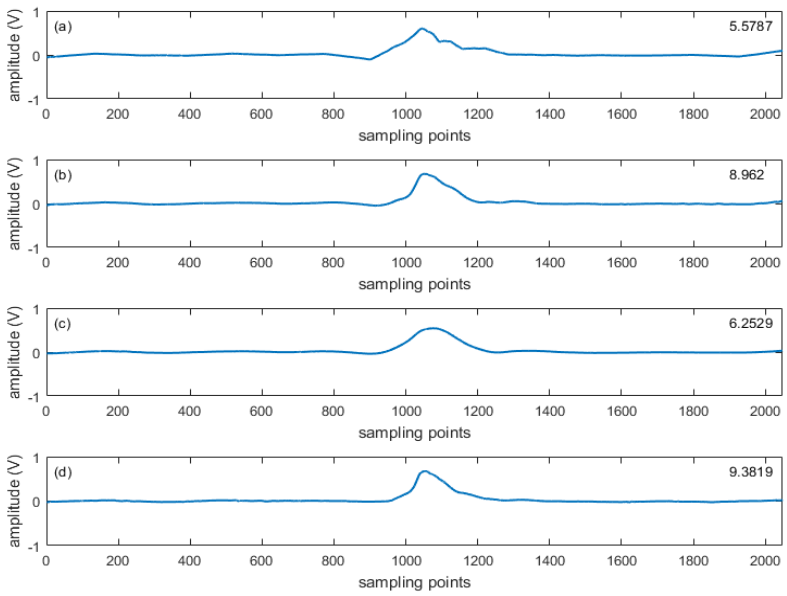

Figure 4-9 Denoised versions of ns_1 using (a) CBWSS, (b) EBWSS, (c) KBWSS, (d) WEBWSS

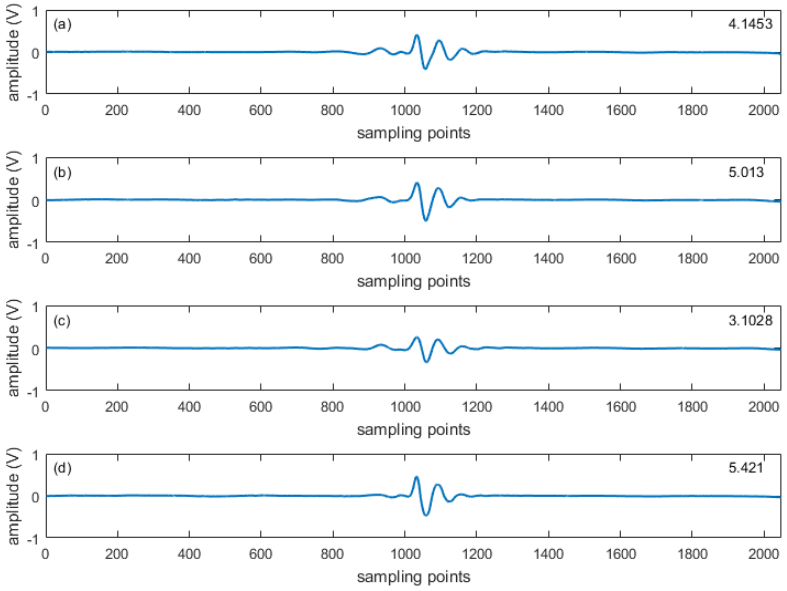

Figure 4-10 Denoised versions of ns_2 using (a) CBWSS, (b) EBWSS, (c) KBWSS, (d) WEBWSS

Table 4-3 Selected wavelet using different WSSs at each decomposition scale for ns_1 and ns_2

		Decomposition Scale						
		1	**2**	**3**	**4**	**5**	**6**	**7**
s_1	**CBWSS**	'db3'	'db3'	'db3'	'db3'	'db3'	'db3'	'db3'
	EBWSS	'db7'	'db5'	'db2'	'db2'	'db7'	'db4'	'db7'
	KBWSS	'db22'	'db22'	'db25'	'db10'	'db8'	'db21'	'db5'
	WEBWSS	'db2'	'db6'	'db2'	'db2'	'db4'	'db2'	'db2'
s_2	**CBWSS**	'db9'	'db9'	'db9'	'db9'	'db9'	'db9'	'db9'
	EBWSS	'db8'	'db13'	'db10'	'db6'	'db5'	'db11'	'db7'
	KBWSS	'db3'	'db3'	'db12'	'db6'	'db22'	'db2'	'db19'
	WEBWSS	'db6'	'db4'	'db2'	'db15'	'db6'	'db9'	'db4'

Admittedly, significant advancement of the proposed scheme, WEBWSS, in the denoising of PD signals with single pulse cannot be directly seen from Figure 4-9 and Figure 4-10. More parameters, i.e., ME, MSE, and XCORR, are listed in Table 4-4 for further performance evaluation of WEBWSS. Two columns in Table 4-4, Improvement 1 and Improvement 2, have presented the improvements by WEBWSS. Improvement 1 is the improved ratio (%) of the use of WEBWSS to CBWSS and Improvement 2 is the improved ratio (%) of the use of WEBWSS to EBWSS. Significant improvements can be seen from these figures of Improvement 1 and Improvement 2 for ME, MSE, and XCORR. The underlying meaning of the improvement of ME is PDs with small magnitude may be picked up by the use of WEBWSS as compared to the other two schemes. This enhanced capability of PD detection has been verified through the application of WEBWSS in on-site PD data, which will be presented in Chapter 8. The improvement of MSE and XCORR indicates that less distortion of the denoised signals can be achieved through WEBWSS. It is good for the accuracy of PD location. In addition, SNR and the parameters in Table 4-4 verify the conclusion presented in [88] that EBWSS outperforms CBWSS in PD denoising.

In the attempt to fully evaluate the performance of the new WEBWSS, PD signals buried in various noise levels are investigated. SNR are set to -8, -6, -4, -2, 0 and 2 for this investigation, representing different noise levels. The parameters used to evaluate the performance of three different schemes are delineated in Figure 4-11 and Figure 4-12 for ns_1 and ns_2, respectively. From the trends of ME, MSE, and XCORR for all denoised signals with the increase of SNR (the higher SNR, the lower noise level is), the new WEBWSS is better than the existing schemes for

wavelet-based denoising of PD detection under various noise levels, and the performance is particularly good when the SNR is low.

Table 4-4 Parameters used to evaluate the performance of WSSs.

Signal	Parameter	CBWSS	EBWSS	WEBWSS	Improvement 1 (%)	Improvement 2 (%)
s_1	**ME**	0.3701	0.3457	**0.3071**	17	11.2
	MSE	0.0055	0.0028	**0.0024**	56.4	14.3
	XCORR	0.8689	0.9358	**0.9495**	8.5	1.4
s_2	**ME**	0.5686	0.5673	**0.5238**	7.9	7.7
	MSE	0.0054	0.0044	**0.0041**	24.1	6.8
	XCORR	0.8757	0.9182	**0.9414**	7	2.5

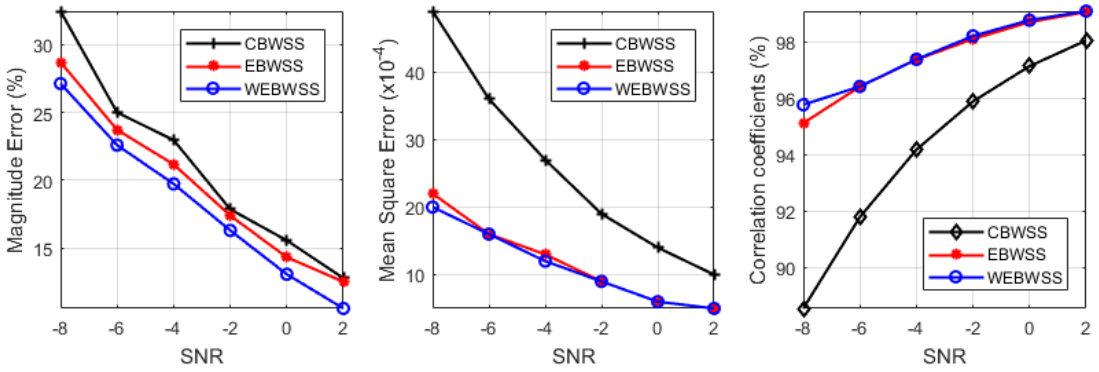

Figure 4-11 ME, MSE, and XCORR between s_1 and denoised s_1 using various WSSs

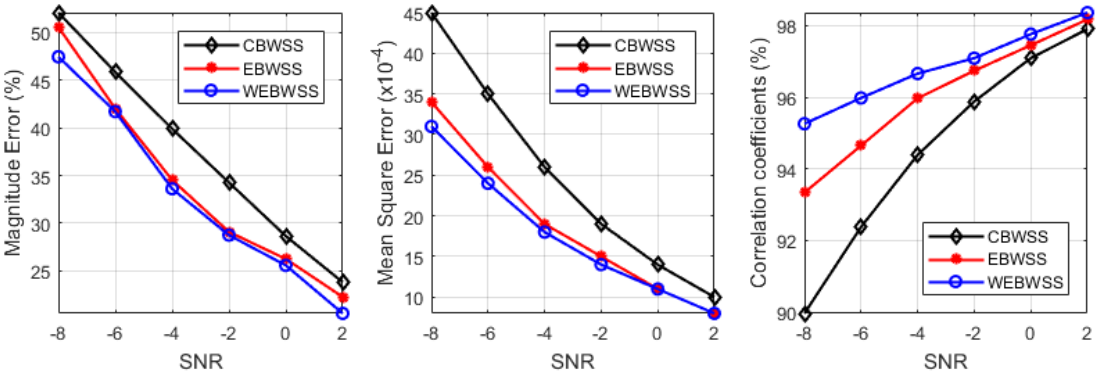

Figure 4-12 ME, MSE, and XCORR between s_2 and denoised s_2 using various WSSs

Although advances of the novel WEBWSS in denoising of PD signals with single pulse have been demonstrated, it is necessary to apply the new method to PD signals with multiple pulses, and then, its performance can be evaluated close to a real case. Tow simulated PD signals,

named spd_1 and spd_2, with 20000 sampling points, together with their noisy versions, $nspd_1$ and $nspd_2$ (SNR = -10) are used for the test of WEBWSS. In spd_1 and spd_2, 5 DEPs and DOPs are randomly distributed along the time scale. The denoising results of $nspd_1$ and $nspd_2$ through three WSSs are illustrated in Figure 4-13 and Figure 4-14, respectively.

In Figure 4-13 and Figure 4-14, the advantages of WEBWSS can be clearly seen from the denoised $nspd_1$ and $nspd_2$. The figures in the top right corner of Figure 4-13 (c) – (e) and Figure 4-14 (c) – (e) are the SNRs of denoised PD signals using three different selection schemes. It directly demonstrates that the application of WEBWSS in PD denoising is better than the existing two schemes. Table 4-4 demonstrates that the new selection scheme has the capability to extract PD signals with small magnitude. The denoised results illustrated in Figure 4-13 and Figure 4-14 are in agreement with that demonstration. It can be seen in Figure 4-13 that two PD pulses with low magnitude can be extracted through WEBWSS (highlighted with red line), while one of them is missing through EBWSS and both of them are missing through CBWSS. Similar situation occurs in Figure 4-14. Also, it is interesting to find that each extracted PD pulse of the denoised $nspd_1$ and $nspd_2$ using WEBWSS suffers less distortion and has higher magnitudes than EBWSS and CBWSS. Further investigation has been implemented through comparisons of the parameters, i.e., MSE and XCORR, results are listed in Table 4-5. These parameters also prove that WEBWSS outperforms EBWSS and CBWSS in denoising of PD signals with multiple pulses.

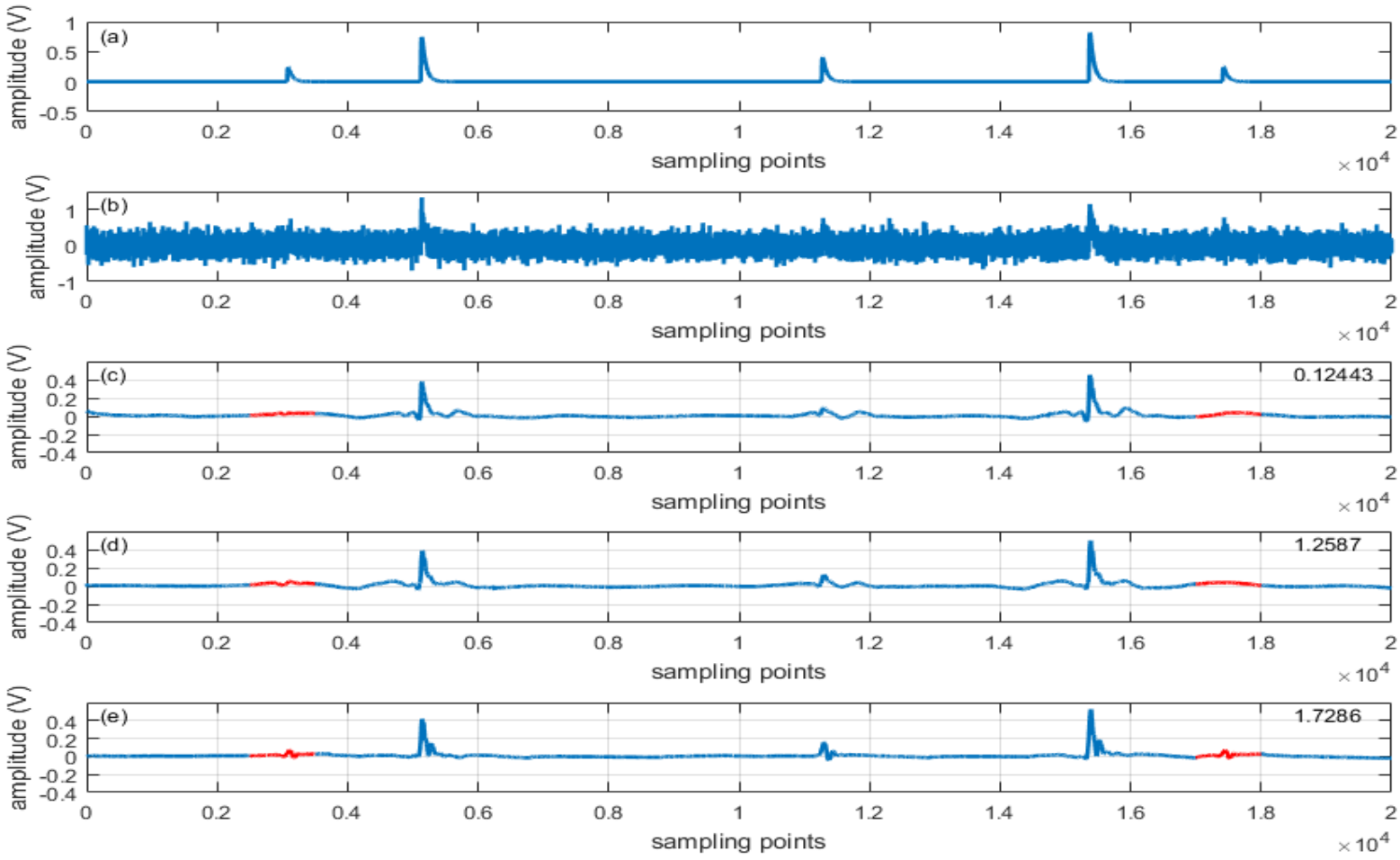

Figure 4-13 (a) spd_1, (b) $nspd_1$, denoising results using (c) CBWSS, (d) EBWSS, (e) WEBWSS

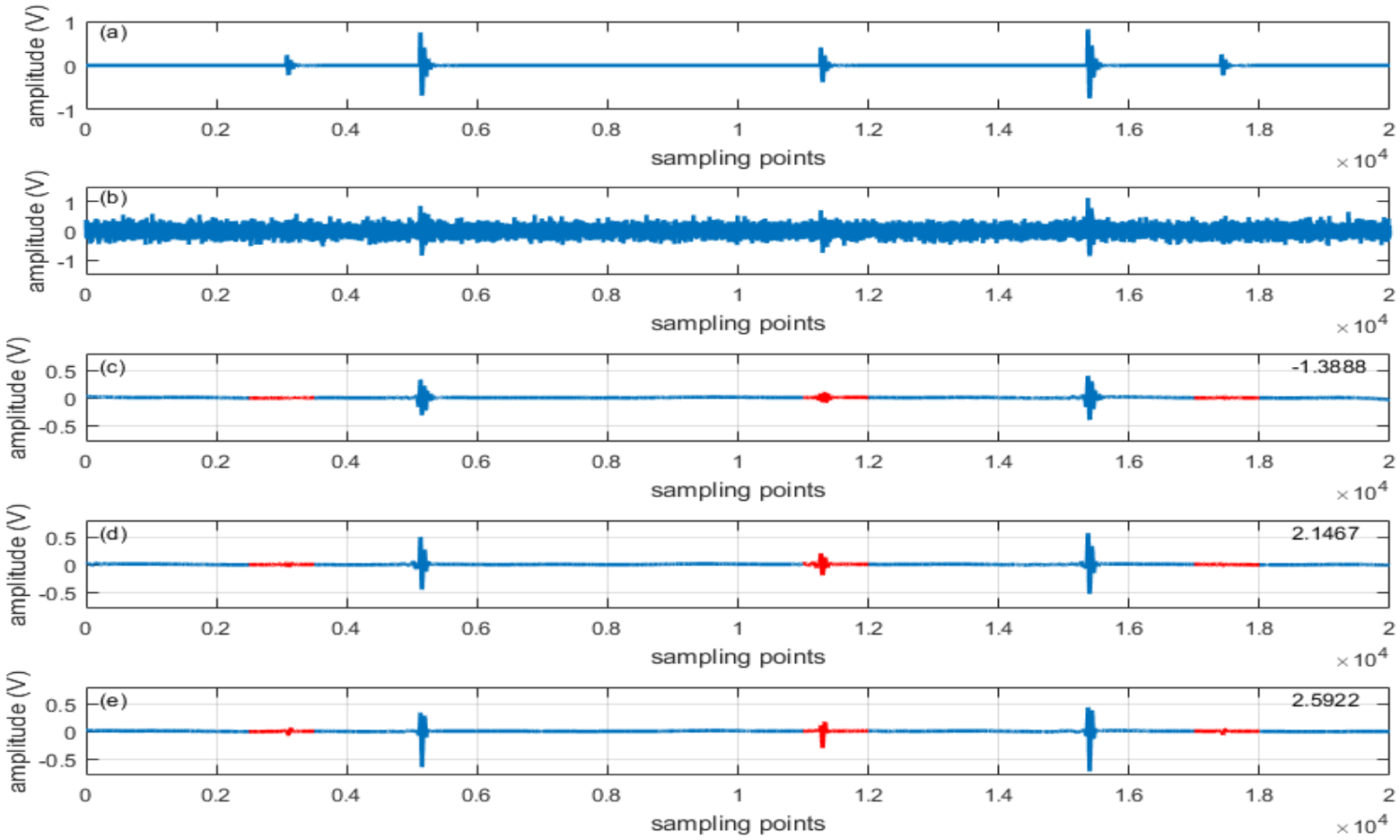

Figure 4-14 (a) spd_2, (b) $nspd_2$, denoising results using (c) CBWSS, (d) EBWSS, (e) WEBWSS

4.6.3 PD Signals Corrupted by White Noise and DSI

The novel wavelet entropy-based wavelet selection scheme is proposed for the suppression of white noise. White noise is not the only noise source that affects PD measurement. In this section, another very common noise source, i.e., DSI, will be involved to test the proposed WSS.

The DSI with 2048 sampling points, as shown in Figure 4-15 (the same as that shown in Figure 2-28), is simulated by the use of (2-9) and parameters in Table 2-2 for this investigation. Table 4-6 lists the entropy value of approximation coefficients at each decomposition scale of DSI. As did in Table 4-2, 'db3' and 'db10' are used for DSI decomposition with 6 scales. Based on the figures of entropy value at each decomposition scale in both Table 4-2 and Table 4-6, difference between PD signals, i.e., DEP and DOP, and DSI is not as big as that between PD signals and white noise. Although the entropy value of DSI at each decomposition scale still indicates the decomposed DSI has relatively high degree of randomness or disorder, it is not very close to its theoretical maximum value. This could be a potential reason that the suppression of DSI using WEBWSS is not as good as that for white noise.

Figure 4-16 illustrates the simulated PD signals s_1 and s_2, as shown in Figure 4-8 (a) and (c), with the DSI and white noise. In both simulated signals, the DSI is added with magnitude of 0.1, and white noise has a mean value of zero and standard deviation of 0.3. The SNRs of noisy s_1 and s_2, based on (4-16), are -6.22 and -8.06, respectively. The novel scheme, WEBWSS, is applied to both noisy PD signals, and the selected wavelet at each scale is listed in Table 4-7. The denoising results of noisy s_1 and s_2 are delineated in Figure 4-17. Based on the denoising results, both DSI and white noise can be reduced by the new WEBWSS. Note that the peak values of both denoised s_1 and s_2 are lowered after denoising. This also can be seen from Table 4-8, the ME figures in which are used to evaluate the denoising performance of WEBWSS. The effective reduction of DSI from noisy s_1 and s_2 may be due to its small magnitude, which will be increased for further investigation. The magnitude of DSI is then increased to 0.3 while white noise is remained as the same. The corresponding

SNRs of noisy s_1 and s_2 decrease to -6.59 and -8.43, respectively. The denoising results of these simulated noisy PD signals are illustrated in Figure 4-18. As highlighted with red circles in Figure 4-18, DSI cannot be completely reduced by WEBWSS when its magnitude increases. The critical value of the magnitude of DSI involved in noisy PD signals is 0.3 after investigation. It means when the figure exceeding this critical value, DSI cannot be effectively suppressed by the novel WEBWSS. A conclusion can be drawn based on this investigation, i.e., WEBWSS is effective for the reduction of white noise and DSI in PD measurement, but the prerequisite is that the DSI level in noisy PD signals cannot be high.

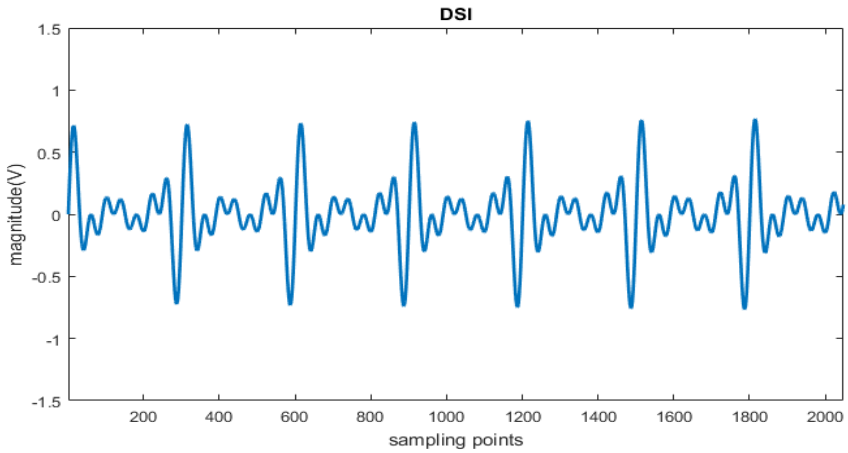

Figure 4-15 Simulated DSI with 2048 sampling points

Table 4-5 Parameters used for performance evaluation of various WSSs

		CBWSS	EBWSS	WEBWSS
spd_1	**MSE**	0.0012	0.0011	**0.0009**
	XCORR	0.8327	0.8516	**0.8791**
spd_2	**MSE**	0.0009	0.0006	**0.0005**
	XCORR	0.8246	0.8823	**0.8847**

Table 4-6 Wavelet entropy values of approximation coefficients of DSI with 6-scale decomposition

	DSI ('db3')		DSI ('db10')	
	Entropy value	**Max**	**Entropy value**	**Max**
a_1	*5.6677*	6.9334	*5.6842*	6.9402
a_2	*4.9764*	6.2442	*5.0233*	6.2653
a_3	*4.2878*	5.5607	*4.3866*	5.6058
a_4	*3.6590*	4.8828	*3.8988*	4.9767
a_5	*3.2023*	4.2195	*3.5549*	4.4067
a_6	*2.4675*	3.5835	*2.8076*	3.9120

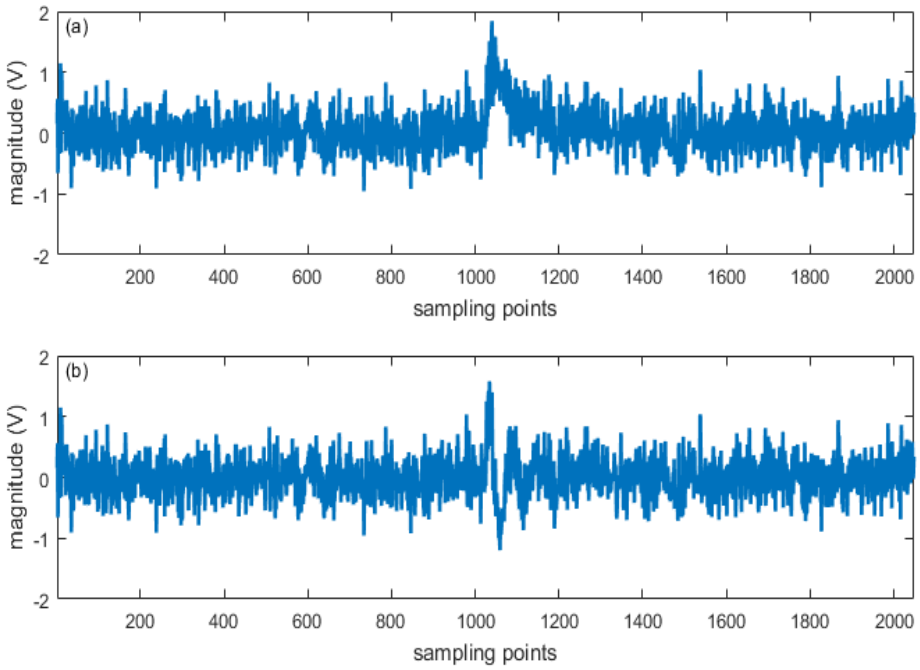

Figure 4-16 (a) noisy s_1 with DSI and white noise, (b) noisy s_2 with DSI and white noise

Table 4-7 Selected wavelet at each decomposition scale using WEBWSS for noisy s_1 and noisy s_2

	Decomposition Scale						
	1	**2**	**3**	**4**	**5**	**6**	**7**
s_1	'db2'	'db2'	'db4'	'db2'	'db5'	'db5'	'db2'
s_2	'db2'	'db2'	'db3'	'db11'	'db5'	'db6'	'db4'

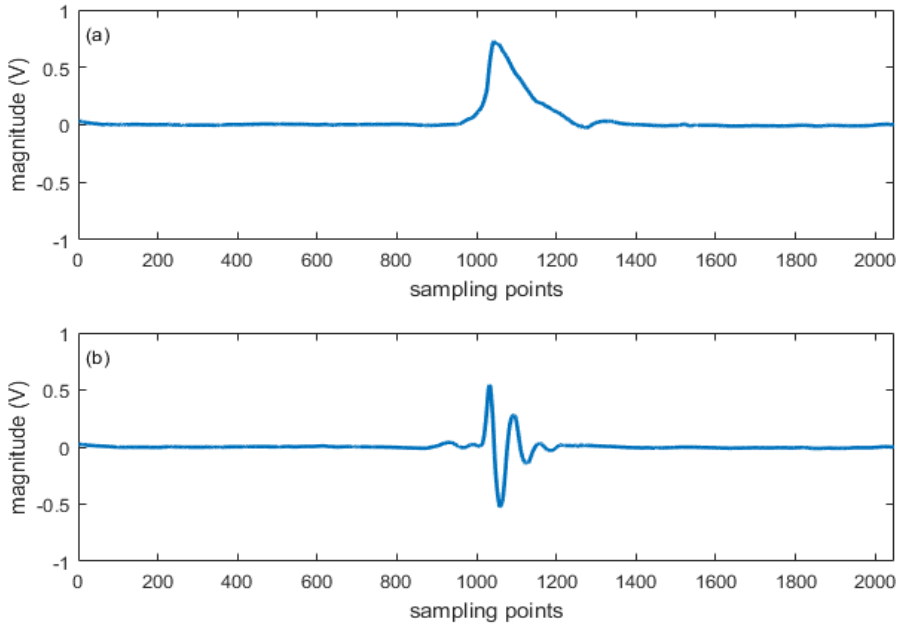

Figure 4-17 (a) denoised noisy s_1 by WEBWSS, (B) denoised noisy s_2 by WEBWSS

Table 4-8 Parameters used for performance evaluation of WEBWSS in noisy s_1 and noisy s_2

	Parameters			
	ME	**MSE**	**XCORR**	**SNR**
s_1	0.2698	0.0019	0.9655	11.1701
s_2	0.4548	0.0035	0.9417	6.1361

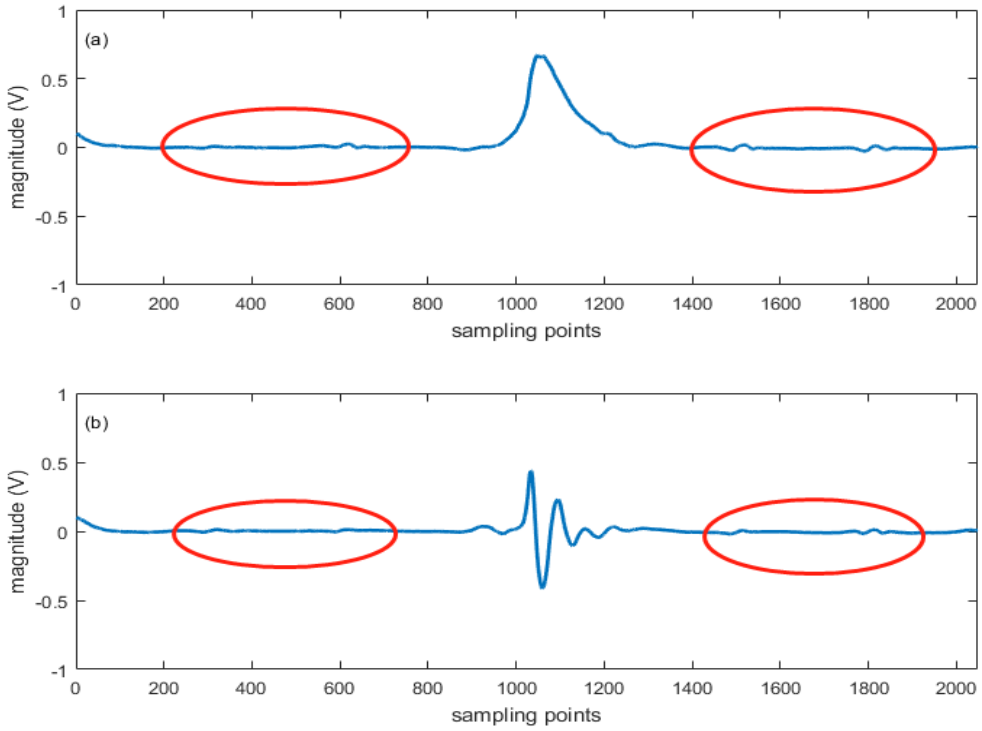

Figure 4-18 High-level DSI in noisy PD signals: (a) denoised noisy s_1 by WEBWSS, (b) denoised noisy s_2 by WEBWSS

4.7 Conclusion

Wavelet selection constitutes a critical step in wavelet-based denoising. Analysis of current existing wavelet selection schemes, i.e., correlation-based and energy-based wavelet selection schemes, revealed their limits for practical uses in this chapter. To improve the accuracy of wavelet selection and provide an alternative to current existing schemes, further study of characteristics of both PD signals and noise was accomplished by the investigations of the high-order statistics and the degree of disorders of their transformed signals through wavelet-based technique. As a result, kurtosis-based (high-order statistics) and wavelet

entropy-based (degree of disorders) wavelet selection schemes were proposed. Simulated PD signals were used for performance test of these new schemes in wavelet-based PD denoising. Results show that the wavelet selected by wavelet entropy-based scheme can present better performance than the others, especially for situations when signal to noise ratio is unity or less. This scale-dependent, automated, and data-driven scheme enables it a promising technique in PD detection. Accordingly, it has the potential to extend the range of PD detection in cables through the use of this new technique.

The analysis of statistical characteristics of PD signals was based on simulated PD signals in this chapter. It was investigated under the same environment as that used in correlated-based and energy-based selection scheme. This new wavelet selection scheme may have limits as the signals used for investigations are ideal in some sense. As aforementioned, the shapes or waveforms of PD signals vary due to the factors, such as signal attenuation and noise level. The suitability of this entropy-based wavelet selection scheme for real PD signal denoising requires the involvement of real PD signals that can have as many waveforms as possible. In Chapter 8, several real PD signals will be used for this purpose, the range of waveform variations may still be far less than those in real situation.

5 Decomposition Scale Determination for Wavelet-based PD Denoising

5.1 Introduction

Wavelet transform has been widely investigated in the area of PD denoising in recent decades due to its advance for the time and frequency representation of signals. As known, the effectiveness of wavelet-based denoising is dictated by the choice of wavelet, decomposition scale, and noise estimation. The choice of wavelet has been fully discussed in last chapter. As a series of improvement of wavelet-based denoising in the area of PD detection, the effect of decomposition scale on the denoising results will be investigated in this chapter.

Decomposition scale plays a significant role in wavelet-based denoising. For a predefined decomposition scale, the noisy PD signal is decomposed from fine scale (high frequency bands) to coarse scale (low frequency bands). The purpose of this is to have sufficient frequency resolutions at coarse scales, and thus, low frequency interferences can be effectively removed. If the scale is too small, it is really difficult to separate the mixed PD signals and low frequency interferences, if the scale is too large, it takes longer computational time and some information of PD signals of interest may be discarded as well. Theoretically, decomposition scale is dependent on the length of the detected PD signals. In practice, it is often determined by trail and errors or experience [12], [21], [23], [119].

Optimal decomposition scale is a desirable property for the application of WT in signal denoising, which is not limited to its use in the area of PD denoising. In [168], an entropy-based selection of optimal decomposition scale was introduced for WT in hydrologic series data. It is a method to measure the entropy differences between the decomposed signal and noise which is artificially generated based on the probability distribution. Optimal decomposition scale is determined when the resultant entropy of the decomposed signal is significantly different from that of noise. Prior knowledge of noise is required as the probability distribution of noise is an essential part of this method. It is not the case in practice since noise is random in nature. In an attempt to overcome the limit of the

entropy-based method, Lei *et al* [169] proposed an alternative method to determine decomposition scale based on sparseness measurement in the area of bio-signal (sparseness measurement is referred to as this method in this chapter). It is known that the basic idea behind wavelet-based denoising is that the wavelet transform can provide a sparse representation for many real-world signals [163]. In this method, the degree of sparseness is quantified for the decomposed signals by evaluating the percentage of the number of zero/near-zero wavelet coefficients. An empirical parameter is adopted in the count of zero/near-zero coefficients. Also, the curve of sparseness of the decomposed signal is monotonically increasing, a suitable threshold of sparseness is required to choose the best decomposition scale. Obviously, the requirement of an empirical parameter and a proper threshold makes this method less robust and more subjective. In relation to PD denoising, Zhou *et al* in [119] claims that a noisy PD signal can be decomposed with a maximum decomposition scale which is determined by both the length of PD signals and the length of the wavelet filter being used (wavelet length is referred to as this method in this chapter). The involvement of the length of the wavelet filter means that a fixed wavelet needs to be applied during the entire decomposition and reconstruction of PD signal denoising. As discussed in last chapter, a scale-dependent wavelet can provide better denoising results. In [170], determination of decomposition scale was based on the energy spectral density of real PD signals. It requires prior knowledge of pure PD signals. However, PD signals vary due to the type and location of PD defects, propagating medium and path, and the detecting circuit. Also, PD signals are normally contaminated by background noise. The variation of PD signals and noise limits this method to be used for online PD monitoring systems.

Based on the discussion above, it is necessary to introduce an automated method to select an optimal decomposition scale without the requirement of any prior knowledge of PD signals or noise as well as the involvement of any empirical parameter and threshold. As such, the new method can provide more accurate denoised PD signals with possibly less computational time. It is one of the motivations for the author in this research project. As a result, a novel method for decomposition scale determination is proposed in this chapter based on the concepts of Shannon entropy in information theory [159] and sparsity in compressive sensing [171].

Wavelet entropy is a relatively new concept derived from Shannon entropy. For a decomposed signal, its wavelet entropy value is highly correlated to the number of wavelet coefficients that preserves energy of the original signal. A smaller wavelet entropy value indicates that energy of the original signal is preserved in fewer number of wavelet coefficients. As such, wavelet entropy can be used to measure the number of significant coefficients in the decomposed signals through WT [172]. As mentioned above, sparsity is the basic idea underlying wavelet-based denoising technique. For a sparse signal, its sparsity defined in [171] is inversely proportional to the number of significant elements. In this thesis, sparsity is reflected by the ratio of the number of significant coefficients to the total number of the wavelet coefficients. Based on this, wavelet entropy is applied to quantify the number of significant coefficients of decomposed PD signal at each decomposition level and then resultant sparsity is used to determine the optimal decomposition scale. Simulated PD signals, s_1, s_2, spd_1 and spd_2 used in last chapter, will be applied to demonstrate the performance of this novel method for the determination of decomposition scale. Results show that it is an effective technique for the selection of decomposition scale.

5.2 Methods for Decomposition Scale Determination

A number of methods used to determine the decomposition scale has been analyzed above. Due to the variations of PD signals and noise, the methods based on the probability distribution of noise or energy spectral density of pure PD signals will not be introduced with more details. Wavelet length method and sparseness measurement method will be discussed in depth. Wavelet length method functions as a reference method to the new proposed method for scale selection. The concept of sparsity used for decomposition scale determination is first introduced in sparseness measurement method. The novel sparsity-based method is inspired by this idea and can present better denoising results.

5.2.1 Wavelet Length Method

The theoretical maximum decomposition scale, J_{max}, in DWT is determined by the length of a signal, as J_{max} is equal to $floor(log_2 N)$, where the function $floor(\cdot)$ rounds the value

of $log_2 N$ to the closest integer ($< log_2 N$) and N is the length of the signal. In [119], it points out that the signal length at the highest decomposition scale should be no less than the length of the wavelet filter being used when the wavelet toolbox is employed in MATLAB. As a result, the maximum decomposition scale for the wavelet-based decomposition of a signal is defined by

$$J_m = fix(log_2(N/N_w - 1)) \tag{5-1}$$

where N_w is the length of the wavelet filter being used, and the function $fix(\cdot)$ is to round the value in the bracket to its nearest integer. For example, if the length of a noisy PD signal is 2048 sampling points and the wavelet 'db3' ($N_w = 6$) is used as the mother wavelet, the decomposition scale of this signal based on (5-1) is 8.

The decomposition scale determined by (5-1) requires a fixed wavelet to be applied in PD signal denoising. However, it has been demonstrated that the denoising result of scale-dependent wavelet is better than scale-independent one. Also, the length of a detected PD signal in practice is often long (millions of sampling points, see in Chapter 8) due to high sampling rate. If a short-length wavelet, e.g., 'db3', is selected for PD denoising, then the decomposition scale determined by (5-1) is very close to its theoretical maximum decomposition scale. In this case, the denoising is time consuming as compared to a small optimal decomposition scale.

5.2.2 Sparseness Measurement Method

Sparseness measurement method is proposed based on the basic idea of wavelet-based denoising, i.e., the coefficients of DWT are generally sparse. The degree of sparseness is quantified by the percentage of the number of zero/near-zero coefficients among the entire transformed coefficients. The degree of sparseness is defined by [169]

$$sp = \frac{N_z}{N - 1} \tag{5-2}$$

where N_z is the number of zero/near-zero coefficients, and N is the length of the original signal. It can be seen that the quantity of N_z plays a key role in (5-2). In sparseness measurement method, an empirical parameter $K \in [5,10]$ is employed to quantify the number of zero/near-zero coefficients. The amplitude of the DWT coefficients less than or

equal to $1/K$ of the largest coefficient at each decomposition scale, i.e., $d_{j,i} \leq max(d_{j,i})/K$, is counted as a zero/near-zero coefficient.

Figure 5-1 illustrates the sparseness of decomposed signals noisy s_1 and s_2 with SNR = -5. An increase of the value of sparseness with the increasing decomposition scale can be seen in Figure 5-1. The increasing rate of the value of sparseness decreases with the increasing decomposition scale, and the value approaches the theoretical maximum value 1 asymptotically. From the distribution of sparseness along the entire signal decomposition, it can be seen that the statistical characteristics of both transformed noisy s_1 and s_2 do not vary much after the 5th decomposition. A suitable threshold is then selected for the determination of optimal decomposition scale.

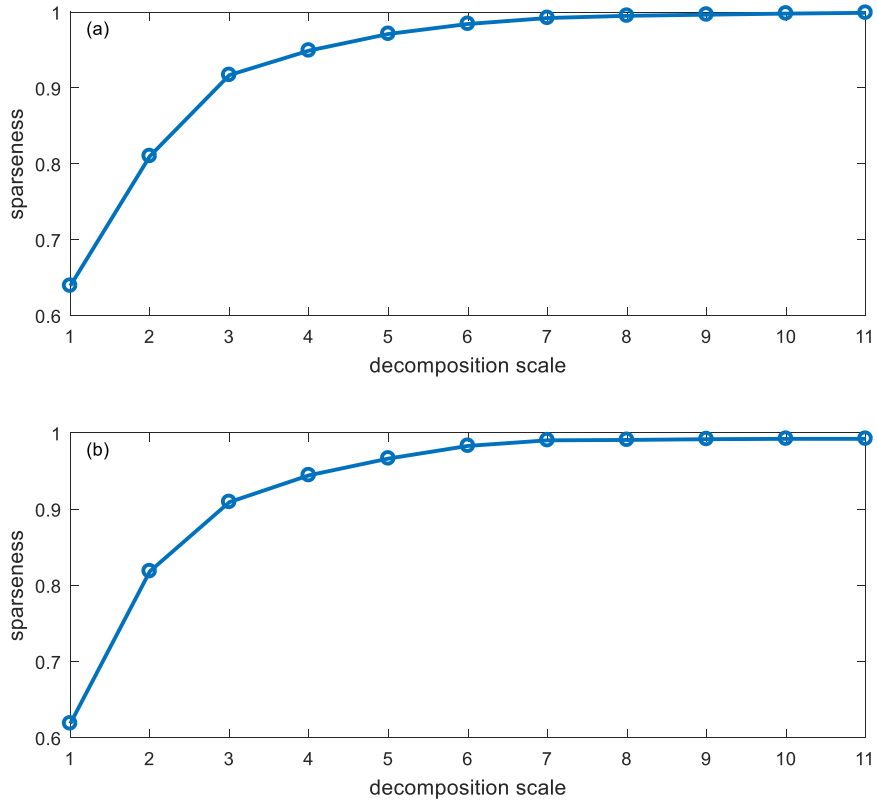

Figure 5-1 Sparseness of the decomposed signal (a) noisy s_1, (b) noisy s_2

The criterion for decomposition scale determination in sparseness measurement method is briefly introduced above. It is not a robust selection criterion due to the involvement of empirical parameters, which determine the quantity of zero/near-zero coefficients and the

threshold of sparseness for the selection of optimal decomposition scale. However, it is a good idea that sparseness of the transformed signal can be used for decomposition scale determination. It can be improved as an adaptive or data-driven method if the number of zero/near-zero coefficients or significant coefficients can be quantified based on the statistical characteristics of the signal.

5.3 Sparsity-based Decomposition Scale Determination

5.3.1 Significant Coefficients

The concept of wavelet entropy has been introduced in 4.5.1 and applied for the selection of optimal wavelet for PD denoising. Given a signal x of length N, the optimal wavelet is the wavelet that can minimize the cost of representing x, i.e., the wavelet that can minimize the entropy, as discussed in last chapter. If the coefficients of transformed x is put into decreasing order of magnitude, the rate at which the coefficients decrease controls the theoretical dimension N_0 [161]. N_0 is a number between one and N describing the number of significant coefficients. Generally, N_0 can be defined in several ways. The simplest is to count the number of the coefficients with magnitude exceeding certain threshold. Another is to exponentiate the entropy of the coefficient sequence. N_0 obtained based on the coefficient sequence itself is more desirable than certain threshold if it is involved in the development of an adaptive and data-driven method for decomposition scale selection. Also, the concept of wavelet entropy in [172] was adopted to measure the number of significant wavelet coefficients N_0 through wavelet expansion of a signal x. N_0 can be calculated by the equation below [172],

$$N_0 = e^H \tag{5-3}$$

where $H = WE$ and WE is a wavelet entropy calculated based on (4-13). The value of N_0 also means the number of the largest wavelet coefficients that needs to be preserved for an adequate representation of x. It has been demonstrated in Figure 5-2, which illustrates s_1 and s_2 with their corresponding N_0 representation in (a) and (b), respectively. Both PD signals are decomposed into 3 scales with the wavelet selected by WEBWSS. Parameters, e.g., XCORR and MSE, are used to investigate the difference between PD signals and the

corresponding N_0 representations, which are listed in Table 5-1. The results are in agreement with the statement above, i.e., the original signal can be sufficiently represented by N_0 coefficients.

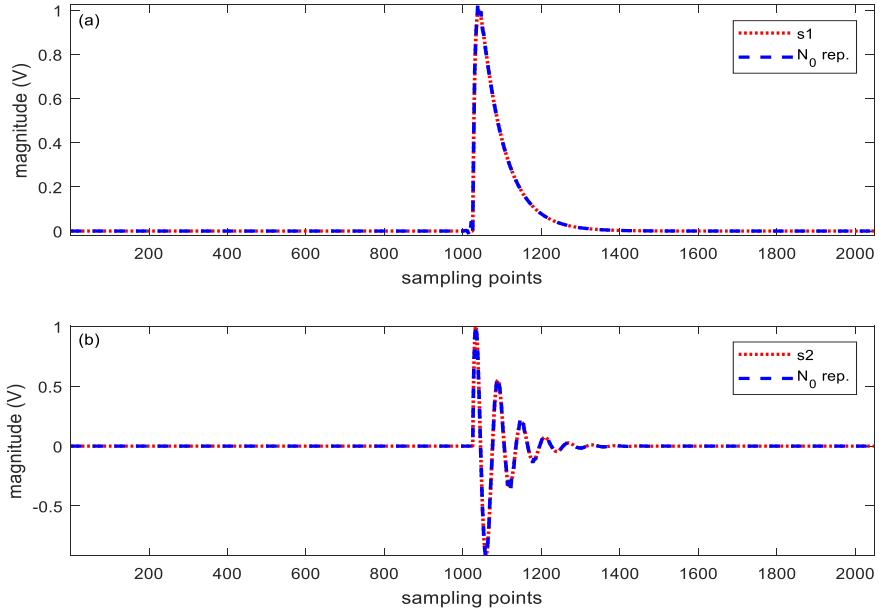

Figure 5-2 (a) s_1 and its N_0 representation, (b) s_2 and its N_0 representation

Table 5-1 Parameters for the difference investigation between PD signals and their N_0 representations

	MSE	XCORR
$\boldsymbol{s_1}$ & $\boldsymbol{N_0}$ ***rep.***	5.6×10^{-6}	0.9999
$\boldsymbol{s_2}$ & $\boldsymbol{N_0}$ ***rep.***	1.6×10^{-5}	0.9994

5.3.2 Sparsity

Sparsity is the prerequisite for the application of compressive sensing in signal processing. Mathematically, for a given *N*-sample signal $[X]$, it is *k*-sparse if it has at most *k* non-zeros, i.e., $\|X\|_0 \leq k$ [173]. The definition is also extended to a nearly sparse signal, that is, X is *k*-sparse if *k* elements are remarkably larger than the other $N-k$ elements. Generally, signals are not sparse themselves, but they may have sparse nature when they are represented in some domains, e.g., wavelet domain. In this case these signals are also referred to as *k*-sparse [173].

The quantification of sparsity is given by [171]

$$sp = (1 - \frac{k}{N}) \times 100\% \tag{5-4}$$

where sp is the sparsity value. In fact, the definition of sparsity in (5-2) and (5-4) are the same in essence but represented by different equations.

In relation to the signal representation in wavelet domain, $\{d_{j,i}\}$ represents the detail coefficients of the j^{th} scale with the length of n_j, and thus, $\mathrm{i} = 1,2,\ldots,\mathrm{n_j}$. Based on the definition of sparsity, $\{d_{j,i}\}$ is sparse if k coefficients in $\{d_{j,i}\}$ are larger than the rest, and the sparsity of $\{d_{j,i}\}$ can be expressed by

$$sp_j = (1 - \frac{k}{n_j}) \times 100\% \tag{5-5}$$

5.3.3 Sparsity-based Method for Decomposition Scale Selection

An inappropriate decomposition may lead to either ineffective reduction of noise (low scale) or distortion of the signal of interest, together with long computational time (high scale). It is therefore necessary to propose a new method that can select the best decomposition scale in an automated way. The combination of the concepts of wavelet entropy and sparsity discussed above sheds light on this. Note that the concept of sparsity will not be directly applied in the new proposed method for scale determination. A new concept, significant coefficients ratio (SCR), will be used to reflect sparsity of the decomposed signals. For an arbitrary signal, N_0 denotes the number of significant wavelet coefficients that can adequately represent it in the wavelet domain. The SCR is then defined by

$$SCR_j = \frac{N_0}{n_j} \tag{5-6}$$

The inverse relationship between sparsity and SCR can be seen from (5-5) and (5-6). The sparsity reaches its maximum value when SCR is its minimum value. Substituting N_0 from (5-3) into (5-6) yields

$$SCR_j = \frac{e^H}{n_j} \tag{5-7}$$

Based on (5-7), SCR reaches its minimum value if e^H is minimum, i.e., the wavelet entropy is minimum. Smaller wavelet entropy value is desirable since it indicates that fewer number of wavelet coefficients that preserve energy as well as less randomness of the signal [161].

SCR in (5-7) can reflect the sparsity of detail coefficients of a transformed signal and enables the identification of the presence of noise in detail coefficients. A large SCR means that noise is presented with a large number of coefficients with small magnitudes, and a small SCR implies that signal is presented with only a few coefficients with large magnitudes. Table 5-2 lists the SCR of transformed s_1, s_2 and white noise with 2048 sampling points. The difference of SCR figures among simulated PD signals and white noise can be seen from Table 5-2. For pure PD signals, the SCR at each decomposition scale is less than 0.1, which means PD signals are presented with only a few large coefficients. In contrast to PD signals, the SCR of white noise at each decomposition scale is larger than 0.4, indicating that noise is presented with a large number of coefficients. After numerous simulations of PD signals and white noise as well as their wavelet-based transforms, it is interesting to find that a detail coefficient sequence at certain scale is noise only if its SCR is larger than 0.4 and is signal only if its SCR is less than or equal to 0.1.

Table 5-2 SCR of transformed s_1 and s_2 with white noise at each decomposition scale

	SCR of Transformed Signals		
	$\boldsymbol{s_1}$	$\boldsymbol{s_2}$	$\boldsymbol{wn}$
$\boldsymbol{d_1}$	0.0009	0.0019	0.4552
$\boldsymbol{d_2}$	0.0039	0.0076	0.4680
$\boldsymbol{d_3}$	0.0077	0.0110	0.5385
$\boldsymbol{d_4}$	0.0152	0.0207	0.4773
$\boldsymbol{d_5}$	0.0294	0.0610	0.4559
$\boldsymbol{d_6}$	0.0556	0.0400	0.6111
$\boldsymbol{d_7}$	0.10	0.0588	0.6500
$\boldsymbol{d_8}$	0.0833	0.0954	0.4167

Based on the analysis above, SCR can be adopted as a criterion for the determination of an appropriate decomposition scale. The general process of this novel method for decomposition scale determination is described as follows,

1. Given an *N*-sample noisy PD signal, ns, the maximum decomposition scale J_{max} is given by $floor(log_2 N)$,
2. DWT of ns generates the detail coefficients $\{d_j\}$ at scale j, $j = 1,2,\ldots,J_{max}$,

3. Based on (5-7), the SCR of each $\{d_j\}$ can be calculated,
4. The selection of an appropriate decomposition scale needs to follow the rules below,
    a) Select the decomposition scale *DS* at which the detail coefficient sequence has the minimum SCR;
    b) If SCR at the selected DS is less than or equal to 0.1, then $DS = DS - 1$.

Figure 5-3 is a flow chart of the general process of the selection of an appropriate decomposition scale for the wavelet expansion of a noisy PD signal.

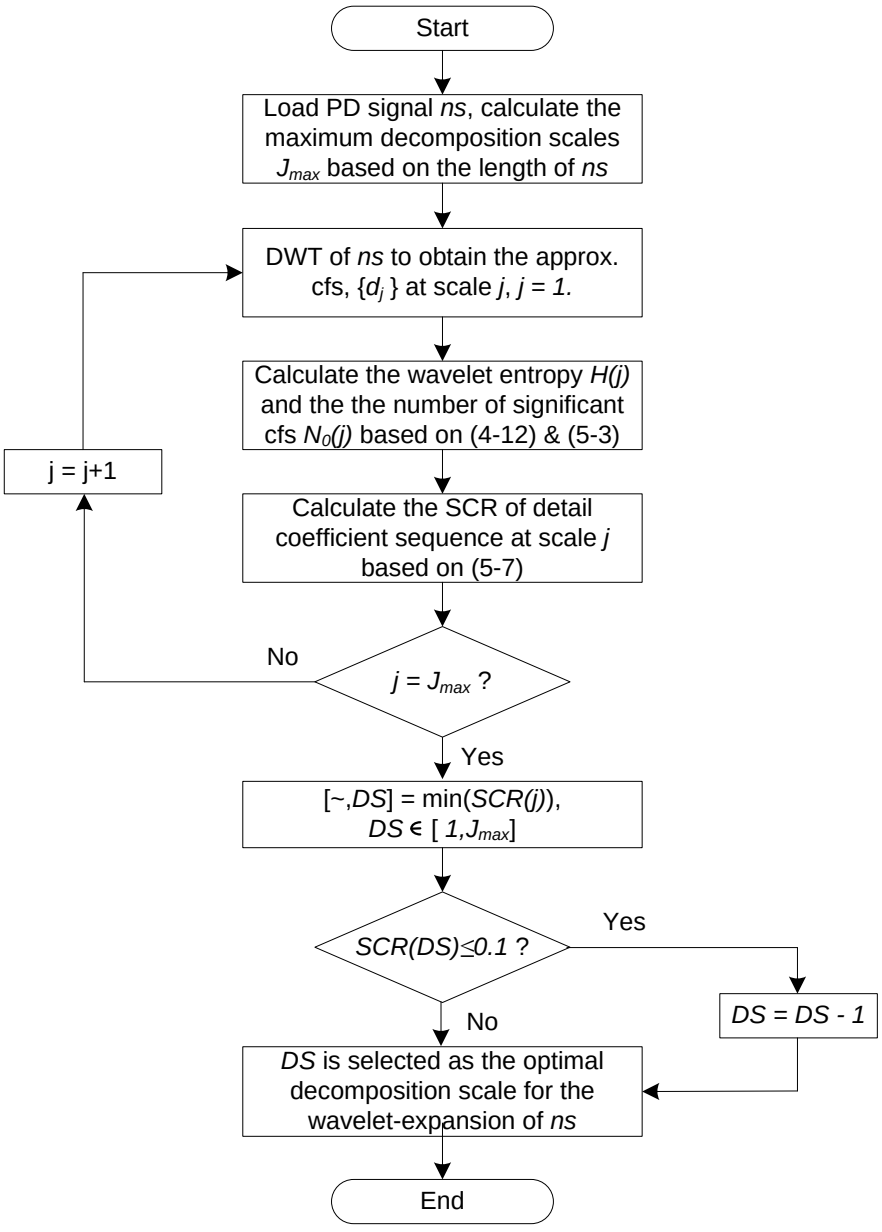

Figure 5-3 Flow chart of the general process of decomposition scale determination

5.4 Performance Evaluation of New Decomposition Scale Determination

Parameters introduced in Section 4.6.1 will be applied to evaluate the performance of sparsity-based method for decomposition scale determination in wavelet-based PD denoising. The new method is initially proposed based on simulated PD signals contaminated by white noise, as did for WEBWSS. The performance evaluation is then extended to PD signals with the involvement of DSI.

5.4.1 PD Signals Corrupted by White Noise

s_1 and s_2 represent typical PD signals with a single pulse, i.e., DEP and DOP. Figure 5-4 delineates s_1 and s_2, and their corresponding noisy signals corrupted by white noise with SNR = -5. As wavelet length method is used as a reference method, the wavelets selected by CBWSS for the wavelet decomposition of noisy s_1 and s_2 are 'db3' and 'db10', respectively.

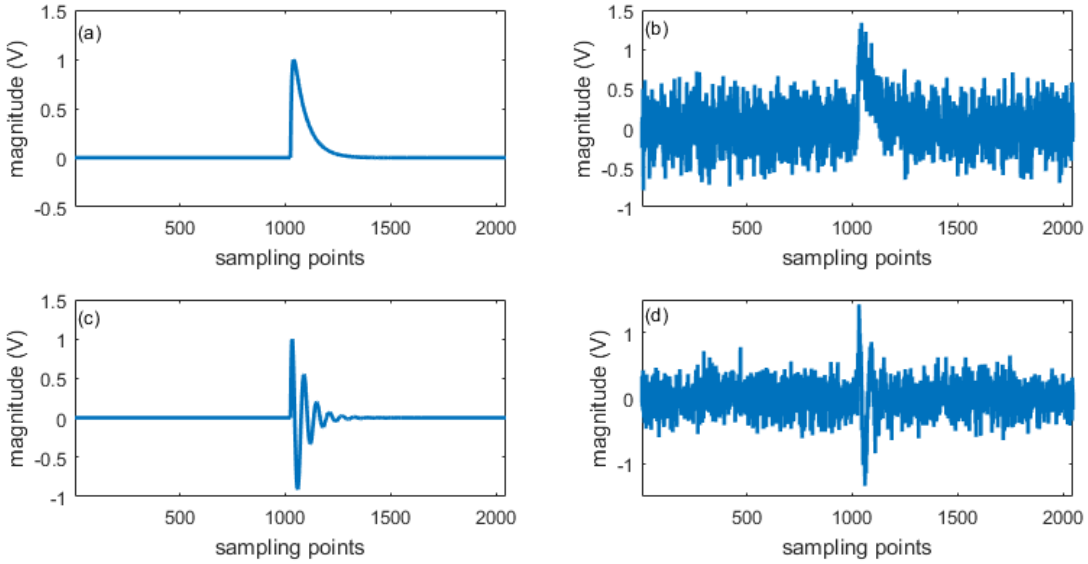

Figure 5-4 (a) s_1, (b) noisy s_1, (c) s_2, (d) noisy s_2

The linked distribution of SCRs of noisy PD signals are illustrated in Figure 5-5 and Figure

5-6. Based on the sparsity-based method for decomposition scale determination, 7 and 5 are the selected decomposition scale for noisy s_1 and s_2, respectively. For wavelet length method, 8 and 7 are the selected decomposition scales based on (5-1) for noisy s_1 and s_2, respectively. To evaluate the performance of different methods for scale selection, Figure 5-7 illustrates the denoising results of noisy s_1 and s_2 with the selected scales mentioned above.

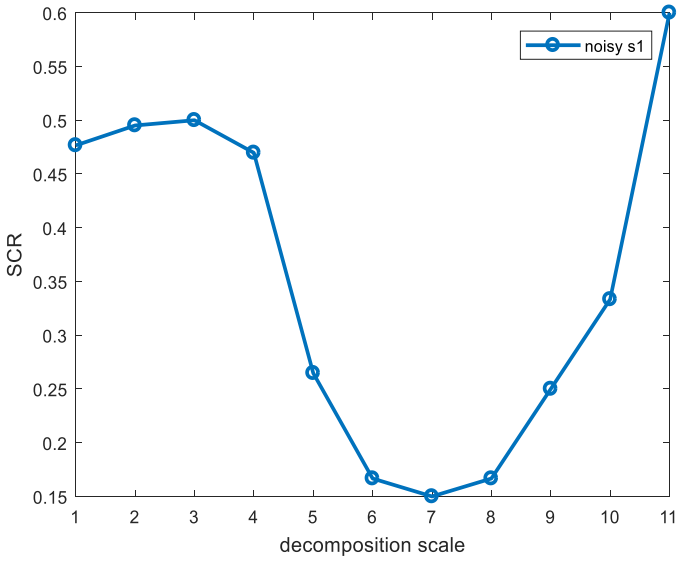

Figure 5-5 SCR distribution of transformed noisy s_1

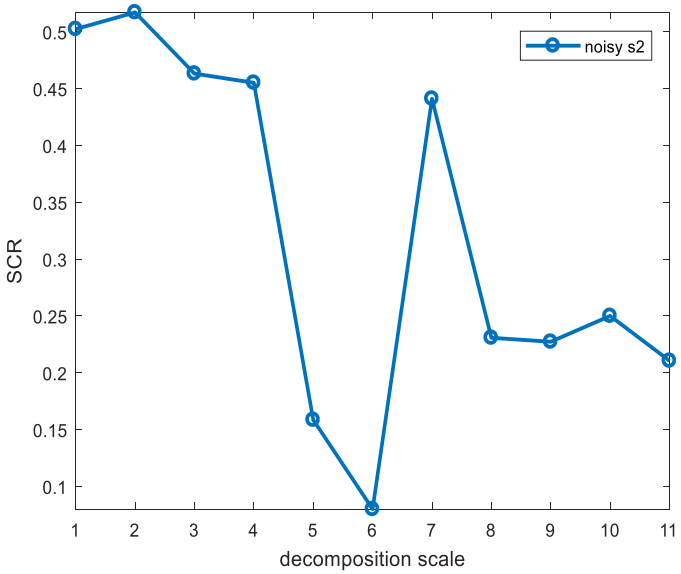

Figure 5-6 SCR distribution of transformed noisy s_2

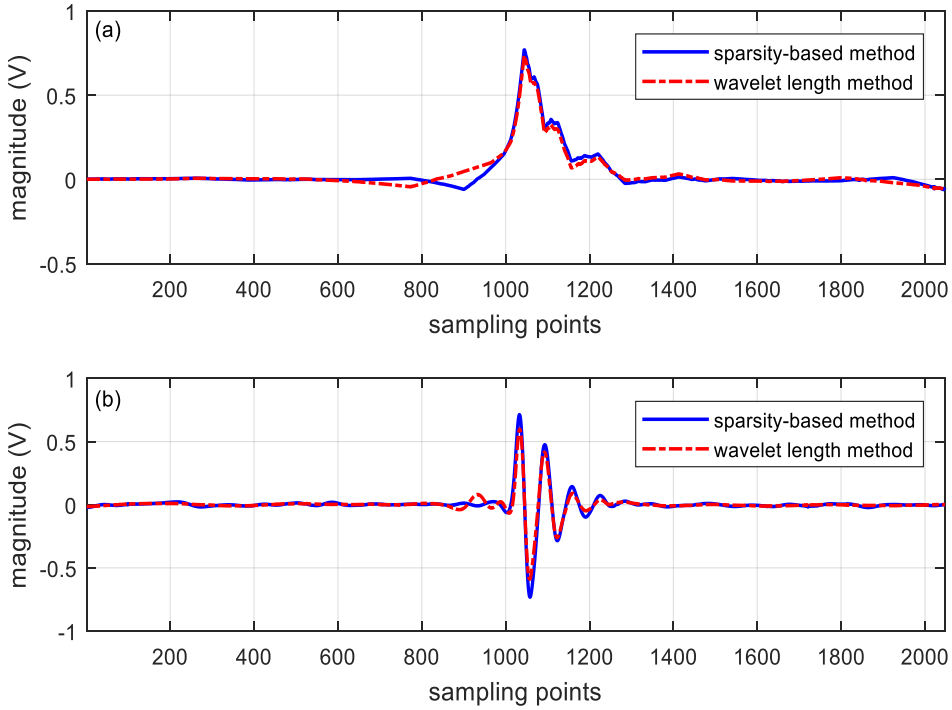

Figure 5-7 Denoising results of (a) noisy s_1, (b) noisy s_2 using different scale selection methods

Better denoising results of both noisy s_1 and s_2 can be seen from Figure 5-7 due to the application of the sparsity-based method for decomposition scale determination. The shapes of denoised PD signals using both methods are similar, but the peak distortion using the new method is reduced as compared to wavelet length method. Parameters regarding the performance evaluation of the sparsity-based method are listed in Table 5-3. Figures in the table are in agreement with the direct observation from Figure 5-7, i.e., the ME of denoised PD signals has been improved when the sparsity-based method is applied. The improvement of ME means that PD pulses with small magnitudes may be successfully extracted from background noise with the application of the new scale selection method. Equally, improvements can be seen from MSE and XCORR of denoised PD signals, which are desirable as relatively lower distortion of PD signals can increase the accuracy for the location of PD defects within insulation systems.

Table 5-3 Parameters used for performance evaluation of different scale selection methods

		Wavelet Length method	Sparsity-based method
s_1	**ME**	0.2650	**0.2308**
	MSE	0.0034	**0.0029**
	XCORR	0.9302	**0.9386**
s_2	**ME**	0.4012	**0.2865**
	MSE	0.0028	**0.0022**
	XCORR	0.9335	**0.9353**

Noisy s_1 and s_2 are processed through one-scale to full-scale wavelet-based denoising to further evaluate the performance of the sparsity-based method for decomposition scale determination. The SNRs after denoising with various scales for noisy s_1 and s_2 are delineated in Figure 5-8 and Figure 5-9, respectively. Based on the SNRs after denoising, 7-scale and 4-scale wavelet-based denoising can obtain the best SNR after denoising for noisy s_1 and s_2, respectively. The scale that can have best denoising result for noisy s_1 is in agreement with the scale selected by the sparsity-based method. The scale selected by the sparsity-based method for noisy s_2 is 5, which is not the scale at which it can obtain the best SNR after denoising. However, the scale selected by the novel method is a scale close to the optimum scale, which means the SNR after denoising is just lower than the best one (see Figure 5-9). Note that variations of SNRs after PD denoising further demonstrate that decomposition scale constitutes a critical component in wavelet-based PD denoising. Denoising with higher decomposition scale may lead to the loss of some information of PD signals of interest. In contrast, denoising with lower decomposition scale may lead to the residual of noise in the denoised PD signals.

A full-scale wavelet expansion of noisy s_1 and s_2 have been implemented, and their approximation and detail coefficients at each decomposition scale are illustrated in Figure 5-10 and Figure 5-11, respectively. Based on the SCR distributions in Figure 5-5 and Figure 5-6, the minimum SCR for noisy s_1 and s_2 are obtained at scale 7 and scale 6, respectively. This can be directly observed from the detail coefficient sequences, as shown in Figure 5-10 and Figure 5-11. In Figure 5-10, D7 is smoother than any other detail coefficient sequence, which means that it has the minimum randomness. Equally, D6 in Figure 5-11 is less random than the others. Low randomness implies that white noise is not remained, or the residual of

white noise is minimum at that scale. The presence of white noise can be determined by the SCR value of that scale. As such, it is reasonable to adopt the SCR as a criterion for decomposition scale determination in wavelet-based denoising.

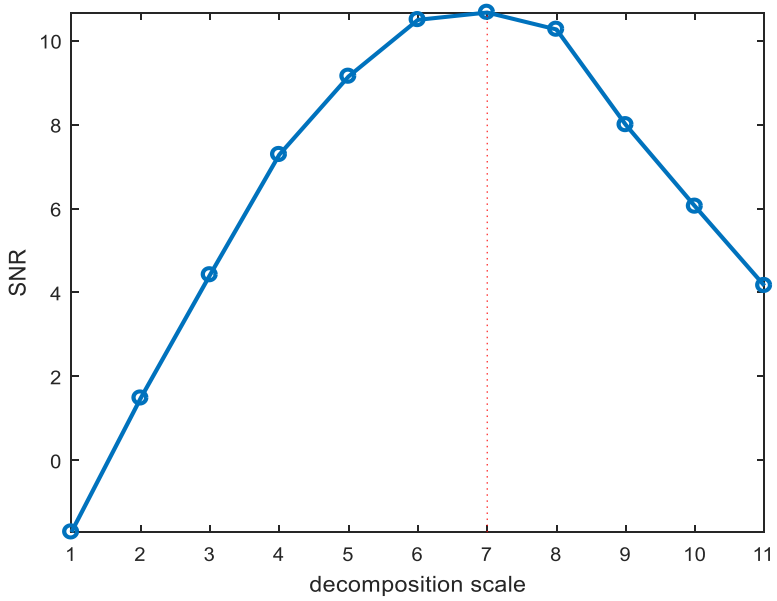

Figure 5-8 SNRs after denoising of noisy s_1 with various decomposition scales

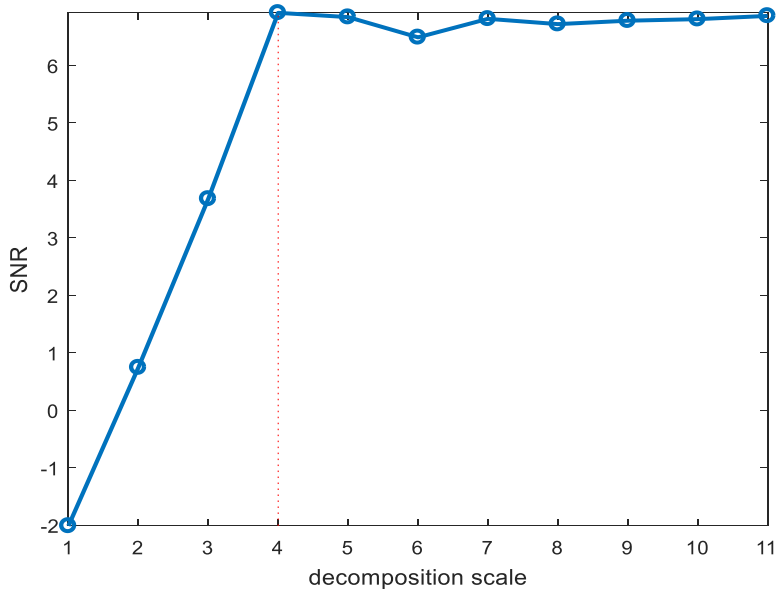

Figure 5-9 SNRs after denoising of noisy s_2 with various decomposition scales

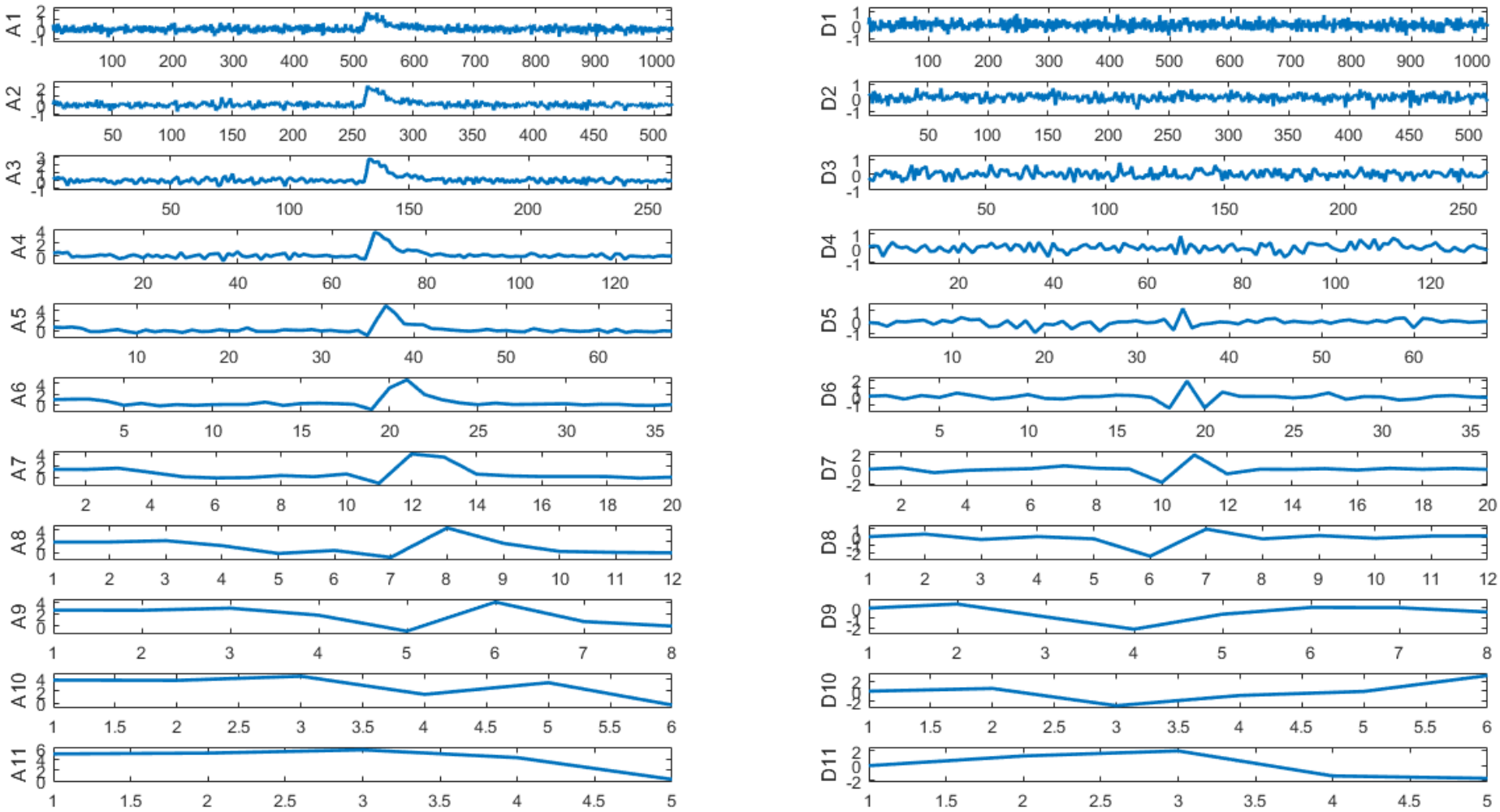

Figure 5-10 A full-scale wavelet expansion of noisy s_1: approximation and detail coefficients at each decomposition scale

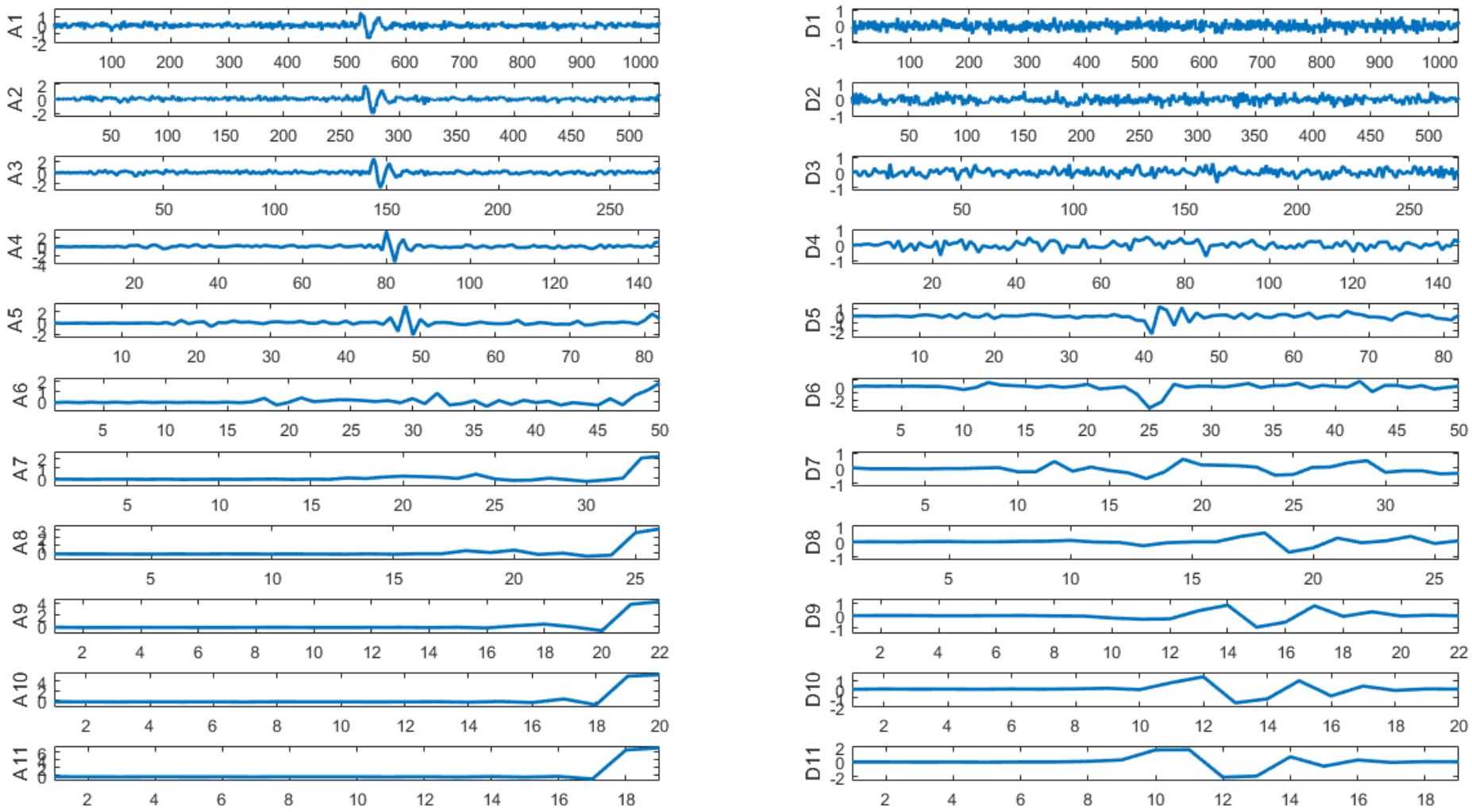

Figure 5-11 A full-scale wavelet expansion of noisy s_2: approximation and detail coefficients at each decomposition scale

The effectiveness of the sparsity-based method for decomposition scale determination has been analyzed with simulated PD signals contaminated by white noise with SNR = -5. In this scenario, it can select an appropriate decomposition scale for wavelet-based PD denoising and outperforms wavelet length method. In the attempt to fully evaluate its performance, PD signals s_1 and s_2 corrupted by different white noise levels are investigated. SNR is then set to -3, -1, 1, and 3 representing noise presence with various levels. Table 5-4 lists the decomposition scale selected by the sparsity-based method for PD signals with various SNRs.

Table 5-4 Decomposition scale selected by sparsity-based method for PD signals with various SNRs

		SNRs			
		-3	**-1**	**1**	**3**
ns_1	**DS**	7	7	7	7
	Best SNR Scale	7	7	7	6
ns_2	**DS**	5	5	5	5
	Best SNR Scale	4	4	4	4

ns_1 and ns_2 refer to as noisy PD signals, and Best SNR Scale is the scale at which the denoised PD signal can obtain the best SNR after denoising. Difference can be seen between the decomposition scale selected by the sparsity-based method and the best SNR scale for ns_1 when the noise level in noisy PD signals is not high. For ns_2, the selected scales are not in agreement with the best SNR scales. It is necessary to mention that the scales selected for ns_1 with SNR = 3 and for ns_2 at each noise level are the scales close to the best one, which means the resultant denoised PD signals using the selected scale can obtain the SNR just lower than the best one. As such, the decomposition scale determined by the sparsity-based method can provide a relatively good denoising results than wavelet length method, even if the scale selected is not the best one in certain scenario. Parameters, such as ME, MSE, and XCORR, used to compare the denoising results between the sparsity-based method and wavelet length method are illustrated in Figure 5-12 and Figure 5-13 for ns_1 and ns_2, respectively. Remarkable differences can be seen from these parameters of denoised PD signals using the sparsity-based method and wavelet length method, especially when SNR is

low. The promising performance of the sparsity-based method enables it to be a good approach for decomposition scale determination in wavelet-based PD denoising.

The sparsity-based method for decomposition scale determination is also applied to noisy PD signals with multiple pulses (see Figure 5-14), where spd_1 and spd_2, as shown in Figure 4-12 (a) and Figure 4-13 (a), are corrupted by white noise with SNR = -5. The SCR distribution of both noisy PD signals are illustrated in Figure 5-15 based on (5-7). Scale 7 and scale 4 are then selected based on the SCR criterion for the wavelet-expansion of noisy spd_1 and spd_2, respectively. Also, the scale at which the denoised PD signals can obtain the best SNR after denoising has been highlighted in Figure 5-16. It demonstrates that the selected scales by the sparsity-based method are in agreement with the best SNR scales. As a reference method, scale 11 and scale 9 are selected by wavelet length method for noisy spd_1 and spd_2, respectively. In fact, the SNRs after denoising shown in Figure 5-16 indicate that the selection of scale by the sparsity-based method is much better than wavelet length method. The denoising results of noisy PD signals with multiple pulses shown in Figure 5-17 further demonstrate that the sparsity-based method is a promising approach that can be applied for scale determination in wavelet-based PD denoising.

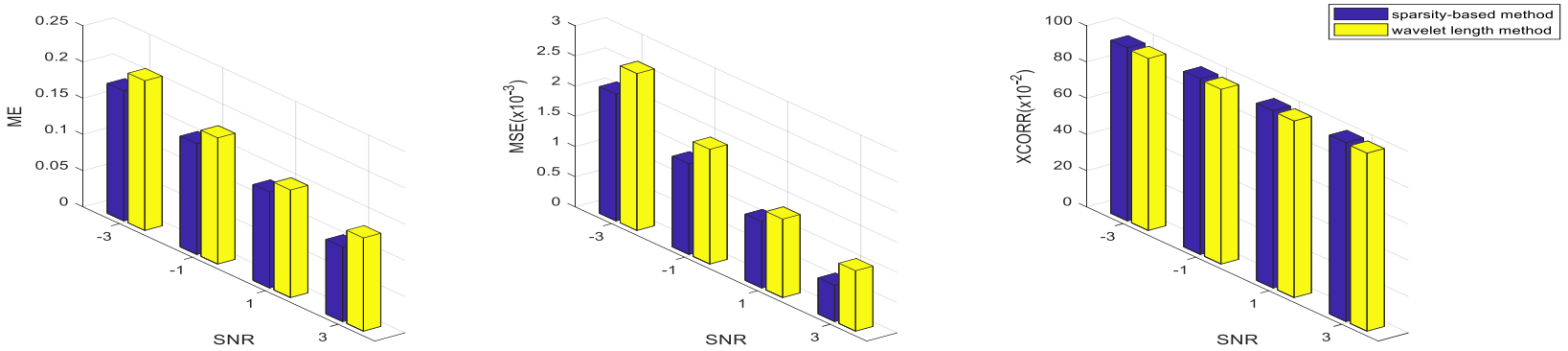

Figure 5-12 ME, MSE, and XCORR of denoised ns_1 using sparsity-based and wavelet length method

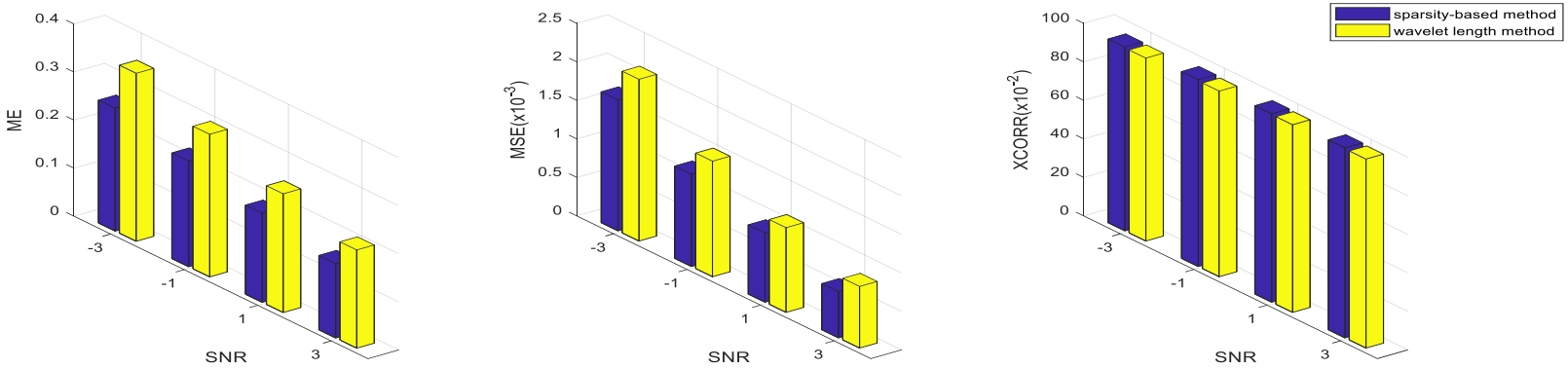

Figure 5-13 ME, MSE, and XCORR of denoised ns_2 using sparsity-based and wavelet length method

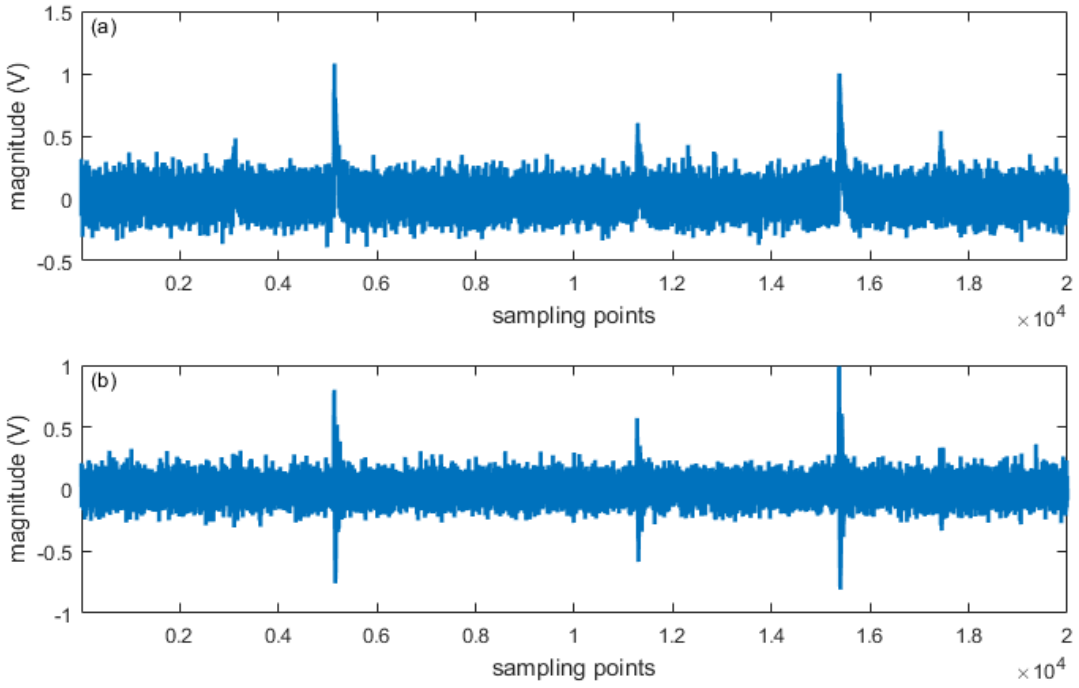

Figure 5-14 (a) spd_1 with SNR = -5, (b) spd_2 with SNR = -5

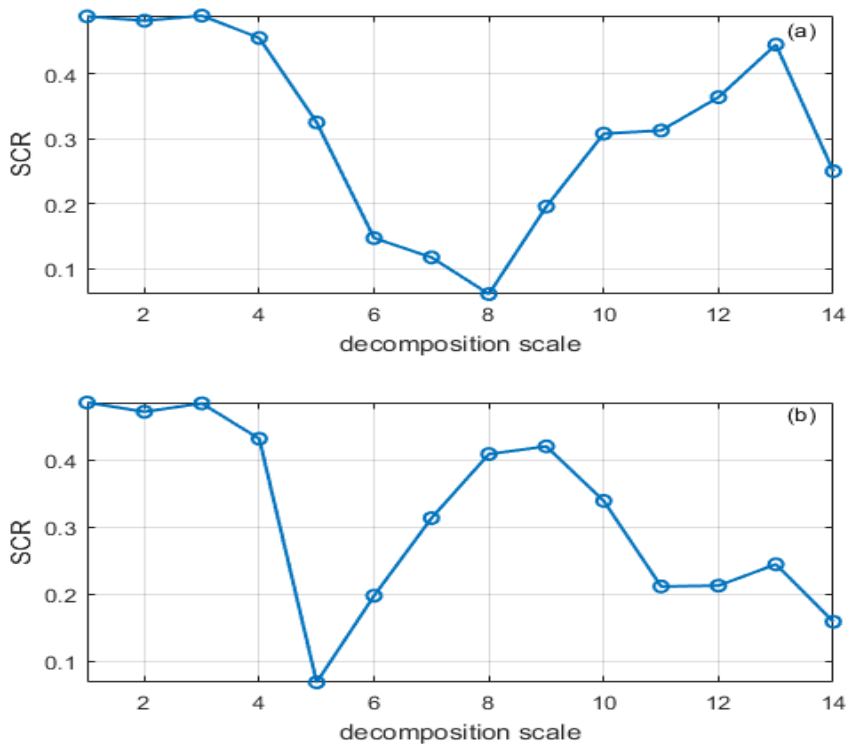

Figure 5-15 SCR distribution of transformed (a) noisy spd_1, (b) noisy spd_2

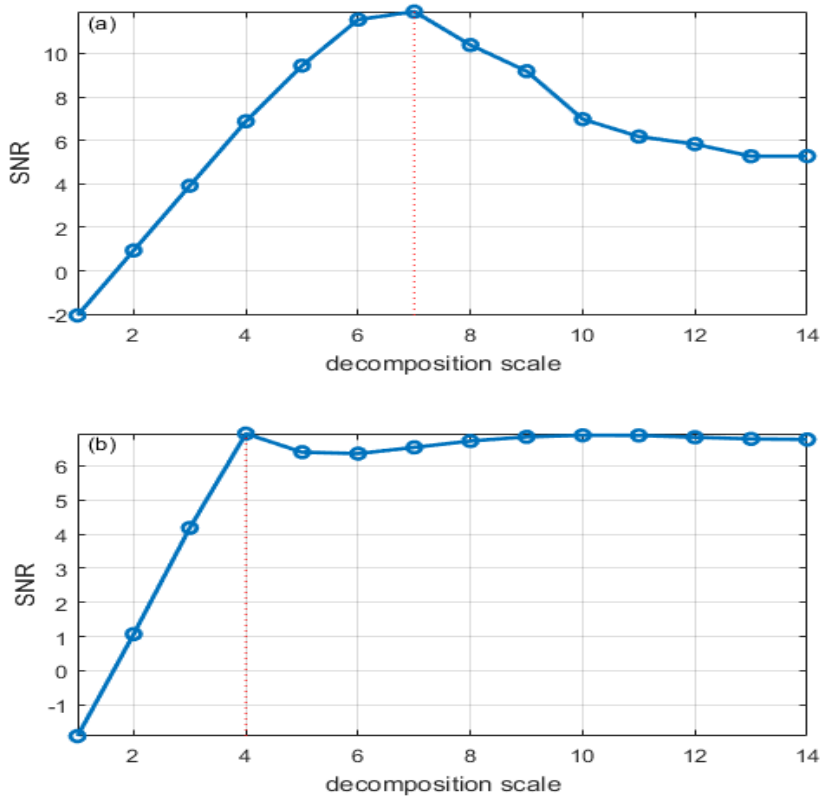

Figure 5-16 SNRs of denoised PD signals with various scales, (a) noisy spd_1, (b) noisy spd_2

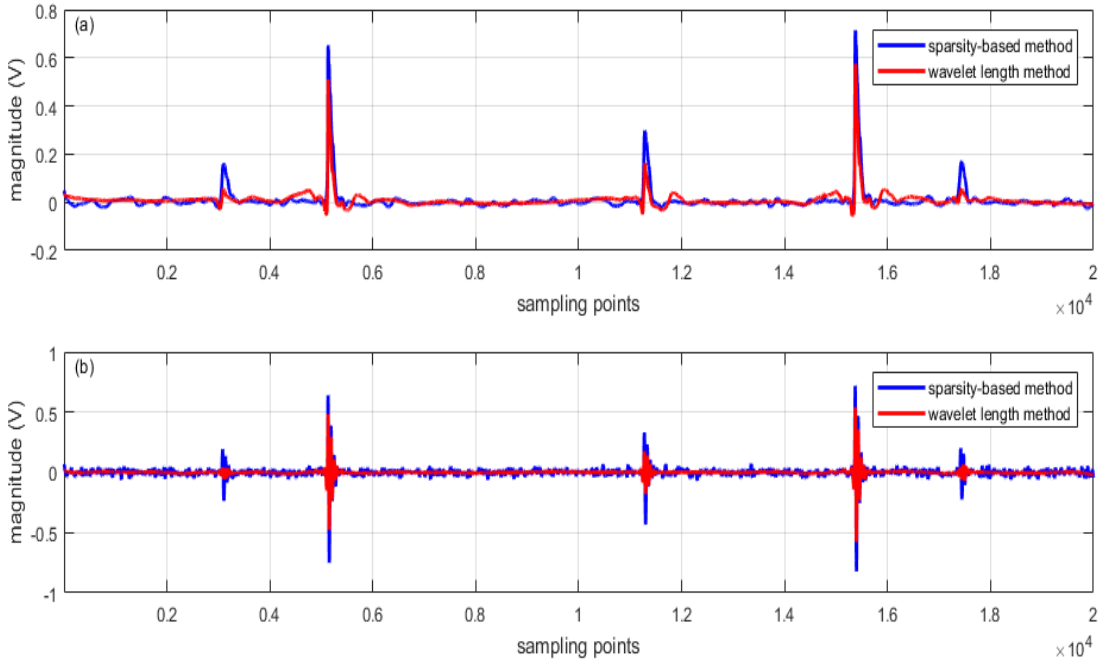

Figure 5-17 Denoising results of (a) noisy spd_1, (b) noisy spd_2 using different scale selection methods

In Chapter 4, a novel wavelet selection scheme, WEBWSS, was proposed based on the concept of wavelet entropy. It is necessary to investigate the performance of the sparsity-based method in combination with WEBWSS. As mentioned before, wavelet length method for the determination of decomposition scale is dependent on a fixed wavelet applied through the whole decomposition and reconstruction of PD signals. The fixed wavelet used in wavelet length method is selected by the correlation-based technique. For comparison, the combination of WEBWSS and the sparsity-based method for scale determination is termed WEBWSS-SP, and the integration of CBWSS and wavelet length method for scale determination is termed CBWSS-WL. PD signals s_1 and s_2 contaminated by white noise with SNR = -5 (shown in Figure 5-18 (a) and (b) respectively) are denoised by WEBWSS-SP and CBWSS-WL, respectively. Table 5-5 lists the decomposition scales selected by WEBWSS-SP and CBWSS-WL for noisy s_1 and s_2. The corresponding denoising results are illustrated in Figure 5-18 (c) – (f). Denoised PD signals shown in Figure 5-18 (c) and (d) using these two different methods demonstrate that denoised PD signals suffer less distortion due to the application of WEBWSS-SP. Residuals between original and denoised PD signals shown in Figure 5-18 (e) and (f) also indicate that denoised PD signals using the new method are closer to the original ones. The parameters used for performance evaluation are listed in Table 5-6, which indicates that WEBWSS-SP is much better than CBWSS-WL in PD signal denoising. In addition, the differences of the denoising results between Figure 5-7 and Figure 5-18 further demonstrate that WEBWSS is more advantageous than CBWSS in wavelet-based denoising of PD signals.

Table 5-5 Decomposition scale selected by WEBWSS-SP and CBWSS-WL for noisy s_1 and s_2

	Wavelet-based Denoising	
	WEBWSS-SP	**CBWSS-WL**
$\boldsymbol{s_1}$	6	8
$\boldsymbol{s_2}$	5	7

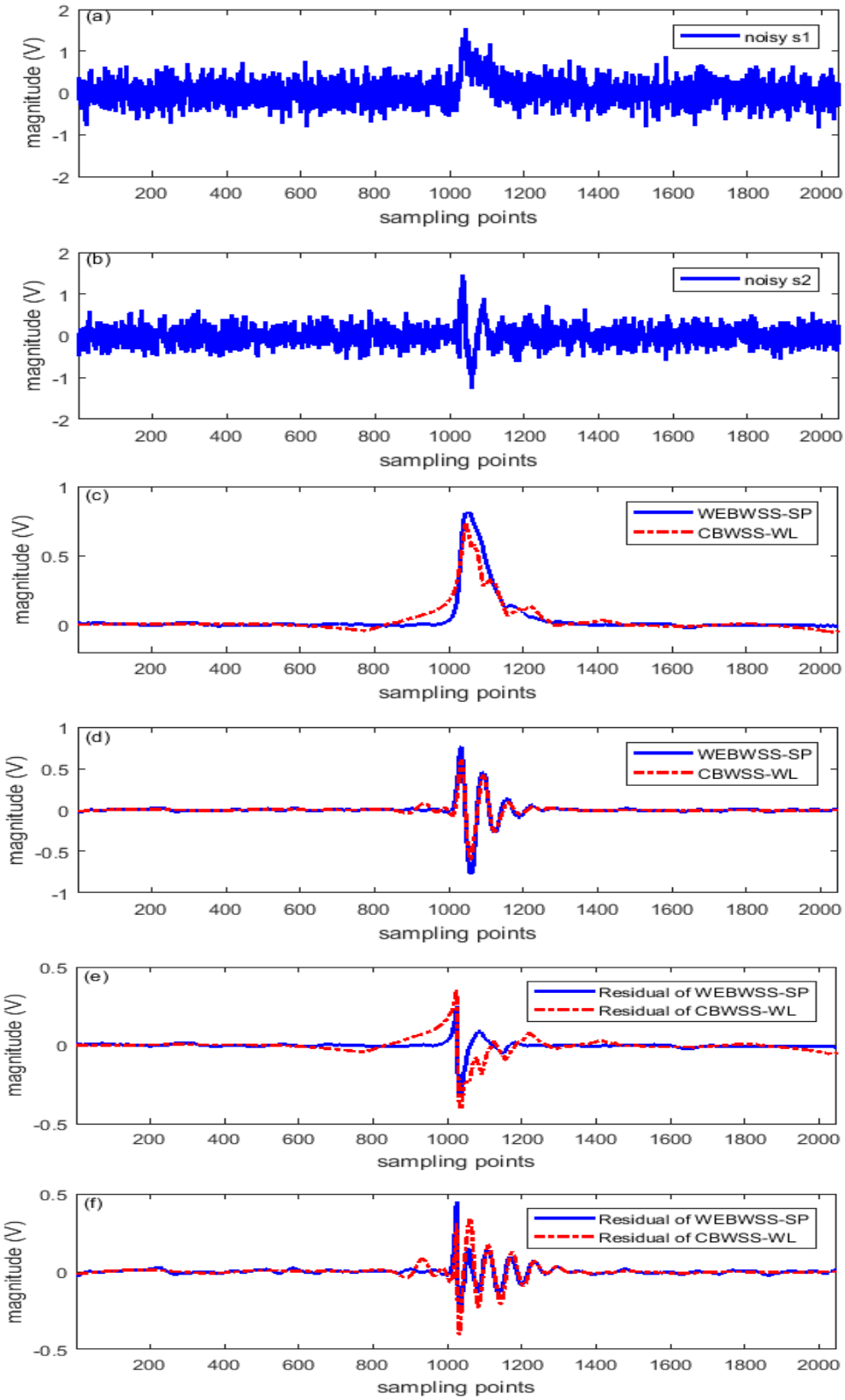

Figure 5-18 (a) noisy s_1, (b) noisy s_2, denoising results of (c) noisy s_1, (d) noisy s_2 using WEBWSS-SP and CBWSS-WL, (e) residuals between original and denoised s_1 using WEBWSS-SP and CBWSS-WL, (f) residuals between original and denoised s_2 using WEBWSS-SP and CBWSS-WL

Table 5-6 Parameters for performance evaluation of CBWSS-WL and WEBWSS-SP in PD denoising

		CBWSS-WL	WEBWSS-SP
s_1	**ME**	0.2650	**0.1821**
	MSE	0.0034	**0.0010**
	XCORR	0.9302	**0.9784**
s_2	**ME**	0.4012	**0.2436**
	MSE	0.0028	**0.0015**
	XCORR	0.9335	**0.9511**

A full comparison between WEBWSS-SP and CBWSS- WL in PD denoising is presented in Table 5-7 with PD signals s_1 and s_2 under various noise levels. Lower ME and MSE, and higher XCORR are desirable by the application of new techniques in wavelet-based PD denoising. The parameters listed in Table 5-7 demonstrate that noisy PD signals processed by WEBWSS-SP can have better denoising results than CBWSS-WL.

Table 5-7 Parameters used for comparisons between CBWSS-WL and WEBWSS-SP in PD denoising with various noise levels

		SNRs for CBWSS-WL				SNRs for WEBWSS-SP			
		-3	**-1**	**1**	**3**	**-3**	**-1**	**1**	**3**
s_1	**ME**	0.2067	0.1741	0.1485	0.1287	**0.1603**	**0.1243**	**0.0411**	**0.0127**
	MSE	0.0026	0.0019	0.0013	0.0010	**0.0008**	**0.0006**	**0.0003**	**0.0002**
	XCORR	0.9494	0.9646	0.9752	0.9826	**0.9822**	**0.9870**	**0.9928**	**0.9947**
s_2	**ME**	0.3495	0.2967	0.2470	0.2043	**0.2007**	**0.1034**	**0.1106**	**0.1901**
	MSE	0.0021	0.0015	0.0011	0.0008	**0.0012**	**0.0004**	**0.0003**	**0.0006**
	XCORR	0.9502	0.9641	0.9745	0.9816	**0.9627**	**0.9878**	**0.9899**	**0.9874**

Based on the analysis above, the sparsity-based method for decomposition scale determination is a promising technique that can be applied in the field of PD denoising. The SCR-based criterion for an appropriate scale selection is a data-driven and automated technique. When incorporated with wavelet-based PD denoising, it provides a reasonable scale that can effectively minimize the negative effects, e.g., signal distortion, on denoised PD signals. Accordingly, PD extraction and location can benefit from the improvement of wavelet-based PD denoising due to the involvement of the sparsity-based method for scale

determination.

5.4.2 PD Signals Corrupted by White Noise and DSI

The application of the sparsity-based method for decomposition scale determination has been fully discussed based on PD signals corrupted by white noise. As did for the test of WEBWSS for wavelet selection, the sparsity-based method is also extended to PD signals contaminated by both white noise and DSI.

The noisy PD signals with white noise and DSI used for the test are shown in Figure 4-14, where the DSI is added with magnitude of 0.1, and white noise has a mean value of zero and standard deviation of 0.3. The SNRs of noisy s_1 and s_2, based on (4-16), are -6.22 and -8.06, respectively. WEBWSS-SP is then applied to both noisy PD signals, and scale 7 (Figure 5-19) and scale 6 (Figure 5-20) are selected for noisy s_1 and s_2, respectively. The selected scale can reflect the randomness of the detail coefficient sequence at that scale, i.e., less random than the others. Denoising results with the selected scale are illustrated in Figure 5-21. The effect of DSI on denoised PD signals can be effectively reduced in this case. Based on this, low DSI level in noisy PD signals can be successfully suppressed by WEBWSS-SP. Further investigation is implemented through the increase of DSI level in original PD signals. As the magnitude of DSI increases to 0.3, scale 8 and scale 6 are the selected scale for noisy s_1 and s_2, respectively. The DSI starts to remain in both denoised PD signals, which can be seen from Figure 5-22. Also, severe distortion of the denoised PD signals can be observed in Figure 5-22 due to the involvement of DSI.

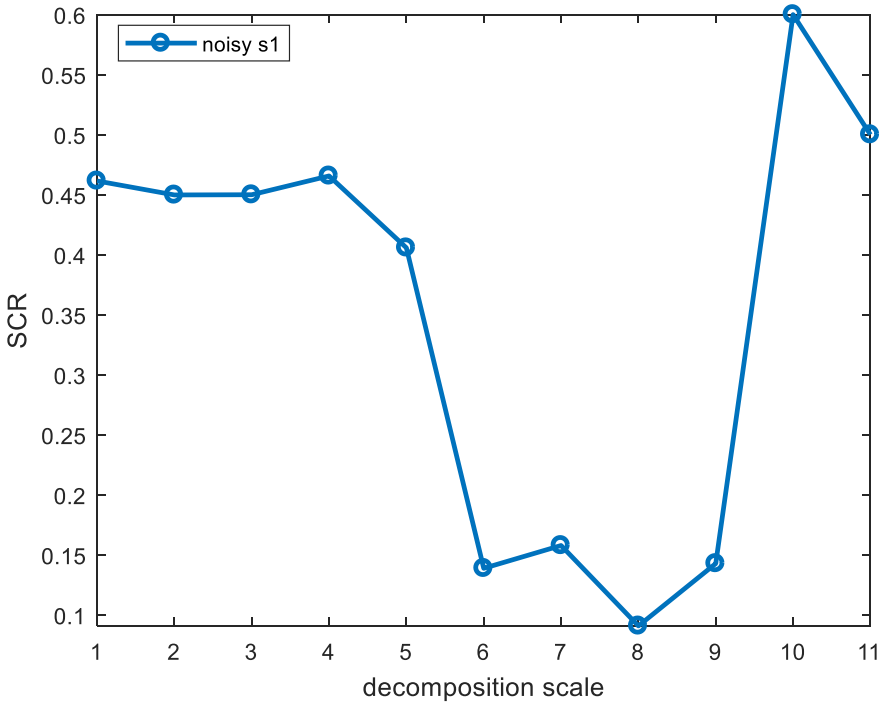

Figure 5-19 SCR distributions of noisy s_1 with white noise and DSI

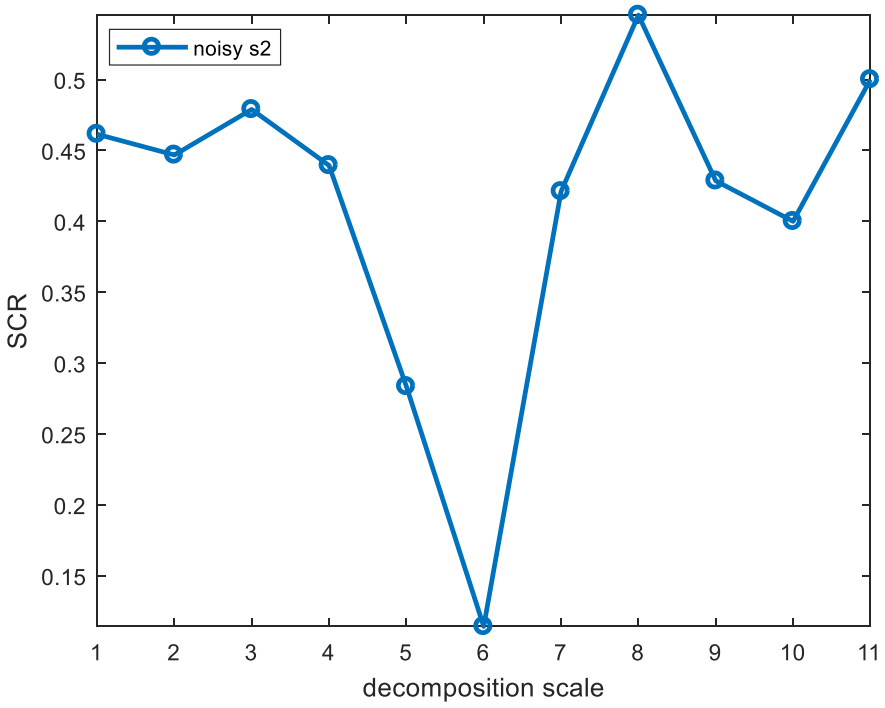

Figure 5-20 SCR distributions of noisy s_2 with white noise and DSI

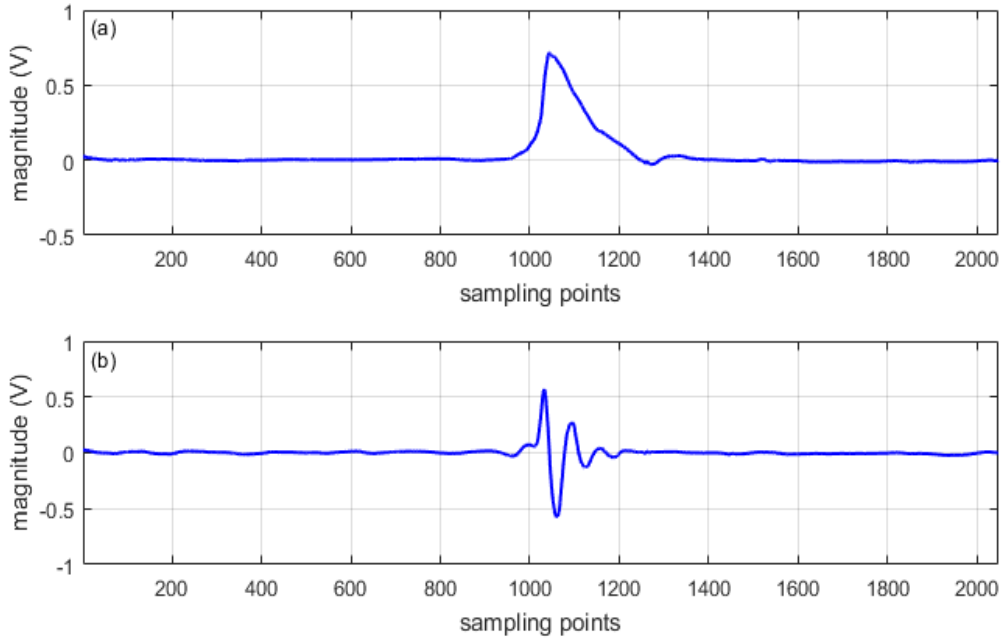

Figure 5-21 Denoising results of (a) noisy s_1, (b) noisy s_2 with white noise and DSI

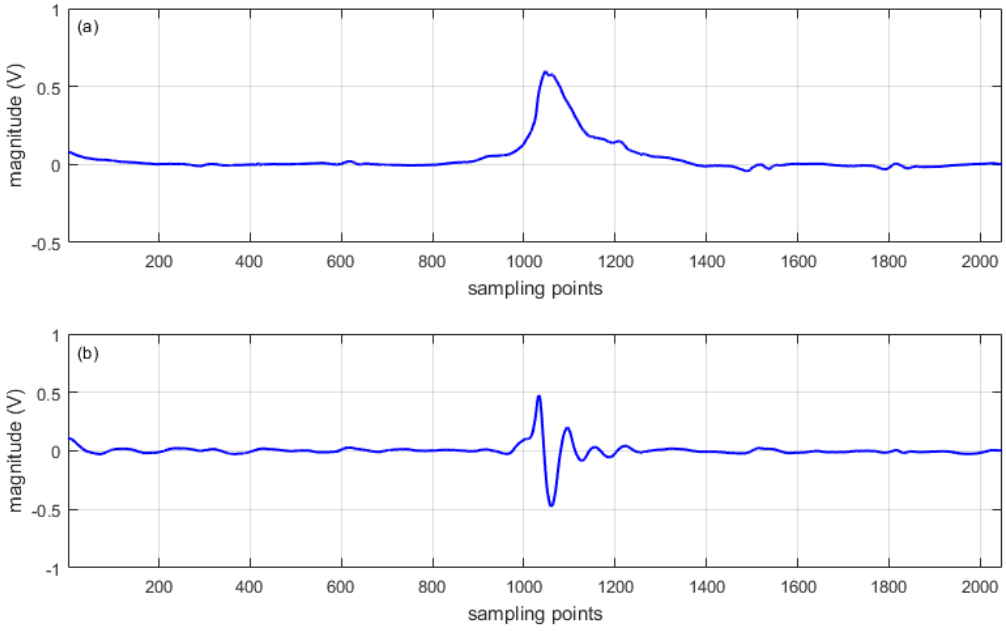

Figure 5-22 Denoising results of (a) noisy s_1, (b) noisy s_2 corrupted by white noise and relatively high-level DSI

Results presented above have seen the influence of DSI on the decomposition scale determination. The selected scale by the sparsity-based method, in combination with WEBWSS, cannot fully suppress DSI with relatively high level in noisy PD signals. Also, the

pulse peak of denoised PD signals are seriously distorted with the increased DSI level in noisy PD signals. The lower effectiveness of the reduction of DSI may be mainly due to the wavelet-based PD denoising technique, as the similar situation also cannot be circumvented in the wavelet selection for PD denoising.

5.5 Conclusion

In the area of PD denoising, decomposition scale of wavelet-based technique is either determined by trial and error, or experience. As a result, the motivation of this research is to propose a new method that can automatically select a reasonable decomposition scale. In this chapter, a SCR-based criterion has been proposed to automatically select an appropriate decomposition scale in the context of PD denoising. The SCR of the detail coefficient sequence at each decomposition scale can reflect the sparsity of that sequence. The lower SCR, the higher the sparsity is. Based on the definition of SCR, high sparsity indicates that low level noise presented in that scale. In the meantime, wavelet length method and sparseness measurement method have been discussed. Sparseness measurement method functions as a part of inspiration for the novel method with significant improvement in scale selection. In relation to PD denoising, wavelet length method is adopted as a reference method to highlight the advantages of the sparsity-based method. Investigations show that the sparsity-based method can select a reasonable decomposition scale for wavelet-based PD denoising, and present better denoising results than wavelet length method.

The aforementioned explication of the sparsity-based method is based on the assumption that PD signals are only contaminated by white noise. To test its robustness in noise reduction, it has been extended to noisy PD signals with white noise and DSI. Results indicate that the integration of WEBWSS and the sparsity-based method for scale selection can suppress both noise if DSI is in a low level. With the increase of DSI level in noisy PD signals, the effectiveness of DSI reduction deteriorates.

It is necessary to mention that the improved effectiveness of this sparsity-based method in wavelet-based noise suppression was derived from simulated PD signals. Further investigation requires to extend it to real PD signals, which can have various waveforms due

to those frequently mentioned factors such as signal attenuation and design of detecting circuits. For this purpose, real PD signals will be applied in Chapter 8 to demonstrate its feasibility in PD denoising in terms of varied waveforms. It may have the same limit as that in the wavelet selection scheme proposed in last chapter, which is caused by the limited range of PD waveforms involved in the investigation.

6 Threshold Estimation for Wavelet-based PD Denoising

6.1 Introduction

Signal denoising can benefit from wavelet expansion as both time and frequency information can be unveiled simultaneously. Wavelet selection and decomposition scale determination have been discussed in last two chapters. Threshold estimation is another critical element in wavelet-based technique for signal denoising. Note that both the signal of interest and noise are decomposed and possibly remained together at each decomposition scale through a wavelet expansion. To remove the noise from the signal of interest, an appropriate threshold needs to be applied to the detail coefficient sequence at each decomposition scale using a thresholding function or policy. Generally, threshold estimation technique is proposed based on the basic idea that the energy of a signal often concentrates in a few coefficients while the energy of noise spreads among all coefficients in wavelet domain. Due to the difference of energy concentration, a few large coefficients representing the signal are remained, while a large number of coefficients with small magnitudes representing noise is removed through the estimated threshold as well as a thresholding function or policy. As such, thresholds of wavelet coefficients determine the quality of the denoised signal. If the estimated threshold is too low, annoying noise will be remained in the denoised signal, while if the estimated threshold is too high, the signal of interest will suffer severe distortion. Based on this, numerous techniques for threshold estimation, including Minimax threshold [140], the Rigorous SURE threshold [140], [174], and the universal threshold [140], have been developed during the past decades.

6.2 Existing Threshold Estimation Techniques

Decomposition, thresholding, and reconstruction are the basic procedures for wavelet-based denoising. The selection of wavelet and decomposition scale, which is closely related to decomposition, has been fully discussed in previous chapters. Reconstruction is not a critical

issue if challenges of decomposition and thresholding have been well addressed. The thresholding normally includes two aspects. One aspect is to define an operator $\Gamma\{\cdot\}$ for the determination of an appropriate threshold thr, the other is to define an operator $T\{\cdot\}$ that reduces noise to obtain the thresholded wavelet coefficients $d'_{j,i}$. These two aspects can be expressed by the following equations,

$$thr = \Gamma\{d_{j,i}\} \tag{6-1}$$

$$d'_{j,i} = T\{d_{j,i}, thr\} \tag{6-2}$$

In this section, the most commonly used threshold estimations that function as the operator $\Gamma\{\cdot\}$ in MATLAB are discussed as follows.

6.2.1 Minimax Threshold

Minimax threshold is a threshold estimation technique proposed to minimize the constant term in an upper bound of the risk involved in estimating a function [140], [174], [175]. Minmax adopts a fixed threshold that can yield minimum/maximum variance. The minimax principle is used in statistics to design estimators. Since the denoised signal can be assimilated to the estimator of the unknown regression function, the minimax estimator is the option that realizes the minimum, over a given set of functions, of the maximum mean square error [163]. Details of the mathematical derivations of Minimax threshold is not in the scope of this thesis and can be seen in [140].

6.2.2 Rigorous SURE Threshold

Rigorous SURE threshold in MATLAB is named ‘rigrsure’. It is a threshold estimation technique that uses the Stein’s Unbiased Risk Estimate(SURE) criterion to get an unbiased estimate [140], [175], [176]. Different from Minimax threshold, ‘rigrsure’ threshold is estimated based on the wavelet coefficients at each decomposition scale, i.e., the estimated threshold is scale-dependent. The estimate of the risk for a particular threshold value λ can be obtained through SURE, and then minimizing the risks in λ gives a selection of the threshold value [163]. More details regarding ‘rigrsure’ threshold can be seen in [140], [174].

6.2.3 Universal Threshold

Universal threshold is proposed by Donoho & Johnstone as an alternative to the use of minimax thresholds in [140]. It suggests thresholding of wavelet coefficients by a fixed threshold given by

$$thr = \hat{\sigma}\sqrt{2logn} \tag{6-3}$$

where n is the number of data points or the length of the signal and $\hat{\sigma}$ is the estimate of the noise level, which is the median absolute deviation of the wavelet coefficients at the finest level, divided by 0.6745. This threshold is easy to remember and its implementation in software requires no costly development of look-up tables as compared to minimax threshold [140], [174]. The universal threshold is substantially larger than the minimax threshold for any particular value of n (see the difference in Figure 6-1). As such, fewer coefficients will be involved in the reconstruction resulting in an approximation that is smoother than the minimax threshold.

For a given signal or data sequence, the parameter n in (6-3) is constant and scale-independent. However, the wavelet patterns of PD signals and noise are scale dependent [22]. As a result, the threshold using (6-3) can reduce too much energy of the PD signal in wavelet-based PD denoising. To mitigate this negative effect, a revised universal threshold (see (3-8)) has been proposed in [22].

6.2.4 Existing Threshold Estimation Techniques in PD Denoising

It is necessary to mention that the universal threshold in MATLAB is equal to $\sqrt{2logn}$, and named 'sqtwolog'. In relation to PD signal denoising, a revised version of the universal threshold has been proposed in [22]. It is a scale-dependent noise estimation technique, and thus, more accurate denoising results can be obtained. Due to the enhanced quality of denoised PD signals, it has been widely used in PD denoising [12], [21], [123], [170], [177]–[179]. Figure 6-1 illustrates the threshold estimation techniques for white noise with 2048 sampling points. The white noise is normal distributed in Figure 6-1. It can be seen that a small portion of coefficients are retained for rigrsure and minimax thresholds, while all values

are rejected in the case of the universal threshold. The difference between sqtwolog and revised universal threshold indicate that the coefficients of noise can be over-smoothed by both thresholds, but the over-smoothing caused by sqtwolog is more serious than the revised universal threshold.

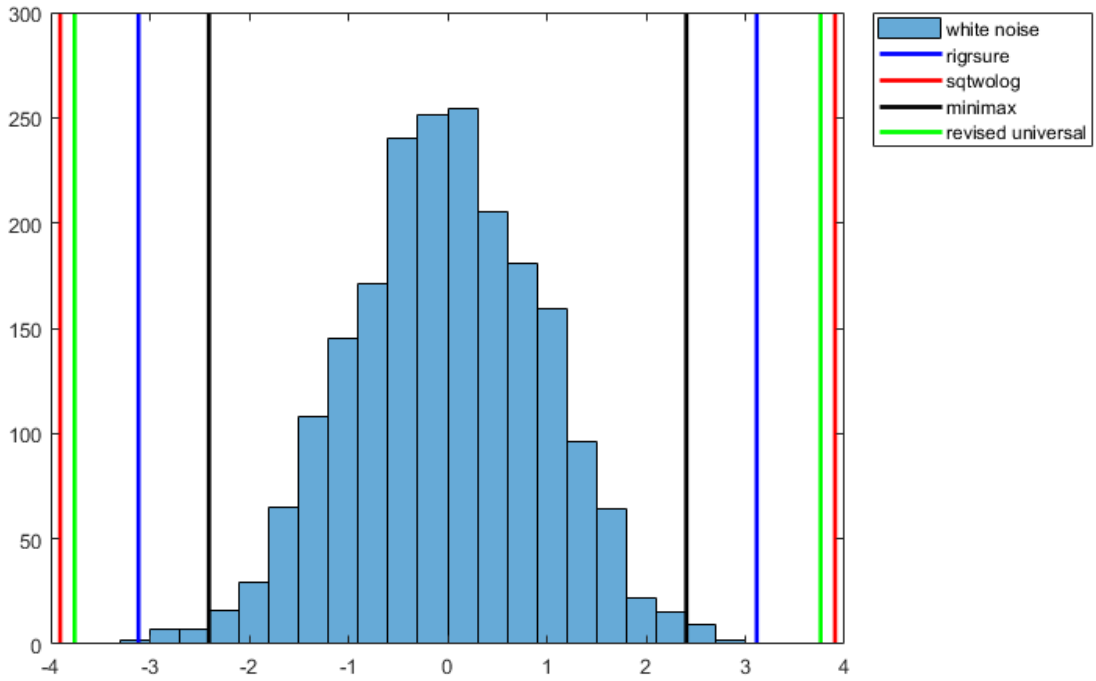

Figure 6-1 Threshold for white noise selected by different threshold estimation techniques

Figure 6-2 and Figure 6-3 delineate the denoising results of PD signals, s_1 and s_2 contaminated by white noise with SNR = 0. For both denoising results, it can be seen that the universal threshold (refers to as revised universal threshold) can present better results than the others. Also, the use of various threshold estimations further demonstrates that the distortion of denoised signals is dependent on the threshold used.

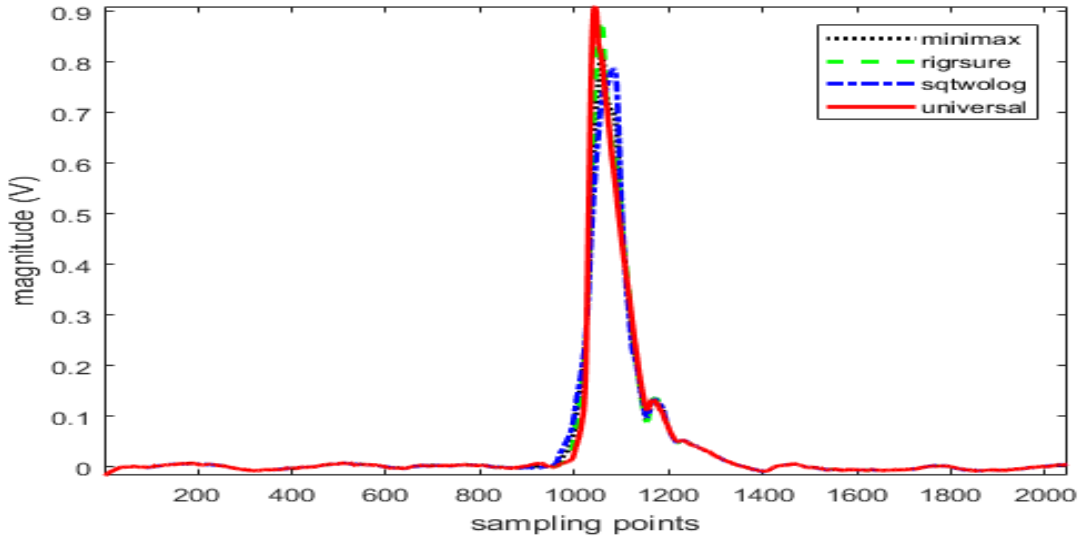

Figure 6-2 Denoising results of s_1 through different threshold estimation techniques

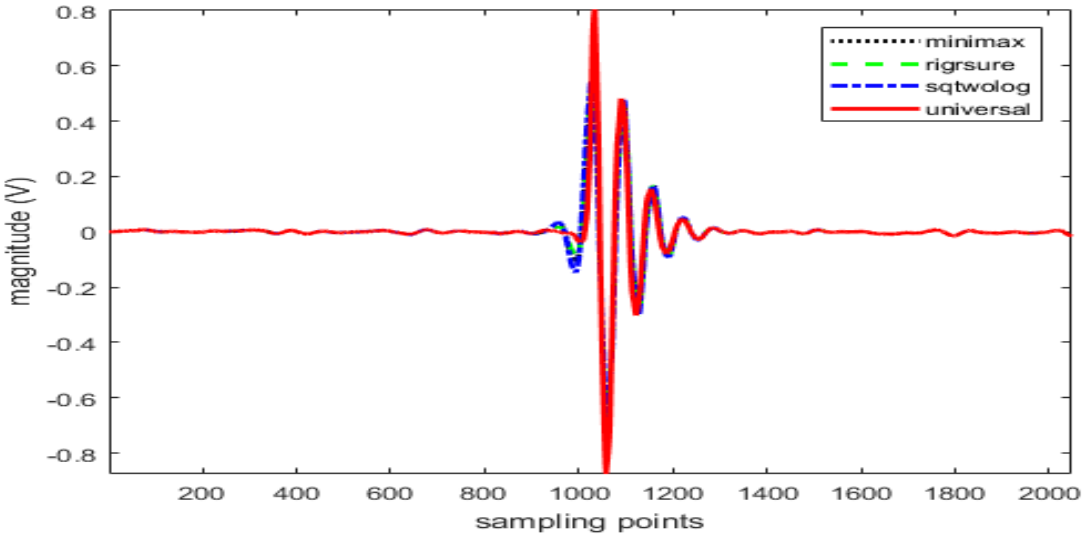

Figure 6-3 Denoising results of s_2 through different threshold estimation techniques

The parameter, MSE, is used to evaluate the degree of distortion of denoised s_1 and s_2 with various threshold estimations, and the corresponding figures are listed in Table 6-1. Parameters in the table show that the distortion of denoised PD signals using the universal threshold is much less than the others. It means that the universal threshold is preferable than the minimax, rigrsure and sqtwolog thresholds in the area of PD denoising. Based on this, the universal threshold is adopted as a reference threshold estimation technique for the novel threshold estimation introduced in next section.

Table 6-1 MSE used for the evaluation of denoising results through various threshold estimations

	Threshold Estimations			
	Minimax	**Rigrsure**	**Sqtwolog**	**Universal**
s_1	0.0022	0.0015	0.0033	**0.0005**
s_2	0.0034	0.0029	0.0044	**0.0008**

6.3 A Novel Threshold Estimation Technique

Among the aforementioned threshold estimation techniques, the universal threshold takes a leading position in signal denoising due to its optimality properties [140], [180]. The universal threshold, however, is estimated based on the noise level, typically a scaled median absolute deviation of the wavelet coefficients [181]. Although the application of the universal threshold 'guarantees' a noise-free reconstruction, it usually underfits the data [181]. Also, it was claimed that the application of the universal threshold can lead to over-smoothed approximations [180], [182]. Either the underfitting or the over-smoothing of the data indicates that unexpected distortion is an inevitable consequence of the application of the universal threshold.

Distortion of a denoised signal is highly correlated to both base wavelet selection and threshold estimation when the wavelet-based technique is used for signal denoising [183]. Due to the particular strengths in the processing of non-stationary signals, e.g., PD signals, the searching for an optimal wavelet and threshold of wavelet-based denoising is still an expanding area of ongoing research. The goal is to recover an unknown function or signal with less distortion from sampled data that are contaminated by noise. In Chapter 4, a wavelet entropy-based scheme was proposed to minimize the distortion of denoised PD signals. However, threshold estimation for wavelet-based PD denoising is still a challenge. It is widely accepted that the choice of threshold is crucial to the success of wavelet-based techniques for signal denoising. As such, a new formula is proposed in this chapter for the selection of an appropriate threshold for the application of wavelet-based denoising in PD signal extraction.

The novel threshold estimation technique is proposed based on the SCR distribution of

transformed PD signals, and thus, named as SCR-based noise estimation. It is expected that the application of this new technique can minimize the distortion of denoised PD signals due to the threshold used. The definition of SCR has been introduced in last chapter for the determination of an appropriate decomposition scale. The SCRs of detail coefficient sequences of a noisy PD signal through wavelet expansion is hereby used to distinguish the level of noise presented in these sequences. By virtue of numerous simulations, the detail coefficient sequence can be divided into four categories (C_1 – C_4) based on the value of SCR:

1. C_1: Detail coefficient sequence only consists of noise coefficients if the SCR is higher than 0.4;
2. C_2: Detail coefficient sequence is mainly dominated by noise coefficients with very few high-magnitude signal coefficients if the SCR is in the range of 0.25 to 0.4;
3. C_3: Detail coefficient sequence is primarily dominated by signal coefficients with small-magnitude noise coefficients if the SCR is in the range of 0.1 to 0.25;
4. C_4: Detail coefficient sequence only contains signal coefficients, i.e., noise is no longer discernable if the SCR is less than 0.1.

The wavelet expansion of noisy s_1 (used for Figure 6-2) is used as an example to explicate the categories above. Figure 6-4 illustrates the detail coefficient sequences of the transformed noisy s_1. The figure on the top right corner of each detail coefficient sequence is the SCR at that scale. The SCR of D1, D2, and D3 is over 0.4, and thus, these detail coefficient sequences are categorized into noise coefficients only. The high degree of randomness of these three coefficient sequences can be seen from a direct observation. For D4, its SCR is 0.28, and thus, is categorized into the coefficient sequence which is primarily dominated by noise coefficients and very few signal coefficients with high magnitudes. D5, D6, and D7 are categorized into the detail coefficient sequences that mainly contain the signal coefficients as the SCRs are in the range of 0.1 to 0.25. The last detail coefficient sequence, D8, is signal coefficients only due to its SCR is less than 0.1. It is worth noting that the detail coefficient sequence in the aforementioned categories occurs in sequential order, as shown in Figure 6-4.

Based on the categories of detail coefficient sequences, a new formula is proposed for the noise estimation, and is given by

$$\lambda = \mathbb{\Gamma}(k) \cdot \sqrt{2logn_j} \tag{6-4}$$

$$k = floor(c_m \cdot N_0), m = 1,2,3,4 \tag{6-5}$$

where $\mathbb{\Gamma}(k)$ is a function to select the k^{th} element of the wavelet coefficients in a descending order of absolute magnitudes, n_j is the length of the detail coefficient sequence, c_m is a constant determined by the category of detail coefficient sequences, and N_0 is the number of significant coefficients of the detail coefficient sequence, which can be calculated based on (5-3).

The parameter c_m is assigned to different values with respect to the category of detail coefficient sequences. The coefficients in C_1 are noise coefficients only. The value of threshold λ is expected to high enough such that all the linked coefficients can be discarded for reconstruction. Based on this idea, λ is equal to the maximum value of the absolute magnitudes of the wavelet coefficients. Then, c_1 for this threshold can be calculated by the equation below,

$$c_1 \approx \frac{max(|d_{j,i}|)}{N_0 \cdot \sqrt{2logn_j}} \tag{6-6}$$

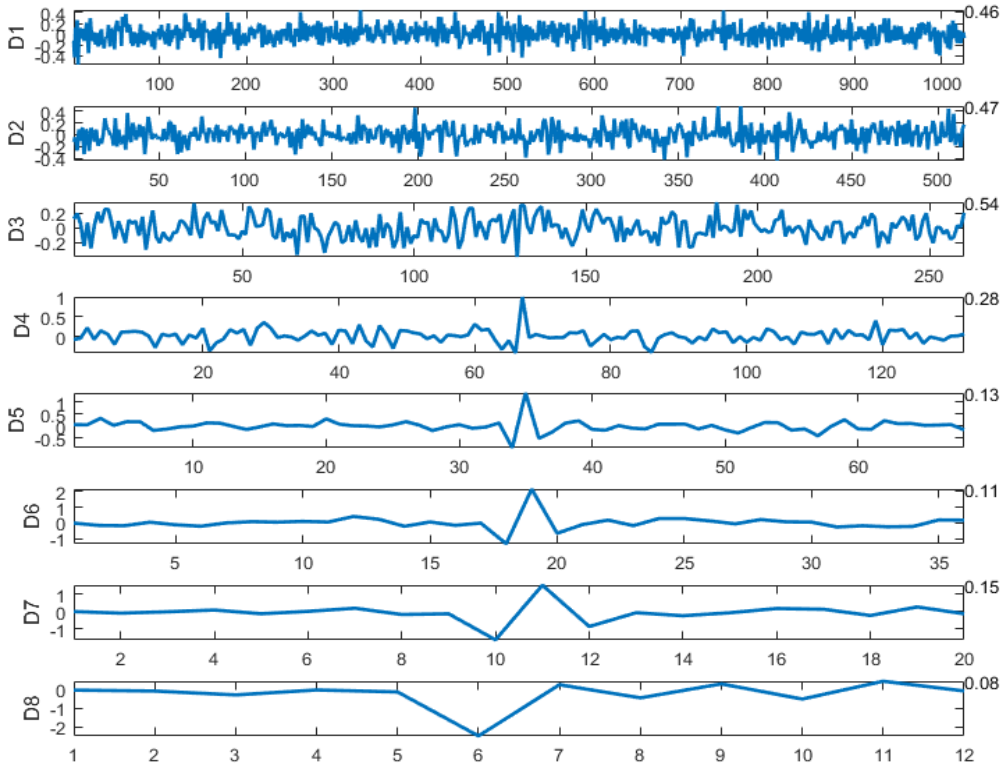

Figure 6-4 The SCRs of detail coefficient sequences of transformed noisy s_1

With numerous simulations of PD signals contaminated by various levels of white noise, c_j equal to 0.6 in (6-6) can present effective suppression of noise in denoised PD signals. Similar idea is applied to the detail coefficient sequences in C_2, C_3, and C_4. The values of c_2 , c_3 and c_4 can be estimated based on c_1 and the SCR range of that category. The equation for the estimation is given by

$$c_m \approx \frac{c_1}{SCR_m} \cdot \frac{max(|d_{j,i}|)}{N_0 \cdot \sqrt{2logn_j}}, m = 2,3,4 \quad (6\text{-}7)$$

Based on (6-7) and results of numerous simulations, the value of c_2 equal to 1 for C_2, c_3 equal to 2.5 for C_3, and c_4 equal to 10 for C_4 can lead to effective reduction of white noise.

The definition of this new threshold estimation denotes that the threshold at each decomposition scale is dependent on the significant coefficient ratio rather than the estimation of noise level. As such, it is expected to remove more noise from noisy PD signals and cause less distortion to the denoised PD signals. This will be demonstrated through simulated PD signals later in this chapter.

6.4 A New Thresholding Function for Wavelet-based Denoising

The reduction of noise in wavelet coefficients at each decomposition scale is implemented through a thresholding function or policy. As shown in (3-9) and (3-10), the hard- and soft-thresholding functions are frequently used in wavelet-based denoising. Both of them have their particular strengths and weakness. Generally, soft-thresholding function is better than the hard-thresholding function as the thresholded wavelet coefficients $d'_{j,i}$ in (3-9) are discontinuous at the points of both $-thr_j$ and thr_j [123], [174], [184], [185]. Due to its discontinuity, hard thresholding function is not stable, i.e., hard thresholding function is sensitive to small changes in the data. This is not the case in soft-thresholding function. Continuous transition can be obtained as the coefficients with magnitudes larger than the threshold shrink by an amount equal to the threshold [180]. However, the derivative of soft-thresholding function is discontinuous, which is undesirable as the continuous derivative or higher order derivatives are often desired for optimization problems [183], [186]. Also, the

difference $\left|d_{j,i} - d'_{j,i}\right|$ is constant (see Figure 3-14 (c)), which can affect the precision of wavelet reconstruction results [184]. To overcome the weakness of both hard- and soft-thresholding functions, a new thresholding function proposed in [184] without introducing any new parameter is applied in this chapter and defined as

$$d'_{j,i} = \begin{cases} sgn(d_{j,i}) \cdot (\left|d_{j,i}\right| - thr_j / exp(\frac{\left|d_{j,i}\right| - thr_j}{thr_j})), if\ \left|d_{j,i}\right| > thr_j \\ 0, \qquad\qquad\qquad\qquad\qquad if\ \left|d_{j,i}\right| \leq thr_j \end{cases} \tag{6-8}$$

The new thresholding function in (6-8)does not include any new parameter, and has continuous high order derivatives [184]. In addition, this smooth thresholding function can reduce or even remove the constant difference between $d_{j,i}$ and $d'_{j,i}$, and thus, improve the precision of signal reconstruction [184]. To differentiate it from other thresholding functions, it is termed as H-S thresholding function in this thesis.

The H-S thresholding function will be applied in wavelet-based PD denoising, and the denoising results will be compared to those using traditional thresholding functions. Details regarding this investigation is presented in next section.

6.5 Performance Evaluation of the SCR-based Threshold Estimation

A novel threshold estimation has been proposed in this chapter. Simulated PD signals with various noise levels are used to investigate its effectiveness in PD denoising. As discussed above, the universal threshold (refers to the revised universal threshold from this section) can present the best denoising results among the existing threshold estimations in the field of PD detection. As a result, the universal threshold is adopted as a reference threshold estimation in this investigation. Initially, the novel threshold estimation technique is applied to PD signals corrupted by white noise. It is then extended to PD signals contaminated by both white noise and DSI, as did for wavelet selection scheme and decomposition scale determination. In the meantime, the H-S thresholding function is used to show its advances as compared to the conventional hard- and soft-thresholding functions.

6.5.1 PD Signals Corrupted by White Noise

6.5.1.1 SCR-based Threshold Estimation

PD signals s_1 and s_2 contaminated by white noise with SNR = -5, as shown in Figure 5-4, are used to test the novel threshold estimation. The wavelet entropy-based selection scheme and sparsity-based decomposition scale determination are applied to noisy PD signals but with different thresholds, i.e., the universal and the novel threshold. Figure 6-5 and Figure 6-6 illustrate the detail coefficient sequences of transformed noisy s_1 and s_2, respectively. Both the universal threshold and the novel threshold have been highlighted at each decomposition scale, the red line denotes the novel threshold while the black dash line is the universal threshold.

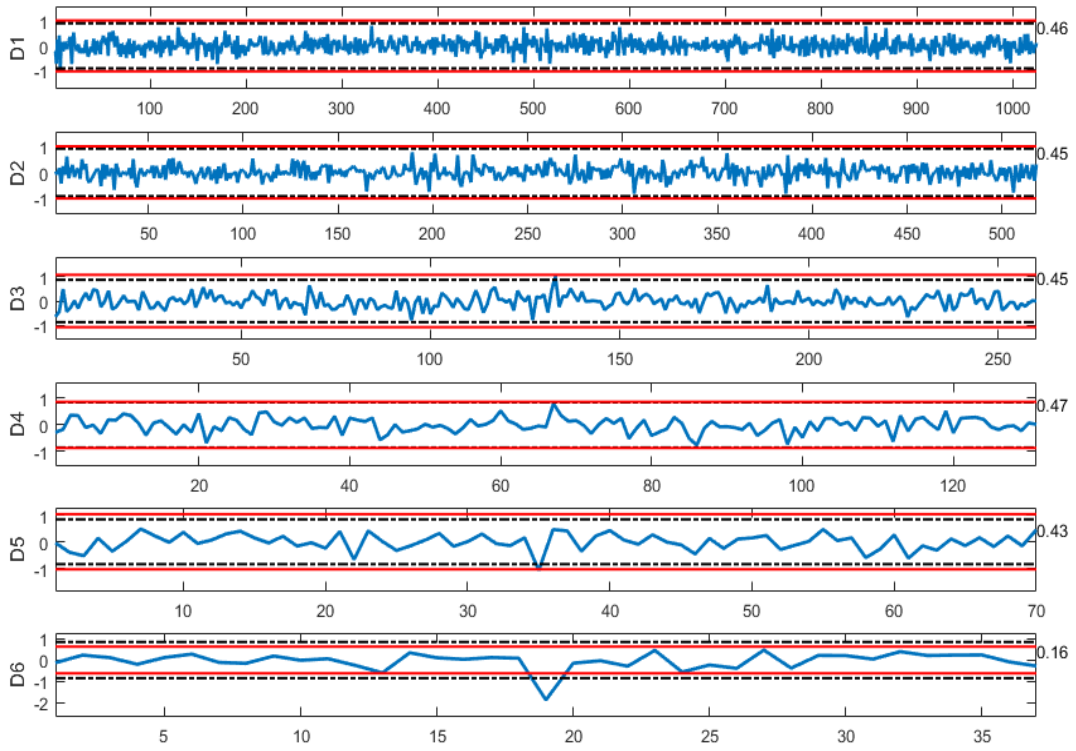

Figure 6-5 Thresholds selected by different threshold estimations for transformed noisy s_1

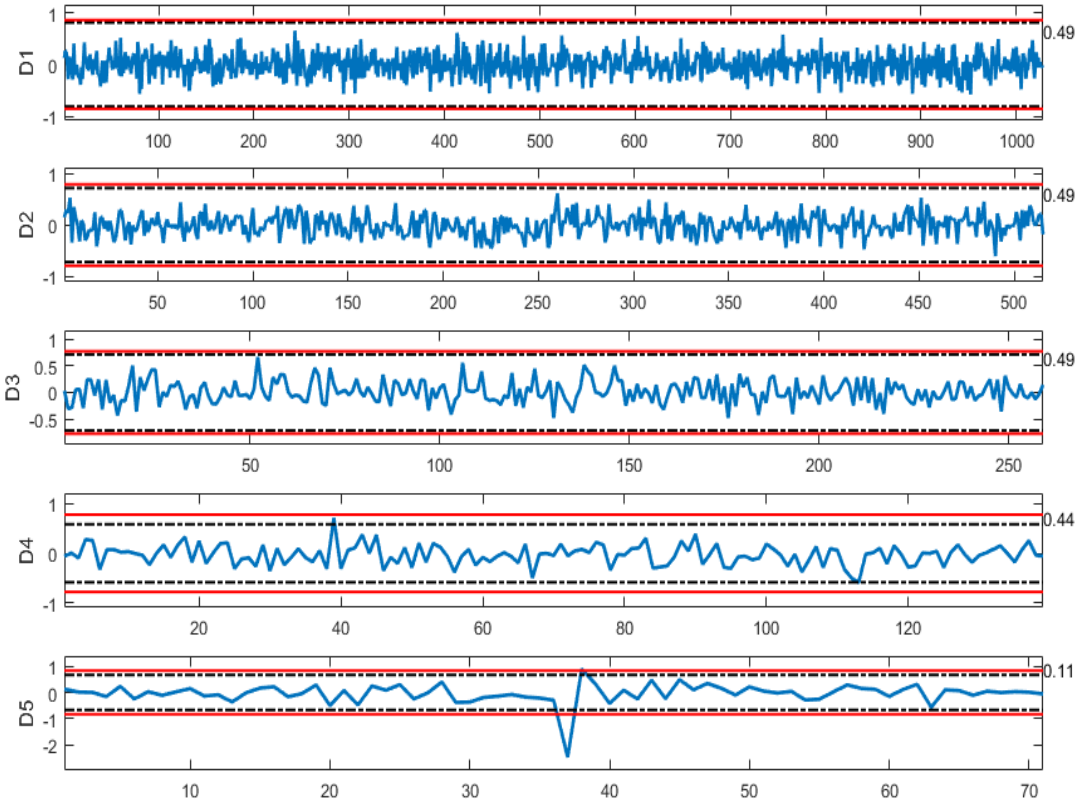

Figure 6-6 Thresholds selected by different threshold estimations for transformed noisy s_2

The figure on the top right corner of each detail coefficient sequence is the SCR of that scale. Difference can be seen from the universal threshold and the novel threshold for the detail coefficient sequences. The detail coefficients at the scale that has the SCR over 0.4 are discarded based on the novel threshold. However, some of them have been retained based on the universal threshold (see D3 and D5 in Figure 6-5).

The denoising results of noisy s_1 and s_2 are delineated in Figure 6-7 (a) and (b), respectively. Admittedly, significant advancements of the novel threshold cannot be directly seen from Figure 6-7. Parameters, such as ME, MSE, and XCORR, are listed in Table 6-2, which demonstrate that the application of the novel threshold estimation can present better denoising results or cause less distortion than the universal threshold estimation in the field of PD extraction. The improvement of ME indicates that PD signals with small magnitudes may be extracted through the application of this novel threshold estimation. Lower MSE and higher XCORR mean that the denoised PD signals suffer less signal distortion as compared to the universal threshold estimation.

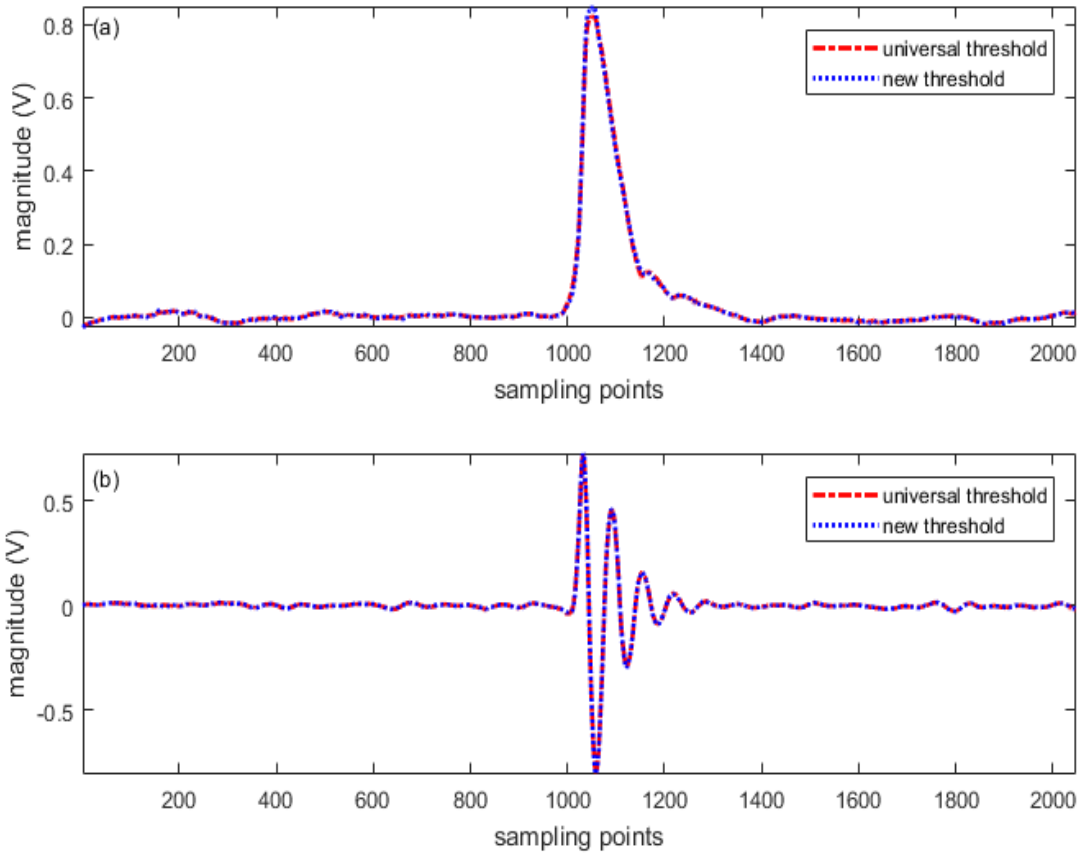

Figure 6-7 Denoised results of (a) noisy s_1, and (b) noisy s_2 using different thresholding functions

Table 6-2 Parameters used for performance evaluation of different threshold estimations

	ME		MSE		XCORR	
	Universal	**Novel**	**Universal**	**Novel**	**Universal**	**Novel**
s_1	0.1745	**0.1497**	0.0010	**0.0008**	0.9781	**0.9824**
s_2	0.2815	**0.2737**	0.0014	**0.0013**	0.9562	**0.9598**

In the attempt to fully test the novel threshold estimation in PD denoising s_1 and s_2 are simulated with white noise contamination under various noise levels. The noise levels of noisy PD signals are presented by setting SNRs equal to -3, -1, 1, and 3. The results of performance test of the novel threshold estimation are illustrated in Figure 6-8 and Figure 6-9. It can be seen that better performances can be achieved through the use of the novel threshold estimation. Based on this, the novel threshold estimation can be a promising approach for the improvement of wavelet-based PD denoising.

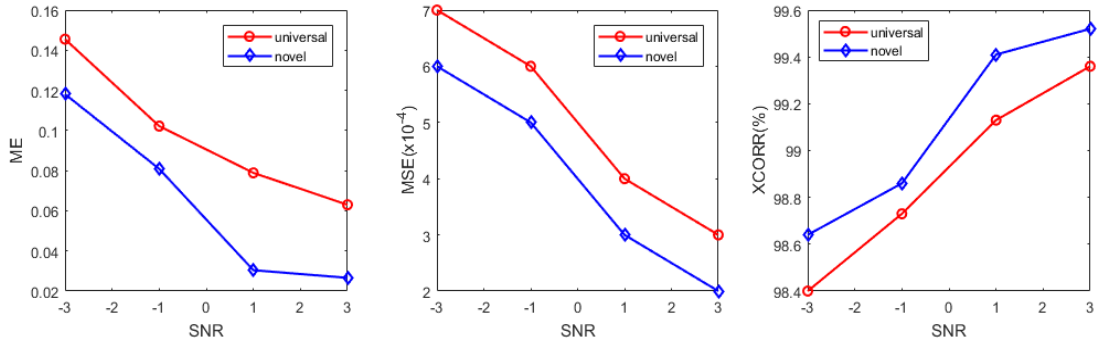

Figure 6-8 ME, MSE, and XCORR between s_1 and denoised s_1 using different threshold estimations

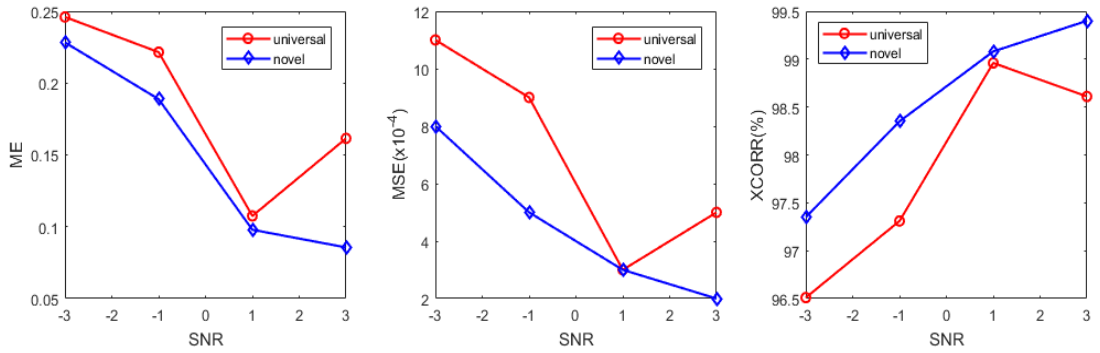

Figure 6-9 ME, MSE, and XCORR between s_2 and denoised s_2 using different threshold estimations

Performance test is also applied to PD signals with multiple pulses, i.e., spd_1 and spd_2, as shown in Figure 4-12 (a) and Figure 4-13 (a), respectively. Both spd_1 and spd_2 are corrupted by white noise with SNR = 0. Denoising results of noisy spd_1 and spd_2 using different threshold estimations are delineated in Figure 6-10. The corresponding parameters used to evaluate the performance of the novel threshold estimation are listed in Table 6-3. Although the improvement for the denoising of PD signals with multiple pulses are not that remarkable, less distortion of the denoised PD signals can be obtained by the application of the novel threshold estimation (see Table 6-3). Also, the improved SNR of denoised PD signals means more noise can be reduced by the use of the novel threshold estimation. This investigation further supports that improvement of wavelet-based PD denoising can be achieved through the application of the novel threshold estimation.

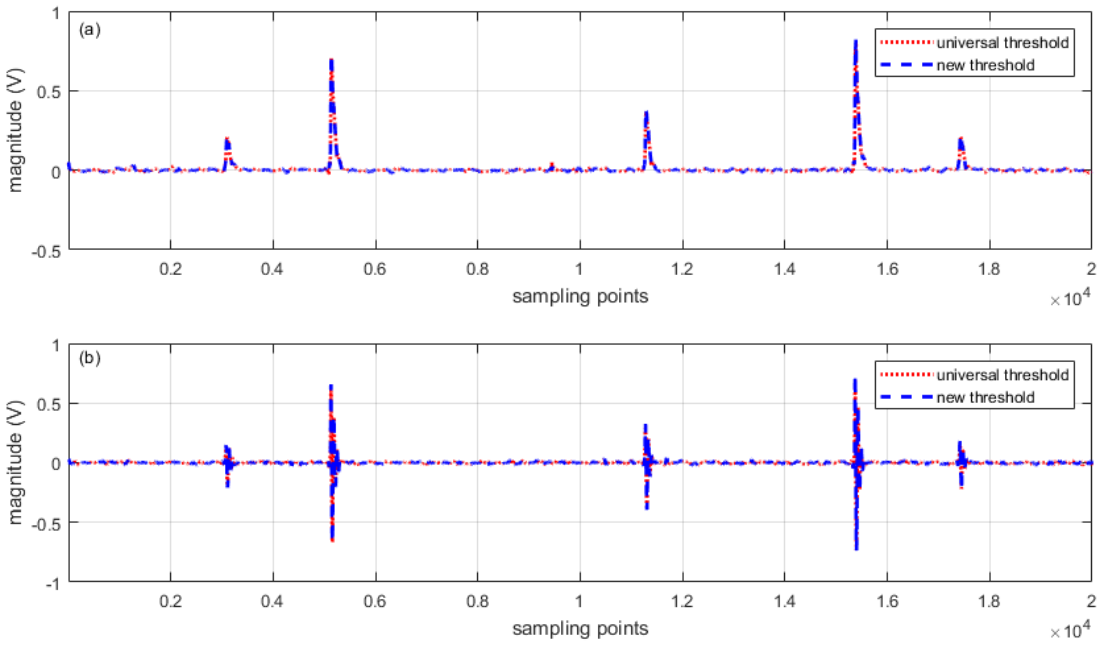

Figure 6-10 Denoising results of (a) noisy spd_1, (b) noisy spd_2 using different threshold estimations

Table 6-3 Parameters used for performance evaluation of different threshold estimations

	MSE		XCORR		SNR	
	Universal	**Novel**	**Universal**	**Novel**	**Universal**	**Novel**
$\boldsymbol{spd_1}$	1.18×10^{-4}	1.06×10^{-4}	0.9827	**0.9845**	14.65	**15.11**
$\boldsymbol{spd_2}$	1.76×10^{-4}	1.34×10^{-4}	0.9608	**0.9700**	11.09	**12.26**

6.5.1.2 H-S Thresholding Function

The advances of the novel threshold estimation have been demonstrated through the denoising results of simulated PD signals. Based on this, the effectiveness of the H-S thresholding function in wavelet-based PD denoising will be investigated in combination of the new threshold estimation. The noisy s_1 and s_2, as shown in Figure 5-4, are used for this investigation. The denoising results of both noisy PD signals are illustrated in Figure 6-11. Direct observation of both denoised PD signals indicates that wavelet-based denoising using the H-S thresholding function can lead to less peak distortion of the denoised PD signals. Further demonstration of the H-S thresholding function is implemented through the parameters listed in Table 6-4. Both ME and MSE of the denoised PD signals using this new

thresholding function can be improved over 35% as compared to the soft thresholding function. This remarkable improvement is desirable as it implies that PD signals can be effectively extracted with less distortion, especially those with small magnitudes.

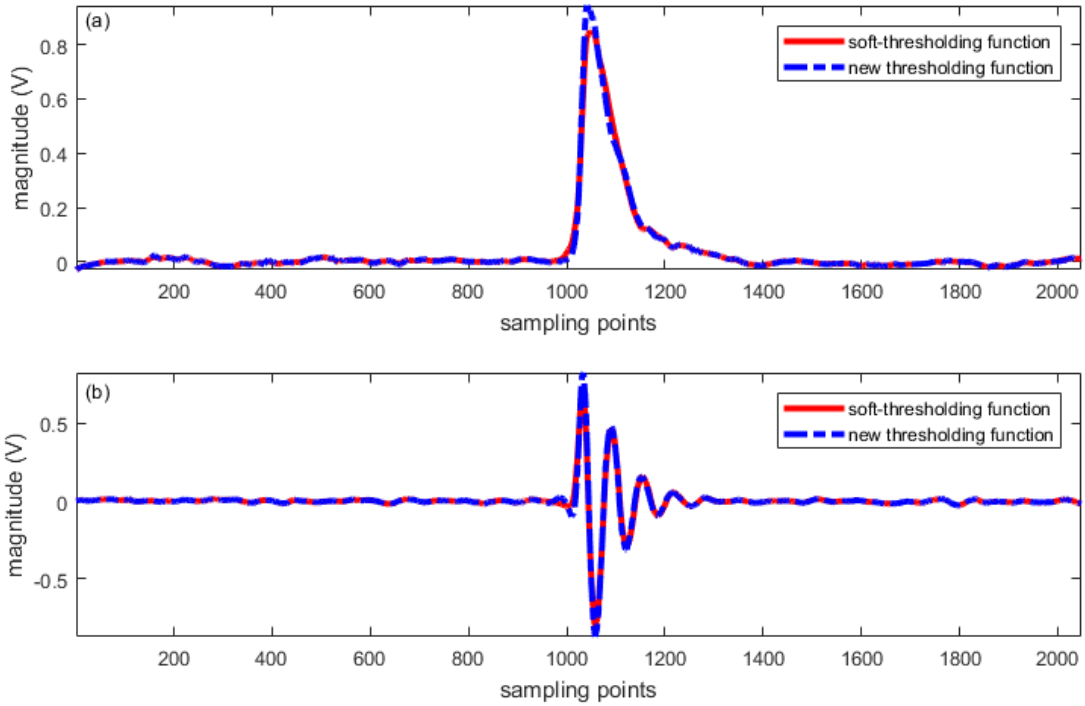

Figure 6-11 Denoising results of (a) noisy s_1, (b) noisy s_2 using different thresholding functions

Table 6-4 Parameters used for performance evaluation of different thresholding functions

	ME		**MSE**		**XCORR**	
	Soft	**New**	**Soft**	**New**	**Soft**	**New**
s_1	0.1497	**0.0576**	0.0008	**0.0004**	0.9824	**0.9908**
s_2	0.2737	**0.1751**	0.0013	**0.0008**	0.9598	**0.9760**

The H-S thresholding function is also applied to PD signals with multiple pulses. The noisy PD signals are the same as those used for Figure 6-10. The denoising results of these PD signals with multiple pulses are illustrated in Figure 6-12, and the associated parameters used for performance test of the H-S thresholding function are listed in Table 6-5. The lower MSE and higher XCORR indicate that the denoised PD signals suffer less distortion by the use of the H-S thresholding function. The higher SNR after denoising denotes that more noise can be

reduced through this new thresholding function.

The denoising results from both PD signals with single and multiple pulses demonstrate that the H-S thresholding function can improve the performance of wavelet-based techniques for PD denoising.

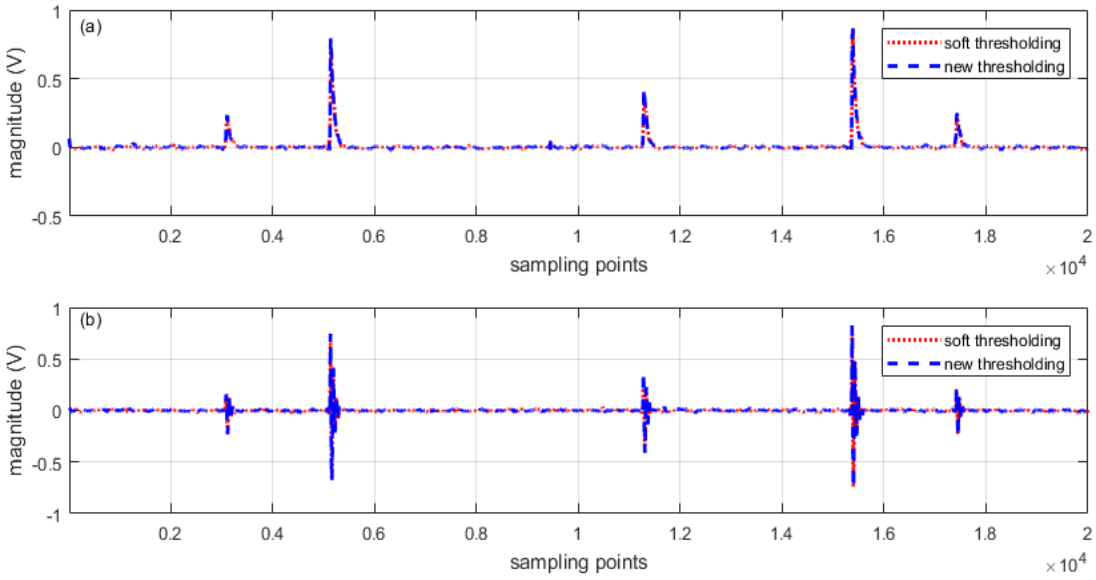

Figure 6-12 Denoising results of (a) noisy spd_1, (b) noisy spd_2 using different thresholding functions

Table 6-5 Parameters used for performance evaluation of different thresholding functions

	MSE		XCORR		SNR	
	Soft	**H-S**	**Soft**	**H-S**	**Soft**	**H-S**
$\boldsymbol{spd_1}$	1.06×10^{-4}	8.34×10^{-5}	0.9845	**0.9880**	15.11	**16.18**
$\boldsymbol{spd_2}$	1.34×10^{-4}	1.02×10^{-4}	0.9700	**0.9779**	12.26	**13.55**

6.5.2 PD Signals Corrupted by White Noise and DSI

Investigations of the novel threshold estimation as well as the new thresholding function are also implemented on PD signals corrupted by both white noise and DSI. The test procedures are the same as those for WEBWSS and sparsity-based scale determination. The noisy PD signals with white noise and DSI used for the test are shown in Figure 4-14, where the DSI is

added with magnitude of 0.1, and white noise has a mean value of zero and standard deviation of 0.3. The SNRs of noisy s_1 and s_2 , based on (4-16), are -6.22 and -8.06, respectively. The denoising results of PD signals corrupted by white noise and DSI are delineated in Figure 6-13.

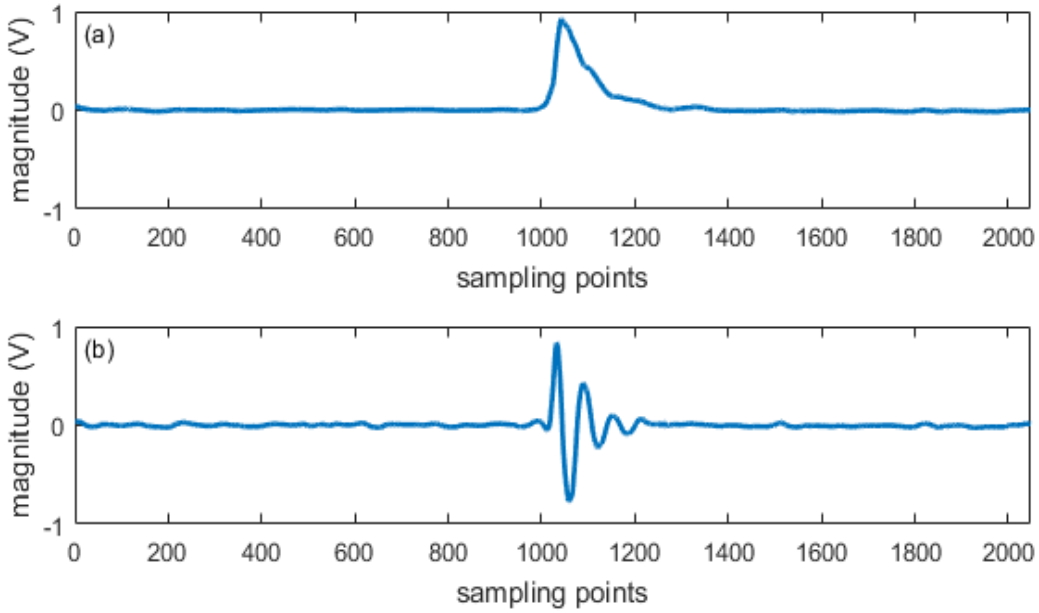

Figure 6-13 Denoising results of (a) noisy s_1, (b) noisy s_2 with white noise and DSI

In this case, the effect of white noise and DSI on PD signals can be minimized through the application of the SCR-based threshold estimation and the H-S thresholding functions in wavelet-based PD denoising. As compared to the denoising results shown in Figure 4-15 and Figure 5-21, it is clear that the peak distortion has been remarkably reduced. Equally, the effectiveness of noise suppression will be impaired if the novel level of DSI increases in the noisy PD signals. The denoising results when the DSI is added with magnitude of 0.3 are illustrated in Figure 6-14. The residual of DSI start to have a minor effect on the denoised PD signals. This effect of DSI becomes serious with the increase of DSI level. It is the same issue that has occurred in WEBWSS and WEBWSS-SP for the denoising of PD signals contaminated by both white noise and DSI. The results indicate that wavelet-based PD denoising with the SCR-based threshold method and the H-S thresholding function can minimize the effect of the DSI, but this is limited to low-level DSI in noisy PD signals. The

residual of DSI can have a major effect on denoised PD signals with the increase of t DSI level in noisy PD signals.

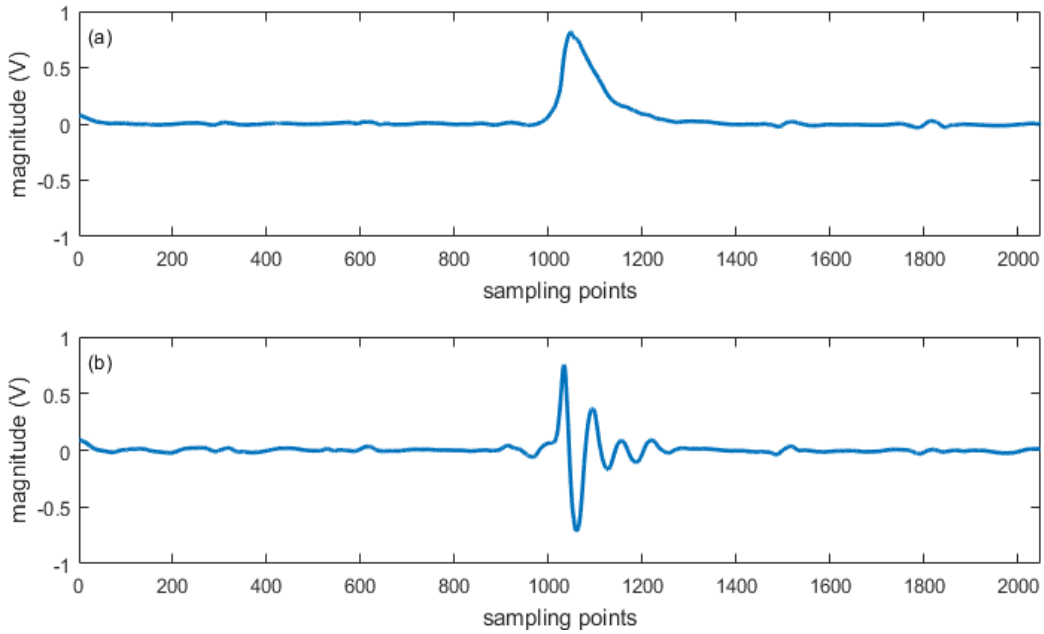

Figure 6-14 Denoising results of (a) noisy s_1, (b) noisy s_2 corrupted by white noise and relatively high-level DSI

6.6 Conclusions

In this chapter, universal threshold estimation was proved to be more effective than other threshold estimations integrated in MABLAB for PD denoising. However, universal threshold is derived from the estimation of noise level, more exactly, white noise level. To estimate the threshold through a signal itself, a new method was proposed based on the findings presented in last chapter. That is, the sparsity of a transformed PD signal at each decomposition was applied for the noise estimation. Denoising results show that the effectiveness of wavelet-based denoising is slightly improved by introducing this new method. Also, less distortion can be achieved for denoised PD signals. It demonstrates that this new method can provide an alternative to current existing threshold estimations.

In the meantime, conventional thresholding functions cannot effectively split the noise from PD signals due to their inherent limits. To overcome these limits, a new H-S thresholding

function used was referenced from others. Results show that this H-S thresholding function can improve the performance of wavelet-based denoising in PD detection.

Wavelet-based denoising, integrated with the proposed wavelet selection scheme, decomposition scale determination, and the findings in this chapter, has been applied to simulated PD signals, results demonstrate that the performance of wavelet-based technique in PD denoising are remarkably enhanced. This further indicates that the wavelet-based technique with these new algorithms applied has a great potential to extend its range into PD detection in power cable systems.

As discussed in Chapter 4 and Chapter 5 for proposed wavelet selection scheme and decomposition scale determination, real PD signals are required for further investigation of this SCR-based threshold estimation, together with the H-S thresholding function due to the variations of PD waveforms in practice. Real PD signals will serve as this function in Chapter 8. However, it may still have limits in real PD denoising due to the limited range of PD waveforms used.

7 Comparison of Wavelet-based PD Denoising using Empirical Mode Decomposition (EMD)

7.1 Introduction

Wavelet-based PD denoising has been studied for decades and still has its inherent challenges, e.g., selection of base wavelet, scale determination, and threshold estimation. Although a series of improvements have been proposed in previous chapters, it is necessary to note that the wavelet is selected from a predefined wavelet library in WEBWSS. Wavelet-based denoising is often criticized for its non-adaptivity due to the use of a predefined wavelet for the signal under processing. This predefined wavelet in some sense cannot be a 'best' match of the signal of interest. Accordingly, signal distortion inevitably occurs and the extent to which the distortion of the denoised signal depends on the matching degree of the wavelet and the signal of interest. This challenge of wavelet-based denoising is the driving force for the investigation of a new technique that can denoise the signal without any prior knowledge or assumption. That is, the new technique is completely adaptive and automated for signal expansion and reconstruction. Empirical mode decomposition (EMD) is such a technique that can decompose the signal with the base function derived from the signal itself. Since the advent of EMD, it has been widely used in the area of signal denoising due to its advantages.

EMD can decompose a given signal into a number of IMFs, from low order mode with high frequency to high order mode with low frequency. In other words, each IMF occupies lower frequencies locally in the time-frequency domain than its preceding ones. Note that the expansion of a given signal is just the first step for EMD-based denoising. The goal of the application of EMD is to split the signal into IMFs, and thus, is convenient for the following thresholding and reconstruction. Similar to wavelet-based denoising, the reduction of the annoying noise presented in the signal by EMD-based denoising requires the application of thresholds to IMFs. Generally, noise is dominated in low order IMFs while signal is mainly presented in high order IMFs. However, whether a specific IMF that carries useful information or noise is not known. Normally, a technique termed relative mode selection

(RMS) is applied to find out the relevant and irrelevant IMF modes, which has been claimed to be significant in EMD-based denoising. In turn, thresholds are only applied to the relative IMFs that are mainly dominated by signal.

Due to the high demand for EMD-based denoising, a number of techniques have been developed for RMS. Noise model-based RMS (NMRMS) was introduced in [139]. The energy distribution of the IMFs of a given noisy signal x and the associated white noise generated by a noise-only model with equal length of x will present difference from certain mode, as shown in Figure 3-18. Generally, the mode that begins to present this difference is adopted as the mode boundary. A stricter criterion for NMRMS (termed SNMRMS) is adopted in [145], [153], [187], i.e., a confidence interval is used as a threshold to select relevant modes and only the IMFs that have energies higher than the threshold will be retained for further processing. In NMRMS and SNMRMS, a noise-only model is required for noise energy estimation, which is usually not the case in practice. To avoid the dependence of a noise-only model, a correlation-based RMS (CRMS) was proposed in [13], [188], [189]. The correlation coefficient between the original noisy signal and IMF modes is adopted as a criterion for the selection of the mode boundary. Although a noise-only model is not involved, this method is very unstable for the noisy signals with different SNRs as the correlation coefficient between the noisy signal and the first IMF is too strong or too weak [188]. Based on this, the motivation of this chapter is to find a new RMS derived from the signal itself, and thus, can provide an alternative to the existing methods.

The basic idea underlying RMS is that most of important information of the signal is often concentrated on the higher order modes and decreases toward lower order ones. In other words, the information of noise is often presented in lower order modes. As mentioned in Chapter 4, Shannon entropy can measure the randomness or the degree of disorder of a signal. The energy entropy of IMF modes is derived from Shannon entropy and can also reflect the linked randomness or degree of disorder. As such, the energy entropy of IMF modes is used as a new criterion for the selection of the mode boundary. This new criterion is derived from the signal or data itself, and thus, it is an adaptive and data-driven method for the distinction of relevant and irrelevant modes for EMD-based PD denoising.

However, it is very interesting to find that RMS is not as important as expected in EMD-based PD denoising during the investigation of the effect of RMS on denoising results. Based on the threshold estimation introduced later, thresholding all IMFs can present better denoising results than any RMS mentioned above. Details regarding RMS in EMD-based PD denoising are discussed in the following section.

7.2 Relative Mode Selection (RMS) for EMD in PD Denoising

Given a noiseless signal $s(t)$ contaminated by white noise $n(t)$, the noisy signal $x(t)$ can be expressed by

$$x(t) = s(t) + n(t) \tag{7-1}$$

The denoising of $x(t)$ is to find an estimation $\tilde{x}(t)$ of the noiseless signal $s(t)$. For EMD-based denoising, one of the critical steps is to distinguish an IMF if it is a relevant or irrelevant mode. This step is so-called relative mode selection. When RMS is applied to IMFs, together with a thresholding technique, an estimation $\tilde{x}(t)$ can be given by

$$\tilde{x}(t) = \sum_{k=i_{th}}^{n} \tilde{h}_k(t) + r_n(t) \tag{7-2}$$

where i_{th} is the mode boundary, which means IMFs from i to n primarily contain the signal information, $\tilde{h}_k(t)$ is the thresholded IMFs. Based on (7-2), the mode boundary has a major effect on the estimation of $s(t)$. As a result, the techniques used to find this mode boundary will be discussed in this section.

7.2.1 Noise Energy Estimate-based RMS

The purpose of significance IMF test implemented by Wu in [153] was to establish a method of assigning statistical significance of information content for IMF components from any noisy data or signal. Numerous experiments on white noise using EMD reveal that the EMD behaviors like a dyadic filter for the decomposition of white noise. This finding is further developed to propose a noise-only model by Flandrin *et al.* in [145], [187] and Kopsinis *et al.* in [139] for EMD-based denoising.

Due to a dyadic filter structure of EMD-based expansion for white noise, the power spectra of

the IMFs exhibit self-similar characteristics, except the first IMF [139], [153]. The self-similar characteristics can also be presented by the IMF energies E_k, which are linearly decreased in a semilog diagram of, e.g., $\log_2 E_k$, with respect to k for $k \geq 2$ [139]. This appealing feature can be seen from Figure 7-1 that the IMF energies E_k linearly decrease from 2nd IMF to 10th IMF of EMD-based decomposition of white noise. The IMF energies shown in Figure 7-1 also indicate that the first IMF carries the highest amounts of energy. EMD is an effective and simple algorithm but lacking theoretical frames, as shown in Figure 3-16. As such, the sifting iterations using EMD vary due to the variation of extrema numbers of different signals. To further investigate the appealing feature mentioned above, white noise is decomposed through EMD using 1 to 15 sifting iterations. The corresponding results of the IMF energies $\log_2 E_k$ are illustrated in Figure 7-2. Obviously, the linear decrease of $\log_2 E_k$ from the 2nd IMF to the 10th IMF can be seen in EMD decompositions using different sifting iterations. Based on this, a noise-only model was proposed in [139] and [187] to estimate the IMF energies of white noise by

$$E_k = \frac{{E_1}^2}{\beta}\rho^{-k}, k = 2,3,4,\ldots \tag{7-3}$$

where ${E_1}^2$ is the energy of the first IMF, and $\beta = 0.719,\ \rho = 2.01.$

The first IMF of a given noisy signal through a EMD-based decomposition is considered a noise-only IMF [139]. For a PD signal contaminated by white noise, the first noise-only IMF is quite important as the subsequent IMF energies of white noise in the PD signal can be estimated using (7-3). As such, the IMF energies of a noisy PD signal and its linked white noise can be obtained for searching the mode boundary. The IMF energies log_2(energy) of noisy PD signals s_1 and s_2 contaminated by white noise with SNR = -5, are used as examples to explain NMRMS in Figure 7-3. For noisy s_1, as shown in Figure 7-3 (a), a significant difference between the IMF energies of real signal and the noise-only model initiates at the 5th IMF. As a result, IMFs 5-11 mainly contain the information of s_1. For noisy s_2, as shown in Figure 7-3 (b), IMFs 1 -3 and the 7th IMF are noise-only IMFs, and thus, will be discarded for further processing. Based on this, the mode boundary selected by NMRMS is the mode that the energy of real signal remarkably deviates from the noise-only model.

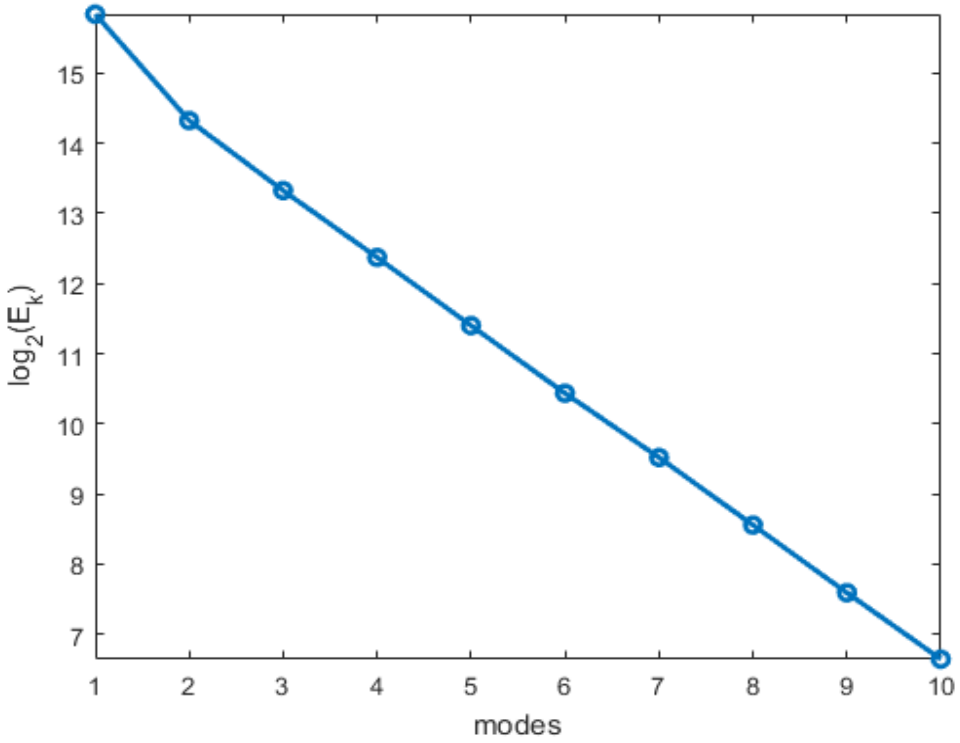

Figure 7-1 The IMF energies E_k linearly decrease in $log_2(E_k)$ with respect to k>1 due to a dyadic filter structure of EMD-based decomposition of white noise

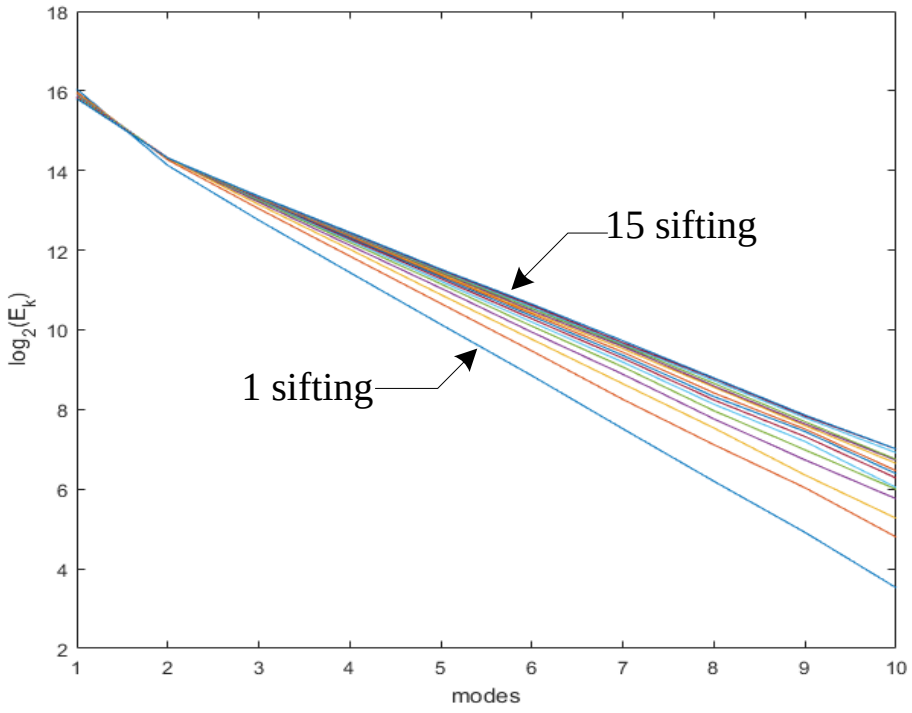

Figure 7-2 The estimated energies of IMFs that correspond to EMDs using 1 to 15 sifting iterations

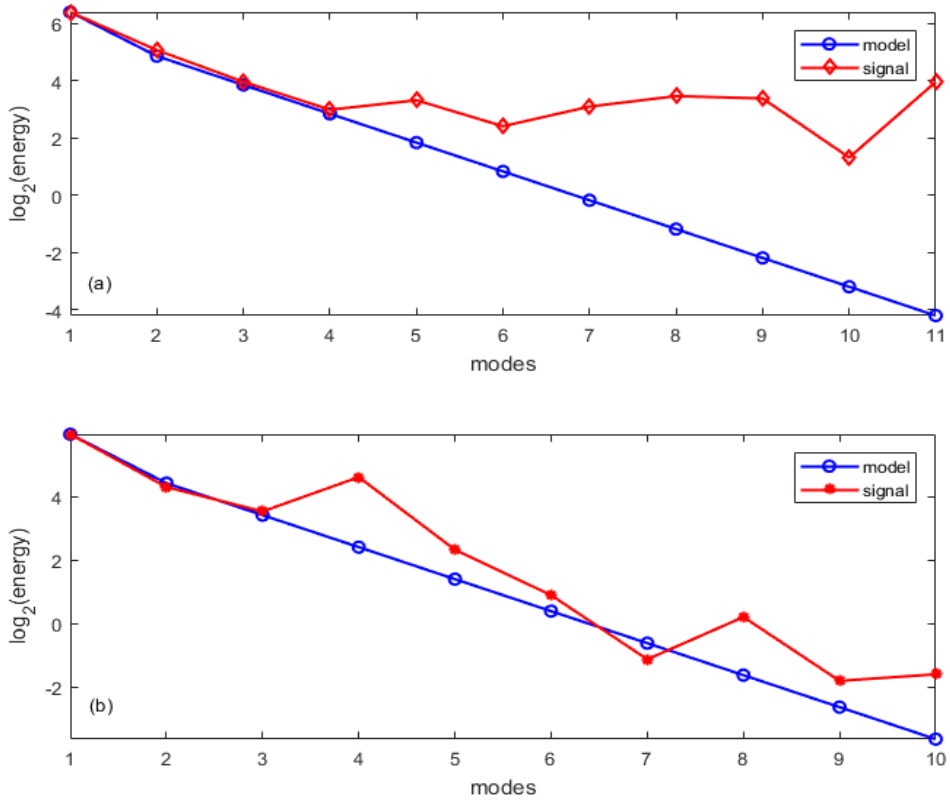

Figure 7-3 (a) The IMF energies of noisy s_1 and the noise-only model, (b) the IMF energies of noisy s_2 and the noise-only model

For SNMRMS, a confidence interval is applied and functions as the threshold to select the IMFs that have energies higher than that threshold. To have a 99% confidence interval, the threshold can be calculated by the equation [187]

$$log_2(log_2(T_k/E_k)) = 0.46 \cdot k - 1.919 \quad (7\text{-}4)$$

where T_k is the threshold for the k^{th} IMF. Noisy PD signals used in Figure 7-3 are also applied to explicate SNMRMS. Figure 7-4 illustrates the IMF energies of noisy PD signals, together with the noise-only model and 99% confidence intervals. Based on SNMRMS, only IMFs 5, and 7-9 are retained in noisy s_1 and only IMF 4 is kept in noisy s_2 for further processing. Differences of the remained IMFs can be seen in both noisy PD signals between NMRMS and SNMRMS.

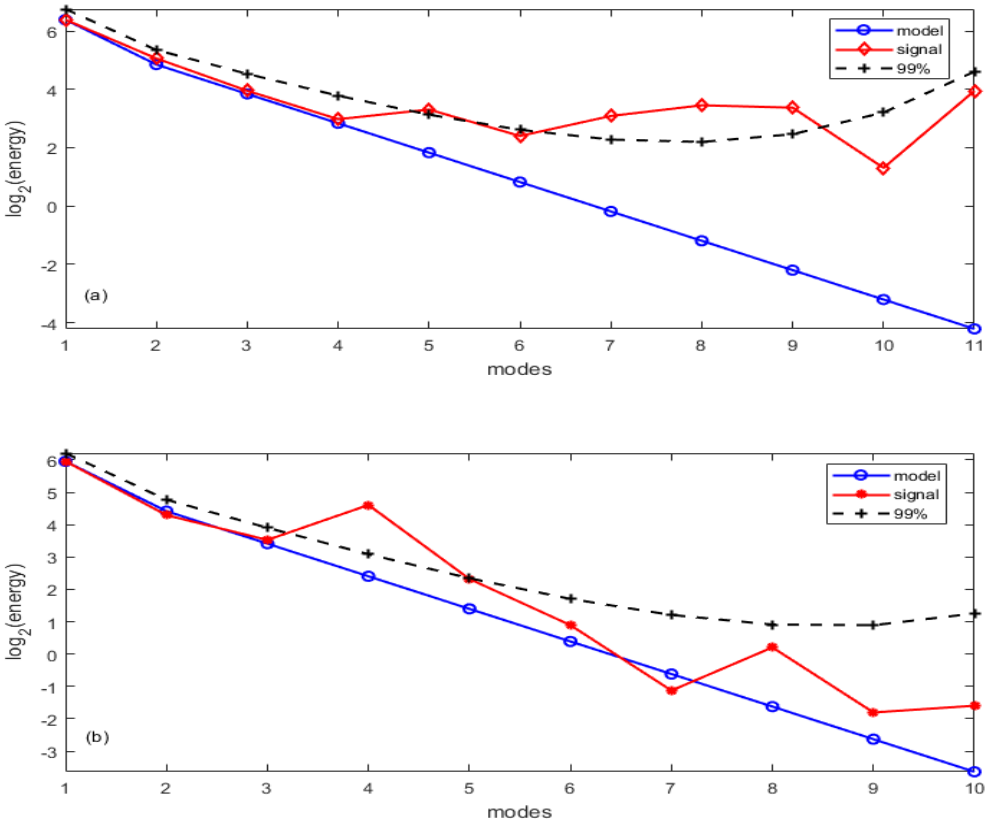

Figure 7-4 (a) The IMF energies of noisy s_1 together with the noise-only model and 99% confidence intervals, (b) The IMF energies of noisy s_2 together with the noise-only model and 99% confidence intervals

7.2.2 Correlation-based RMS

The correlation-based RMS (CRMS), as its name implies, adopts correlation coefficient as a criterion to identify the degree of relevance between the IMFs that carry information and the original signal. Given that i_{th} is the mode boundary for EMD-based denoising of an arbitrary signal $x(t)$, the summation of relevant modes can be defined by [188], [189]

$$x_m(t) = x(t) - \sum_{k=1}^{m} h_k(t) \tag{7-5}$$

where $m = i_{th} - 1$. In turn, the correlation coefficient between $x(t)$ and $x_m(t)$ can be calculated by [188], [189]

$$\rho(m) = \sum_{t=1}^{N} x(t) \cdot x_m(t) / \left(\sqrt{\sum_{t=1}^{N} x^2(t)} \cdot \sqrt{\sum_{t=1}^{N} {x_m}^2(t)} \right) \tag{7-6}$$

where N is the length of the signal $x(t)$, m is the mode number when $\rho(m)$ starts to be less than an empirical value in the range of [0.75 0.85].

As mentioned before, the CRMS is not stable due to the variations of SNRs of the original noise signal. As did for NMRMS, noisy PD signals with SNR = -5 will be used to test the feasibility of CRMS in EMD-based PD denoising. The correlation coefficients for both noisy PD signals are shown in Figure 7-5. Based on the criterion of CRMS, all the IMFs of both noisy s_1 and s_2 are retained for further processing if the empirical value is set to 0.8. It also indicates that the first IMF has a strong correlation with the original signal when the SNR is low.

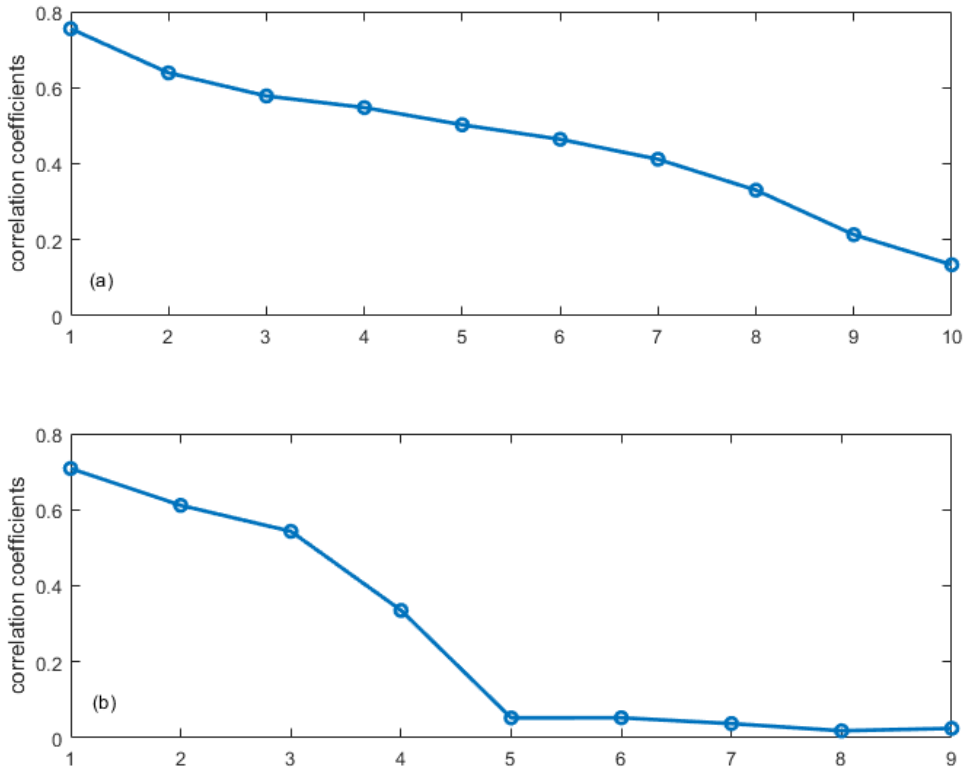

Figure 7-5 Correlation coefficients for (a) noisy s_1, (b) noisy s_2

7.2.3 A Novel Energy Entropy-based RMS

7.2.3.1 Energy Entropy

The definition of energy entropy is similar to that of wavelet entropy in Chapter 4. For a *N*-sample signal $x(t)$, the distribution of energy probability for $x(t)$ can be derived by

$$p_j = \frac{x^2(j)}{\sum_{j=1}^{N} x^2(j)} \tag{7-7}$$

where $\sum_j p_j = 1$. Then, the energy entropy of $x(t)$ can be calculated by

$$EE = -\sum_j p_j \cdot ln\, p_j \tag{7-8}$$

Similar to the Shannon entropy in information theory, the energy entropy of $x(t)$ can measure the randomness or the degree of disorder of $x(t)$. For signals corrupted by white noise, the energy entropy can also reflect the presence of white noise in the signal. The lower the entropy value, the less white noise is presented in that signal.

7.2.3.2 Energy Entropy-based RMS

The energy entropy-based RMS (EERMS) is inspired by CRMS. The energy entropy is measured by gradually removing the IMFs (fine to coarse) from the original signal. Given that i_{th} is the mode boundary for EMD-based denoising of $x(t)$ using EERMS, the summation of relevant modes can be calculated based on (7-5). The energy entropy of $x_m(t)$ can be calculated based on (7-7) and (7-8). The value of m is determined when $x_m(t)$ has the minimum energy entropy. The minimum energy entropy indicates that the randomness of $x_m(t)$ is minimum, i.e., the noise remained in $x_m(t)$ is minimum. Noisy s_1 and s_2, as used for NMRMS and CRMS, are also used as examples to show the mode boundary selected by EERMS in Figure 7-6 (b) and (d). Based on EERMS, IMFs from 5 to 9 for noisy s_1, and 4 to 8 for noisy s_2 are remained for further thresholding and reconstruction.

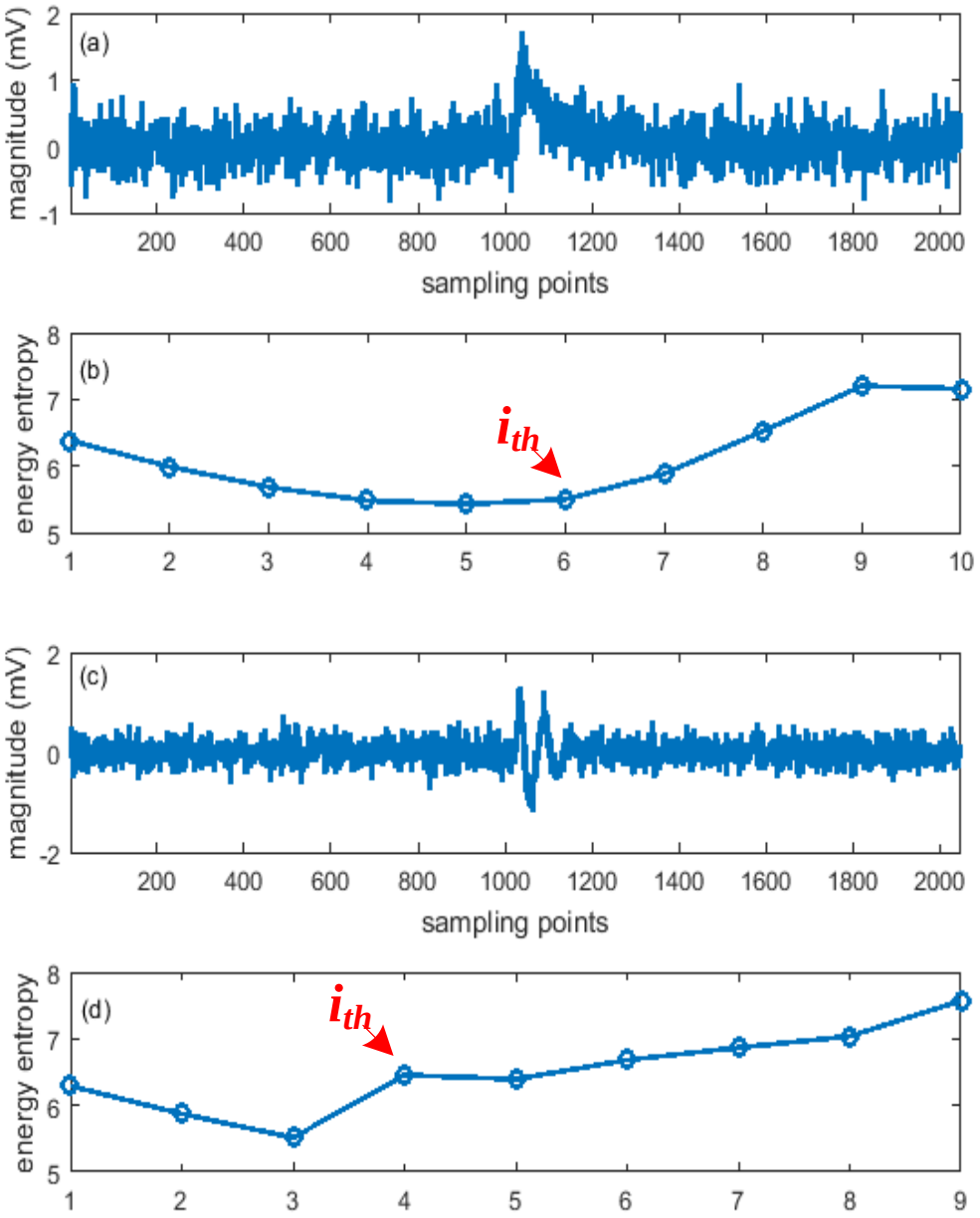

Figure 7-6 (a) noisy s_1, (b) the mode boundary i_{th} selected by EERMS for (a) (c) noisy s_2, (d) the mode boundary i_{th} selected by EERMS for (c)

7.3 EMD for Partial Discharge Denoising

The EMD-based denoising methods, e.g., EMD-TR, EMD-DT, and EMD-IT, have been fully discussed in Chapter 3. The denoising results suggest that EMD-IT outperforms EMD-TR and EMD-DT in signal denoising. However, the selection of the threshold applied in these methods has not been mentioned. This will be mainly discussed in this section. With selected

thresholds, the RMSs and EMD-based denoising methods adopted in the area of PD denoising will be compared.

7.3.1 Threshold Estimation

The threshold applied in EMD-based denoising, as derived from (3-13) and (3-14), is a direct use of the threshold defined as the universal threshold in wavelet-based denoising. EMD-based denoising may have a serious problem in this case due to two major differences between wavelet- and EMD-based denoising [139]. Firstly, the object that needs to be processed by the threshold is different, i.e., the threshold is applied to the wavelet components in wavelet-based denoising, while the threshold in the EMD case is applied to each IMF, which basically contains the signal portion in each adaptive subband. As a result, the thresholding in EMD-based denoising is equivalent to the procedure in the wavelet method that thresholding is performed on the reconstructed signal after synthesizing the function on each scale separately. Secondly, the IMF samples are not Gaussian distributed with variance equal to the noise variance, which is, however, the basic idea for the introduction of the universal threshold in wavelet-based denoising. In other words, noise can be estimated through the variance of wavelet components irrespective of scale. In fact, the noise contained in each IMF is colored, leading to the difference of noise energy concentrated on each mode [139]. Based on this, a new formula was proposed in [139] to calculate multiples of the IMF-dependent universal threshold. The formula for the threshold is defined as

$$\lambda_k = C\sqrt{E_k \cdot 2\,ln\,N} \tag{7-9}$$

where C is a constant, N is the length of the signal, and E_k is the IMF energies that can be computed based on (7-3). Note that E_k in (7-9) is estimated based on white noise, it is the main cause of ineffectiveness of EMD-based technique in PD denoising when other type of noise, e.g., DSI, is involved.

7.3.2 Denoising Results Analysis

7.3.2.1 Denoising Results of Various RMSs

Four methods, i.e., NMRMS, SNMRMS, CRMS, and EERMS, have been presented in Section 7.2. EERMS is a novel method proposed by the author for searching the mode boundary in EMD-based denoising. Tests are implemented to evaluate the performance of this novel RMS with comparisons of the other RMSs.

PD signals s_1 and s_2 are contaminated by white noise with SNR = -5. With an EMD-based expansion, noisy s_1 is split into 10 IMFs and one residual, and noisy s_2 is decomposed into 9 IMFs and one residual. The relevant IMFs selected by various RMSs are listed in Table 7-1. Note that all the IMFs are selected for denoising by the use of CRMS. It is in agreement with the drawback claimed for CRMS, i.e., the correlation coefficient between the original PD signals and the first IMF is too strong when SNR of noisy PD signals is very low. With the relevant IMFs, the threshold estimated by (7-10) and EMD-IT are applied to obtain the denoising results. Parameters, i.e., ME, MSE, XCORR, and SNR (after denoising), used for the performance evaluation of various RMSs in EMD-based PD denoising are listed in Table 7-2. It is found that the novel RMS cannot provide better denoising results than the others for noisy s_1, while it can provide the same denoising results as NMRMS and CRMS for noisy s_2. In the meantime, CRMS can provide the best denoising results for both noisy s_1 and s_2 among all the RMSs.

Table 7-1 IMFs of noisy s_1 and s_2 selected by various RMSs for further denoising (SNR=-5)

	RMS			
	NMRMS	**SNMRMS**	**CRMS**	**EERMS**
s_1	5-10	5,7,8,9	1-10	6-10
s_2	4,5,6,8,9	4	1-9	4-10

Table 7-2 Parameters used for the performance evaluation of various RMSs in EMD-based denoising

s_1	Parameters				s_2	Parameters			
	ME	MSE	XCORR	SNR		ME	MSE	XCORR	SNR
NMRMS	0.533	0.0067	0.8533	5.182	**NMRMS**	0.517	0.0049	0.8892	4.762
SNMRMS	0.552	0.0078	0.8202	4.562	**SNMRMS**	0.540	0.0050	0.8883	4.650
CRMS	0.533	0.0067	0.8533	5.182	**CRMS**	0.517	0.0049	0.8892	4.762
EERMS	0.608	0.0081	0.8127	4.368	**EERMS**	0.517	0.0049	0.8892	4.762

With the attempt to further test the novel RMS in EMD-based PD denoising, PD signals are simulated with different SNRs to represent PD signals detected in different noisy environment. For this purpose, the SNRs of simulated PD signals are set to -3, -1 ,1, 3, and 5. The parameters, as used in Table 7-2, are applied to evaluate the performance of various RMSs and illustrated in Figure 7-7, Figure 7-8, Figure 7-9 and Figure 7-10, respectively. Based on these diagrams, the novel RMS can have the same denoising results as NMRMS or SNMRMS for noisy s_2 with low SNRs, and it can present better denoising results for noisy s_2 with high SNRs (over 5 in Figures). Also, it can be seen that CRMS can present better denoising results of both noisy PD signals when the SNR of original signals is low (less than 1 in Figures) while NMRMS can provide better denoising results when the SNR is relatively high (over 1 in Figures). It is interesting to find that CRMS can deliver better denoising results when SNR of the noisy PD signals is low, but worse denoising results when the SNR is high. As known, the correlation coefficients between original PD signals and IMF modes are highly dependent on the SNR of the original PD signals. In other words, all the IMFs can be selected by CRMS for denoising when the SNR is very low, e.g., IMFs selected for both noisy s_1 and s_2 with SNR = -5 (see Table 7-1). However, fewer IMFs will be selected by CRMS if the SNR is relatively high, as shown in Table 7-3, where the SNR for both noisy s_1 and s_2 is 5. Noisy s_1 is decomposed into 9 IMFs and one residual while noisy s_2 is split into 10 IMFs and one residual. Based on the parameters for performance evaluation of RMSs and the number of IMFs remained for denoising, it indicates that better denoising results can be obtained if the number of IMFs retained is high. As such, more details regarding the effect of RMS on EMD-based PD denoising will be explored in depth.

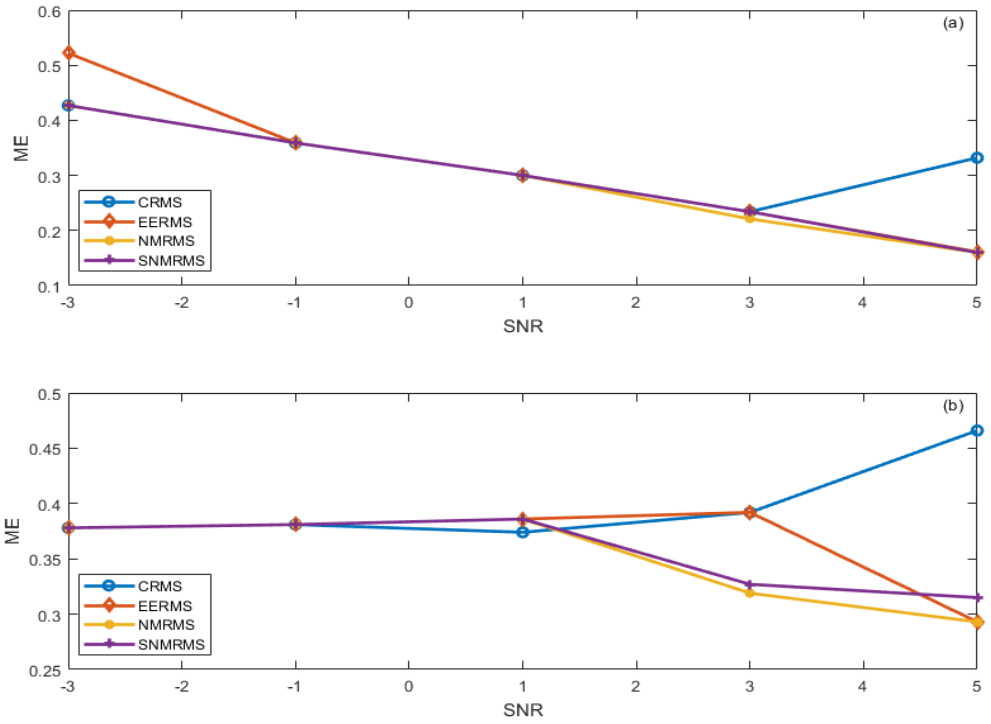

Figure 7-7 ME of denoised PD signals (a) s_1, (b) s_2 using various RMSs at different noise levels

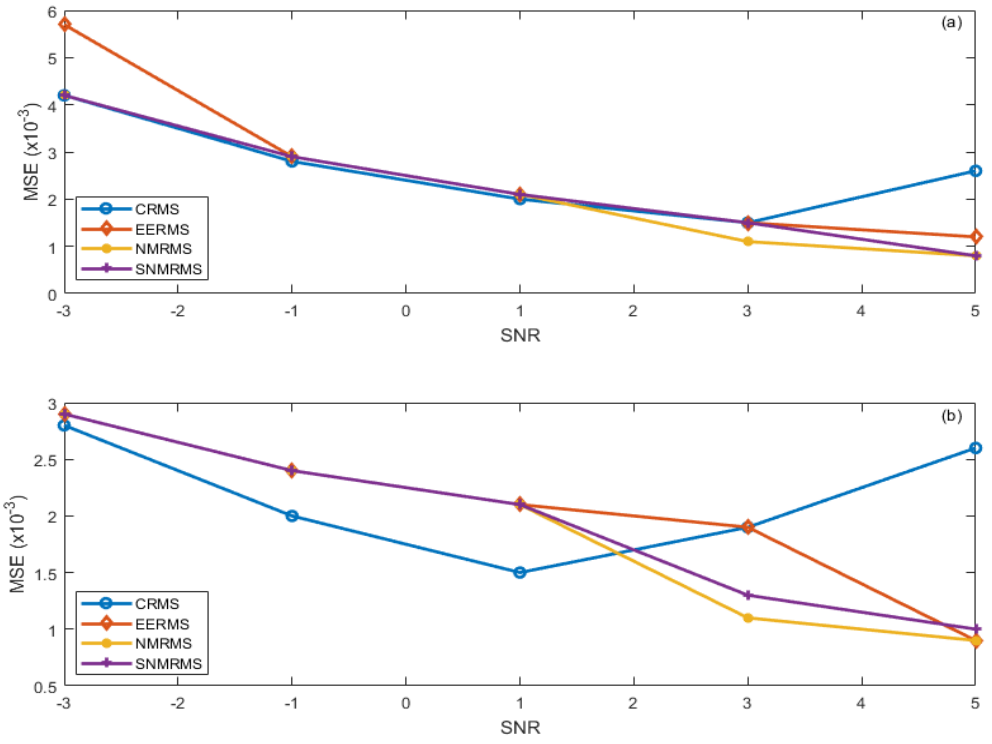

Figure 7-8 MSE of denoised PD signals (a) s_1, (b) s_2 using various RMSs at different noise levels

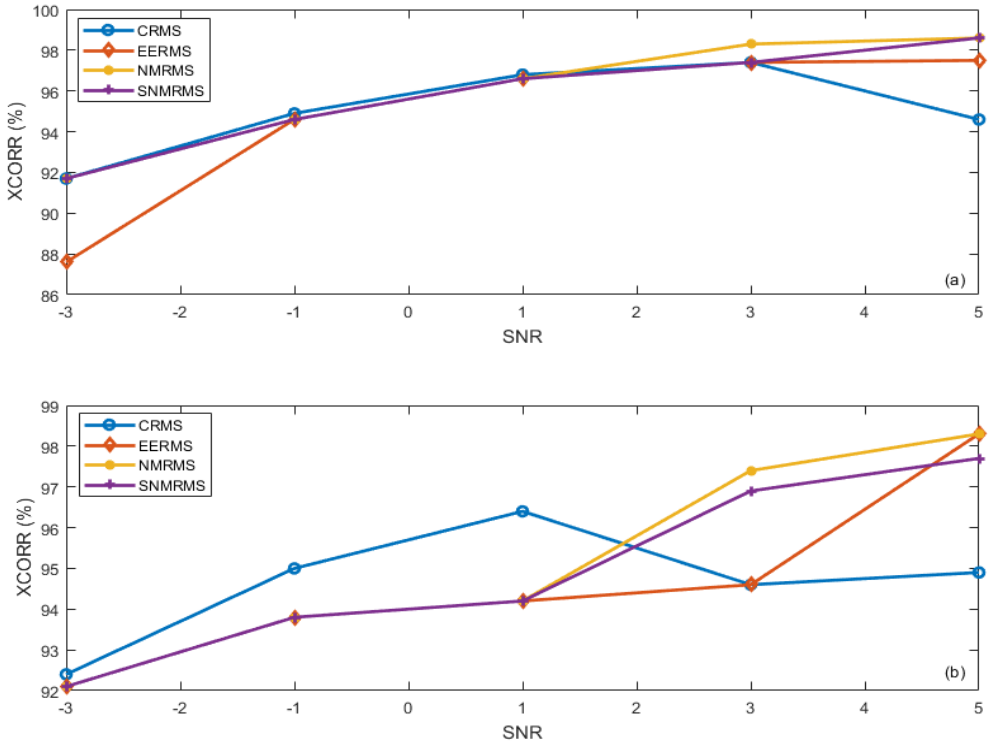

Figure 7-9 XCORR of denoised PD signals (a) s_1, (b) s_2 using various RMSs at different noise levels

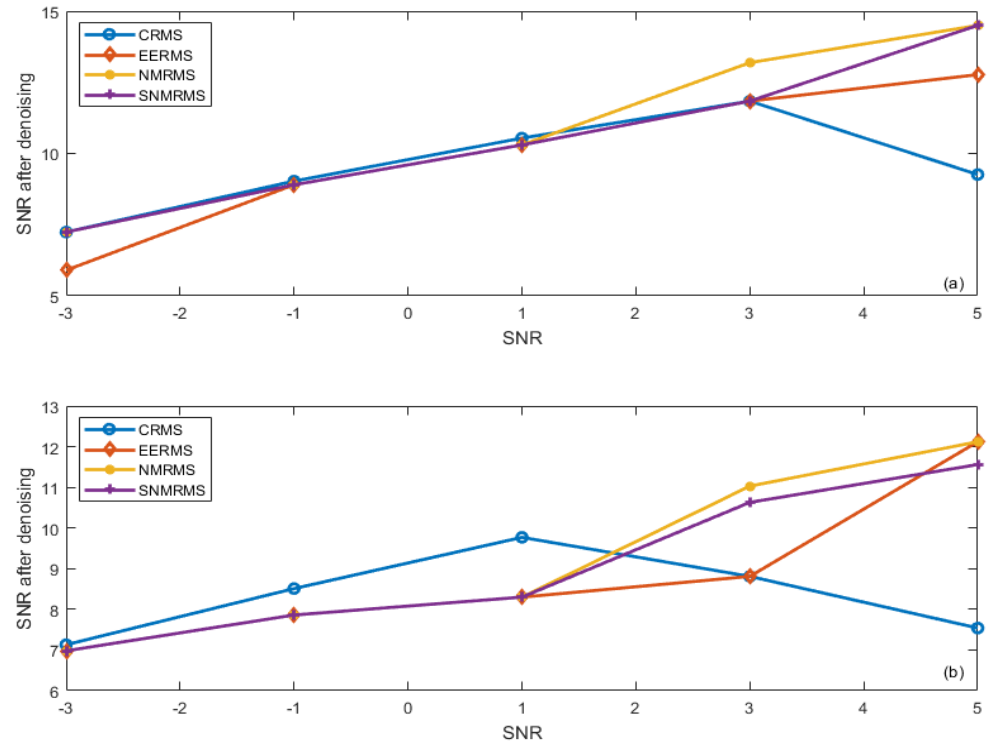

Figure 7-10 SNR of denoised PD signals (a) s_1, (b) s_2 using various RMSs at different noise levels

Table 7-3 IMFs of noisy s_1 and s_2 selected by various RMSs for further denoising (SNR =5)

	RMS			
	NMRMS	**SNMRMS**	**CRMS**	**EERMS**
$\boldsymbol{s_1}$	4-9	4-9	6-9	5-9
$\boldsymbol{s_2}$	3-10	3-5&8	4-10	3-10

To further investigate RMSs, noisy s_1 and s_2 with SNR = 1 are used as examples. 10 and 9 IMFs are obtained through EMD-based expansions of noisy s_1 and s_2, respectively. The IMFs selected by various RMSs are listed in Table 7-4. Also, the IMFs and the associated thresholds based on (7-10) for noisy s_1 and s_2 are illustrated in Figure 7-11 and Figure 7-12.

Table 7-4 IMFs of noisy s_1 and s_2 selected by various RMSs for further denoising (SNR=1)

	RMS			
	NMRMS	**SNMRMS**	**CRMS**	**EERMS**
$\boldsymbol{s_1}$	5-10	5-10	3-10	5-10
$\boldsymbol{s_2}$	4,5	4,5	2-9	4-9

The red line at each IMF in both Figure 7-11 and Figure 7-12 is the threshold estimated for that mode. Based on the relationship between threshold and IMF, IMF4 to IMF10 of noisy s_1 need to be remained for denoising while IMF1 to IMF3 are noise-only IMFs in Figure 7-11. As a result, IMFs 4-10 should be selected by an appropriate RMS. Table 7-4 demonstrates that no RMS investigated can be considered an optimal technique for the selection of relevant modes. However, the IMFs selected by CRMS include 4-10 and CRMS provides better denoising results than the others, as shown in Figure 7-7 to Figure 7-10 when SNR = 1. This indicates that the redundant IMFs selected by CRMS has no effect on the denoising results. Equally, IMF3 to IMF5 of noisy s_2 need to be selected for denoising based on Figure 7-12. IMF3 is discarded as it is assumed to be a noise-only IMF by NMRMS, SNMRMS, and EERMS, while it is retained by CRMS. CRMS coincidently presents the best denoising results than the others for noisy s_2 (see Figure 7-7 to Figure 7-10 when SNR = 1), even if IMF2, IMF6 to IMF9 are also retained. It is found that the IMFs retained by CRMS for both

noisy PD signals contain the IMFs that need to be denoised (see Figure 7-11 and Figure 7-12). Also, the redundant IMFs selected by CRMS have no effect on the denoising results as they are noise-only IMFs (see Figure 7-11 and Figure 7-12). When the SNR is relatively high, e.g., SNR = 5, NMRMS can provide better denoising for noisy s_1 and s_2. Based on Figure 7-13 and Figure 7-14, IMFs 3-9 and 3-10 of noisy s_1 and s_2 respectively should be retained for further denoising. Table 7-3 demonstrates that only NMRMS can select the IMFs close to the numbers mentioned above, even if some IMFs are missing. Based on this, the RMS that can keep more IMFs will have better denoising result.

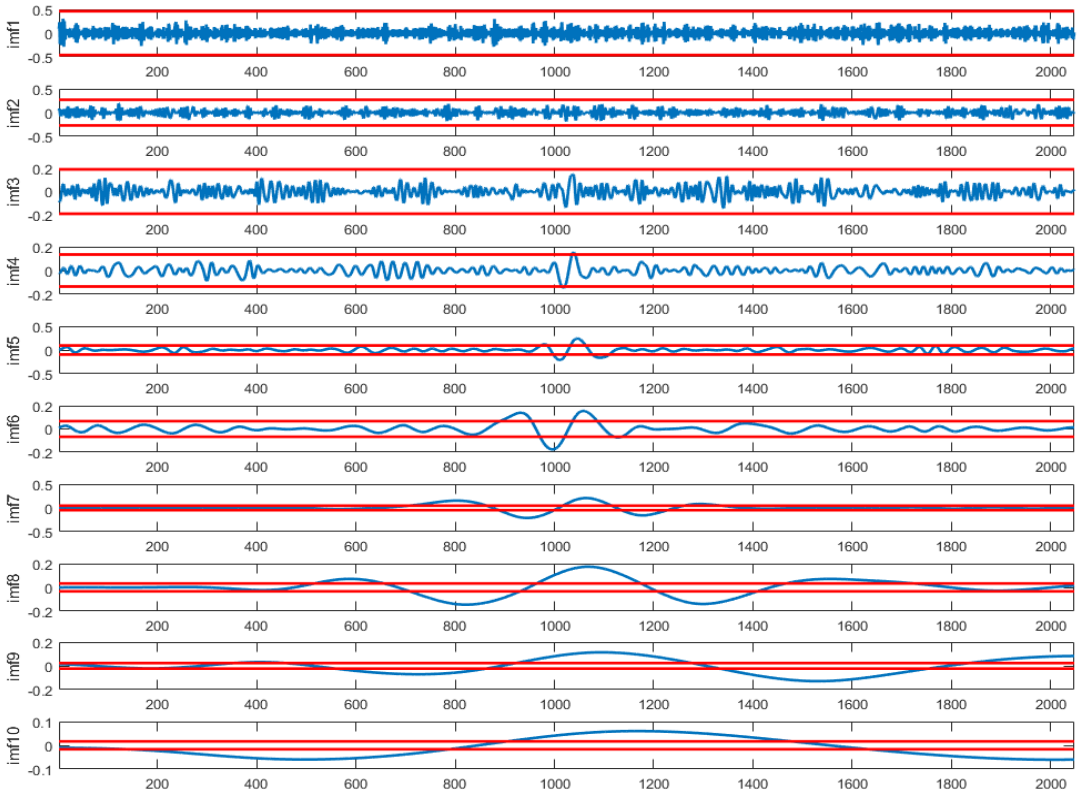

Figure 7-11 IMFs and the associated threshold for noisy s_1 with SNR = 1

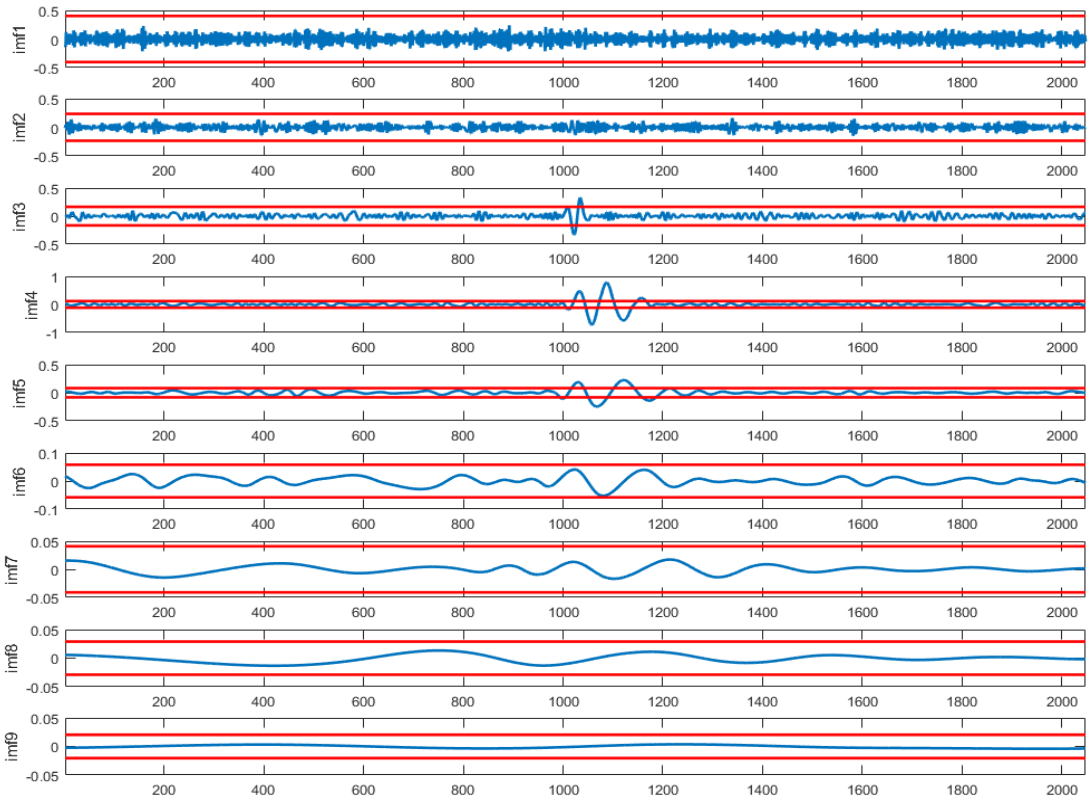

Figure 7-12 IMFs and the associated threshold for noisy s_2 with SNR = 1

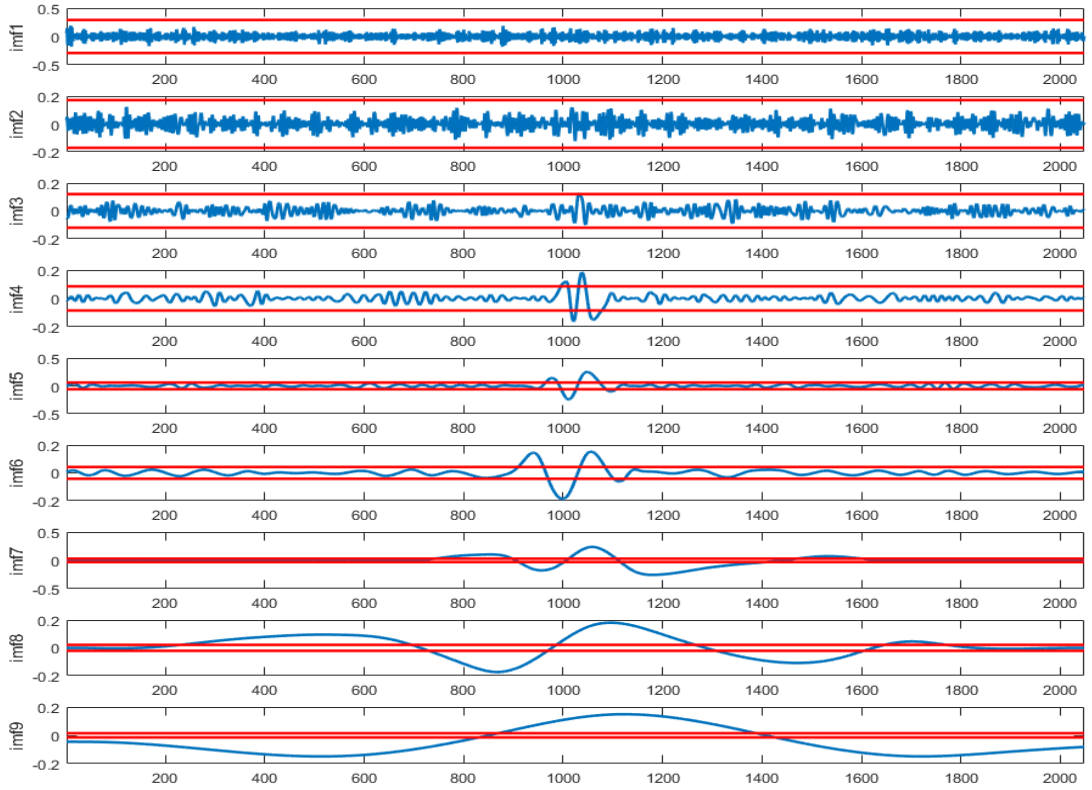

Figure 7-13 IMFs and the associated threshold for noisy s_1 with SNR = 5

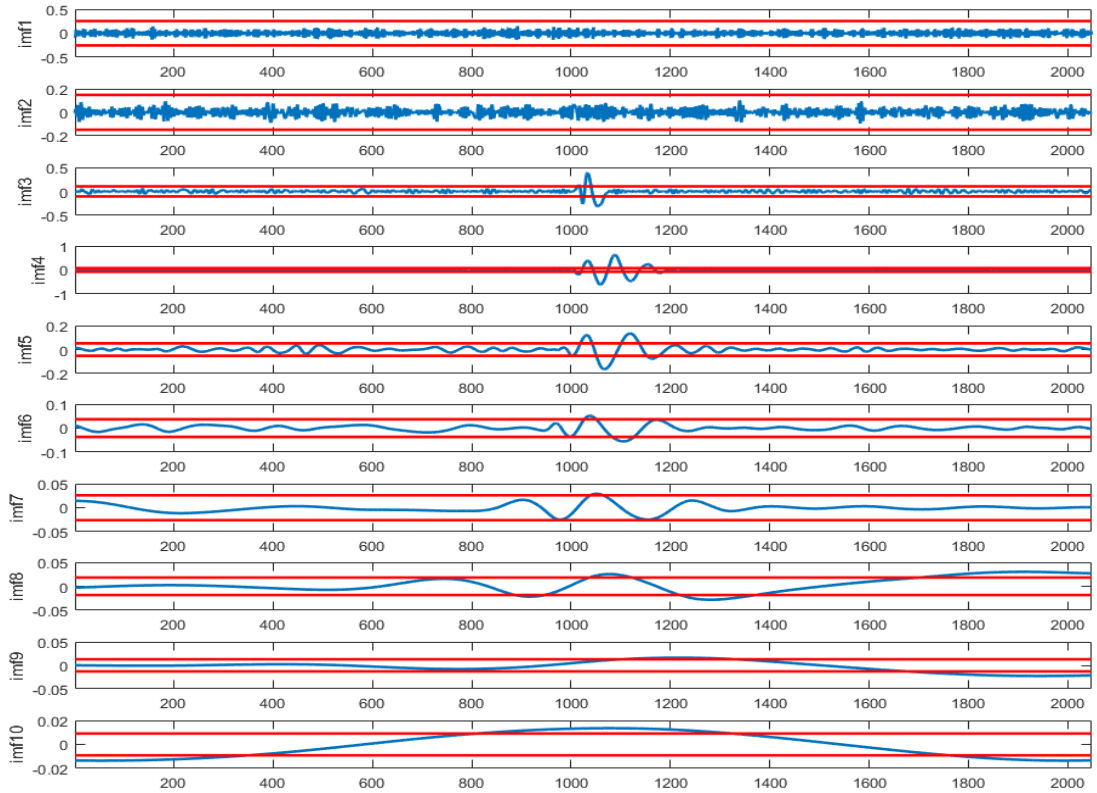

Figure 7-14 IMFs and the associated threshold for noisy s_2 with SNR = 5

The findings presented above indicate that the number of IMFs that need to be retained for further denoising varies due to the SNR of original PD signals. In other words, the RMS techniques discussed above cannot be solely applied to select the optimal number of IMFs for PD signals under various noise levels. Only the RMS that can keep sufficient IMFs will achieve better denoising results, although it may not be an optimal one. Based on this, it is recommended that all the IMFs can be remained for denoising as the redundant IMFs normally are the noise-only IMFs which have no effect on the denoising results. Accordingly, the best results of EMD-based PD denoising can be obtained in this scenario.

7.3.2.2 Denoising Results: Wavelet-based Technique versus EMD-based Technique

EMD-based denoising is applied as an alternative to wavelet-based denoising in the field of PD detection, and is claimed with the advantage of adaptivity. In this section, simulated PD signals with noise are adopted to investigate the effectiveness of wavelet-based and

EMD-based techniques for PD denoising.

In wavelet-based denoising, soft-thresholding function can cause less discontinuity in denoised PD signals than hard-thresholding function. However, peak distortion caused by soft-thresholding function in EMD-IT is more serious than the discontinuity caused by hard-thresholding function in denoised PD signals. The serious peak distortion can be seen in Figure 7-15, where PD signals s_1 and s_2 are corrupted by white noise with SNR = 0. The peak distortion can be remarkably improved by the use of hard-thresholding function in EMD-IT. Parameters listed in Table 7-5 further demonstrate that hard-thresholding function cause less distortion than soft-thresholding function in EMD-IT for PD denoising. Based on this, EMD-based PD denoising using hard-thresholding function will be applied to noisy PD signals, and then, compared to wavelet-based PD denoising.

PD signals s_1 and s_2 corrupted by white noise with SNR = -5 are used as the testing signals for EMD-based and wavelet-based PD denoising. The denoising results are illustrated in Figure 7-16 and Figure 7-17 for noisy s_1 and s_2, respectively. Figure 7-16 (c) and Figure 7-17 (c) are the denoising results using EMD-based technique, and Figure 7-16 (d) and Figure 7-17 (d) are the denoising results using wavelet-based technique. It is necessary to mention that the wavelet-based technique used for comparison is an improved one based on the previous 3 chapters. Direct observation from Figure 7-16 cannot distinguish major difference between EMD-based and wavelet-based denoising technique for noisy s_1. However, Figure 7-17 reveals that the wavelet-based PD denoising presents better results than EMD-based PD denoising for noisy s_2, since EMD-based technique cause more serious distortion in the denoised s_2. Parameters used to evaluate the performance of these two techniques in PD denoising are listed in Table 7-6, which demonstrates wavelet-based technique can present better denoising results than EMD-based technique, except the ME of noisy s_1.

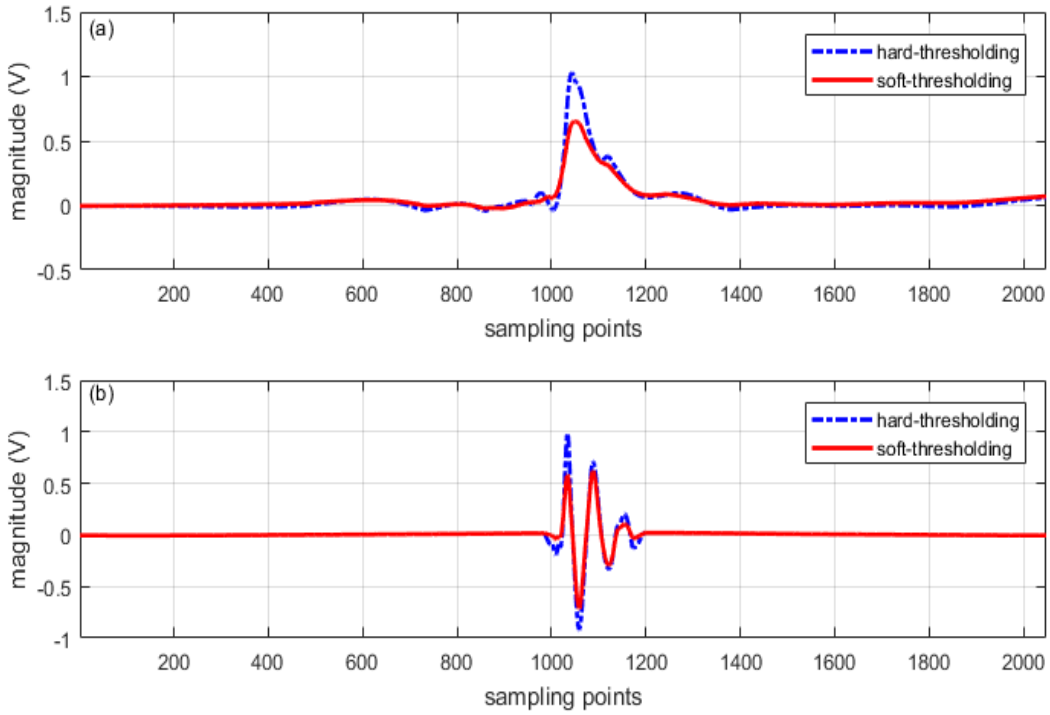

Figure 7-15 EMD-based denoising of (a) noisy s_1, (b) noisy s_2 using hard- and soft-thresholding function

Table 7-5 Parameters used for performance evaluation of EMD-based denoising through hard- and soft-thresholding function

	ME		**MSE**		**XCORR**		**SNR**	
	soft	**hard**	**soft**	**hard**	**soft**	**hard**	**soft**	**hard**
s_1	0.349	**-0.028**	0.0027	**0.0011**	0.953	**0.978**	9.19	**13.08**
s_2	0.379	**0.009**	0.0018	**0.0007**	0.958	**0.978**	9.12	**13.42**

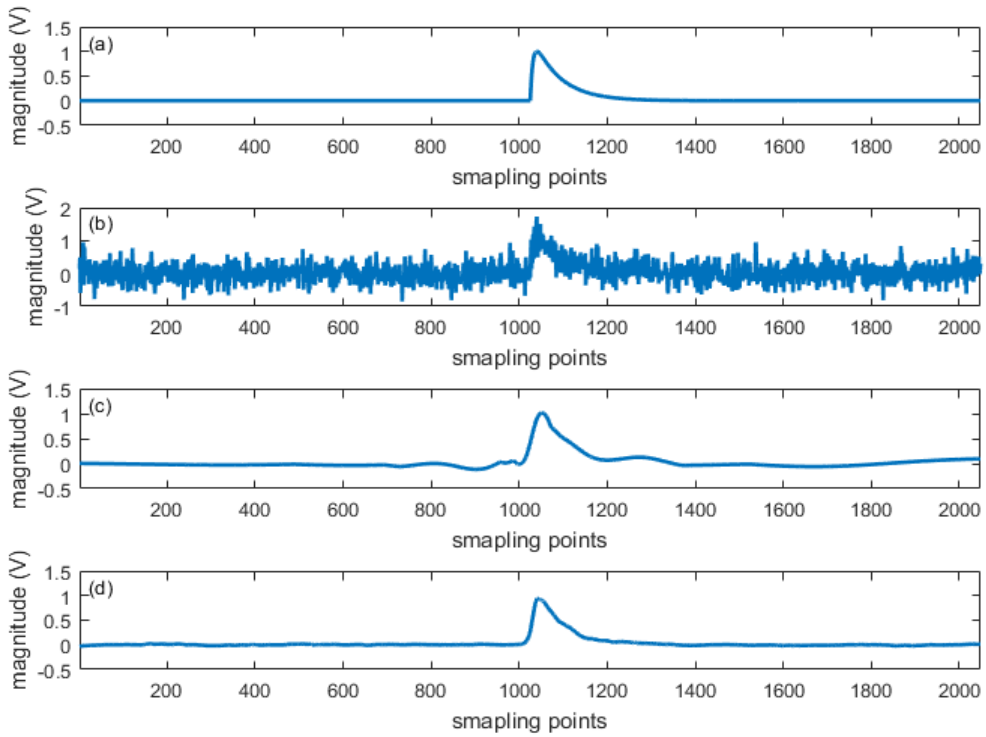

Figure 7-16 (a) s_1, (b) noisy s_1, (c) EMD-based denoising result, (d) wavelet-based denoising result

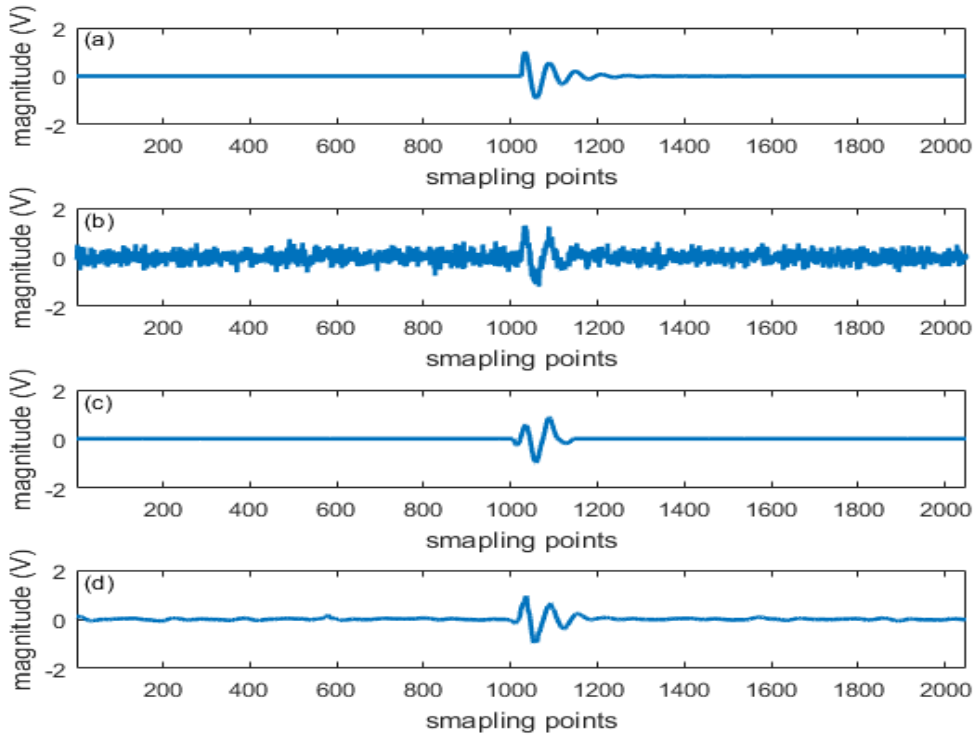

Figure 7-17 (a) s_2, (b) noisy s_2, (c) EMD-based denoising result, (d) wavelet-based denoising result

Table 7-6 Parameters used for performance test of EMD-based and wavelet-based PD denoising

	ME		MSE		XCORR		SNR	
	EMD	**Wavelet**	**EMD**	**Wavelet**	**EMD**	**Wavelet**	**EMD**	**Wavelet**
s_1	-0.017	-0.057	0.0029	**0.0025**	0.951	**0.954**	8.84	**9.95**
s_2	0.121	**0.075**	0.0029	**0.0023**	0.897	**0.931**	7.01	**8.47**

Further comparisons between EMD-based and wavelet-based PD denoising are implemented on PD signals simulated with various noise levels. The SNR of simulated PD signals is set to -3, -1, 1, 3, and 5, representing various noise levels. Figure 7-18 and Figure 7-19 illustrate the parameters used to evaluate the performance of both denoising techniques for noisy s_1 and s_2 respectively at various noise levels. For noisy s_1, neither EMD-based technique nor wavelet-based technique presents remarkable advances than the other in denoising results (see Figure 7-18). The distortions of denoised PD signals caused by both techniques are very close to each other. However, EMD-based technique provides better denoising results than wavelet-based technique for noisy s_2, as shown in Figure 7-19.

Based on the analysis above, wavelet-based denoising outperforms EMD-based denoising when the SNR of the original PD signal is very low, as shown in Figure 7-16 and Figure 7-17. With the increase of SNR, EMD-based denoising can present similar denoising results as wavelet-based denoising for noisy s_1, while the denoising results of noisy s_2 are slightly improved by EMD-based denoising.

The comparison between EMD-based and wavelet-based denoising is also implemented on PD signals with multiple pulses. spd_1 and spd_2, as shown in Figure 4-12 (a) and Figure 4-13 (a), are corrupted by white noise with SNR = -5. Denoising results of spd_1 and spd_2 using EMD-based and wavelet-based technique are delineated in Figure 7-20 and Figure 7-21, respectively. A direct observation of Figure 7-20 (c), (d) and Figure 7-21 (c), (d) show that all PD pulses of both spd_1 and spd_2 have been successfully extracted through EMD-based and wavelet-based denoising. The parameters, e.g., MSE and XCORR, of denoised spd_1 and spd_2 are listed in Table 7-7. It can be seen that EMD-based denoising has similar effectiveness as wavelet-based denoising in the extraction of spd_1, but can improve the

denoising results for spd_2 as compared to wavelet-based denoising.

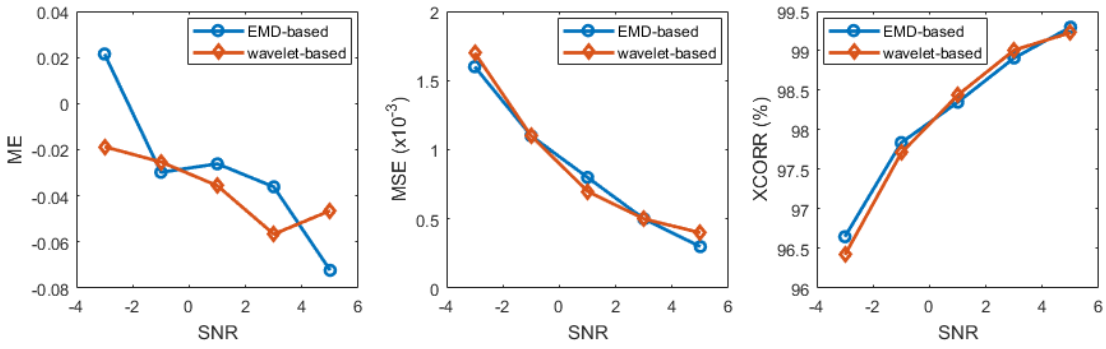

Figure 7-18 ME, MSE and XCORR of denoised s_1 using EMD-based and wavelet-based denoising

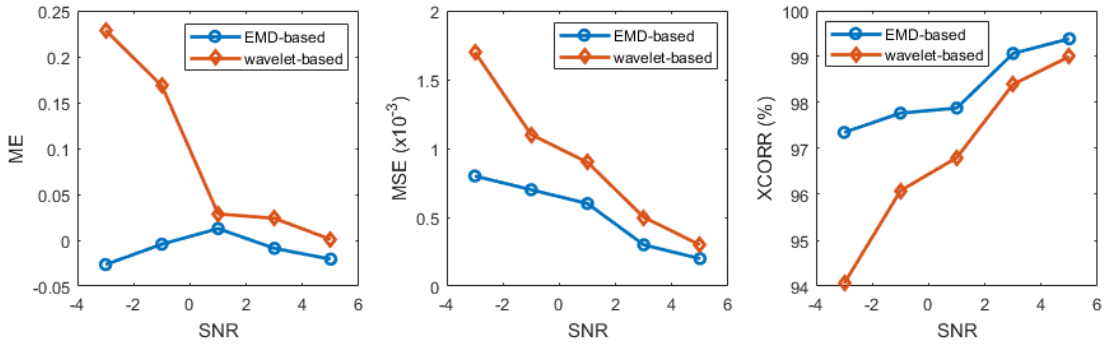

Figure 7-19 ME, MSE and XCORR of denoised s_2 using EMD-based and wavelet-based denoising

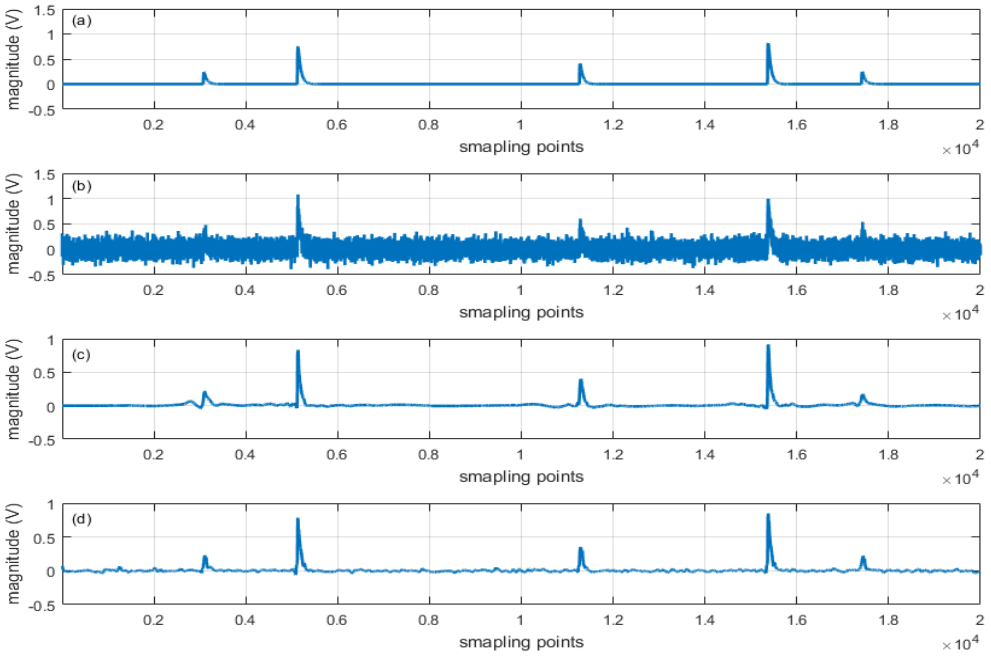

Figure 7-20 (a) spd_1, (b) noisy spd_1, (c) EMD-based denoising result, (d) wavelet-based denoising result

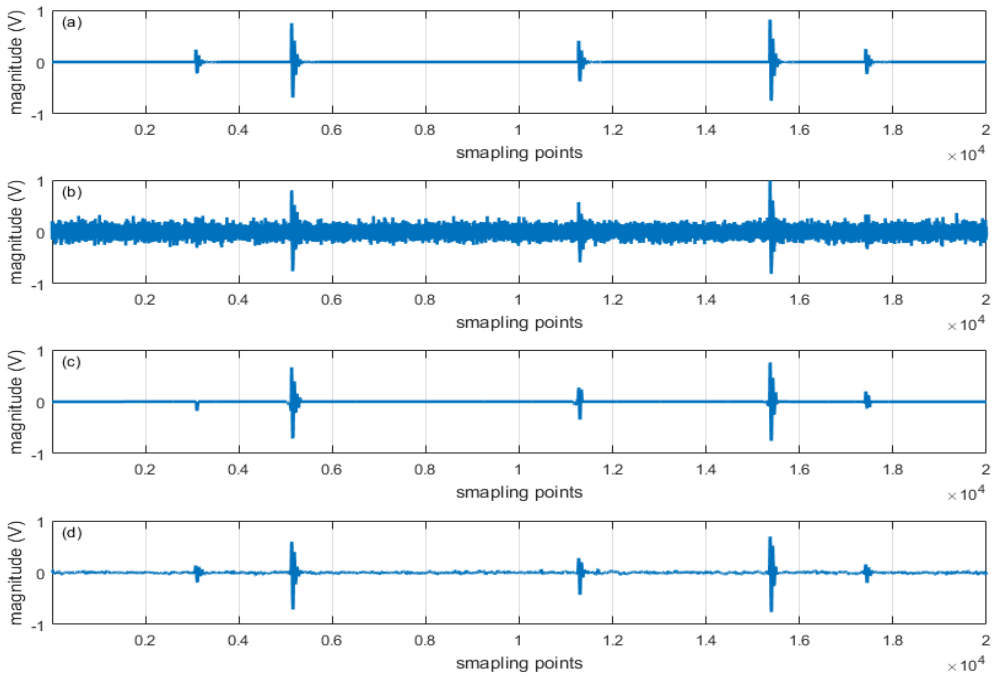

Figure 7-21 (a) spd_2, (b) noisy spd_2, (c) EMD-based denoising results, (d) wavelet-based denoising results.

Table 7-7 Parameters of denoised spd_1 and spd_2 using EMD-based and wavelet-based denoising

	MSE		XCORR	
	EMD	**Wavelet**	**EMD**	**Wavelet**
$\boldsymbol{spd_1}$	0.00026	0.00025	0.9652	0.9643
$\boldsymbol{spd_2}$	0.00019	0.00034	0.9578	0.9233

The findings in the denoising results of PD signals with single pulse and multiple pulses indicate that EMD-based technique could be an effective alternative to wavelet-based technique in the area of PD denoising. One of the most important aspects is EMD-based expansion is a data-driven and adaptive, while wavelet-based denoising requires an assumption of the wavelet. However, a close observation of the denoised PD signals shows that the PD pulses with small magnitude, e.g., the 1st, the 3rd, and the 5th PD pulse in spd_1 and spd_2, suffers serious distortion under EMD-based denoising when SNR is low, although EMD-based denoising presents a better denoising result for the whole PD signal. Figure 7-22 and Figure 7-23 illustrate the 1st, the 3rd and the 5th PD pulse of denoised spd_1 and spd_2 respectively using EMD-based denoising. It can be seen that the PD pulses with small

magnitudes suffer serious distortion after denoising, especially for DOP-type (s_2) PD pulses (see Figure 7-23). This serious distortion of PD pulse is not expected as it can negatively affect the following PD data analysis, e.g., lower the accuracy of PD defect location and pattern recognition. Note that this distortion reduces with the increase of the SNR of noisy PD signals.

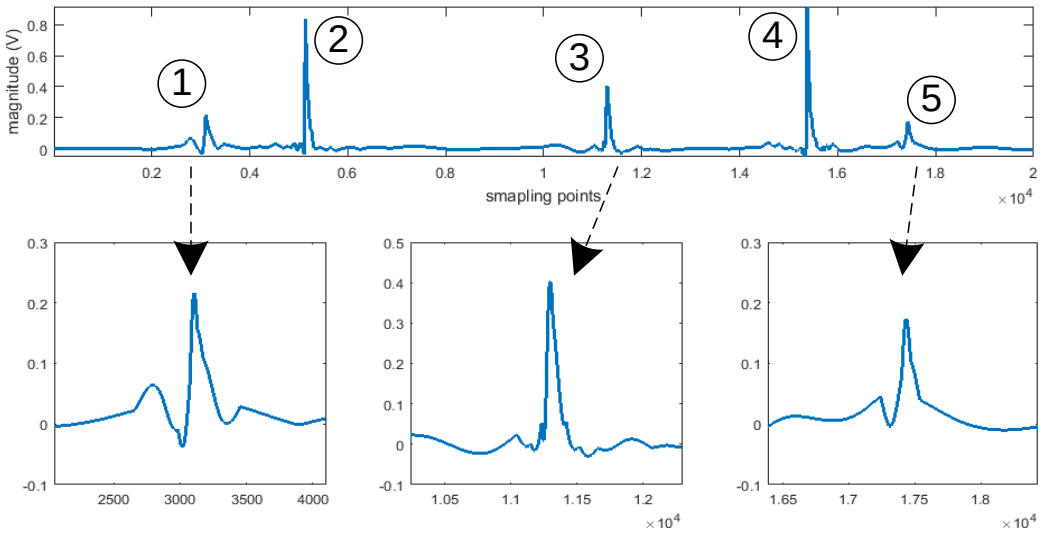

Figure 7-22 The 1st, the 3rd and the 5th PD pulse of denoised *spd1* using EMD-based denoising

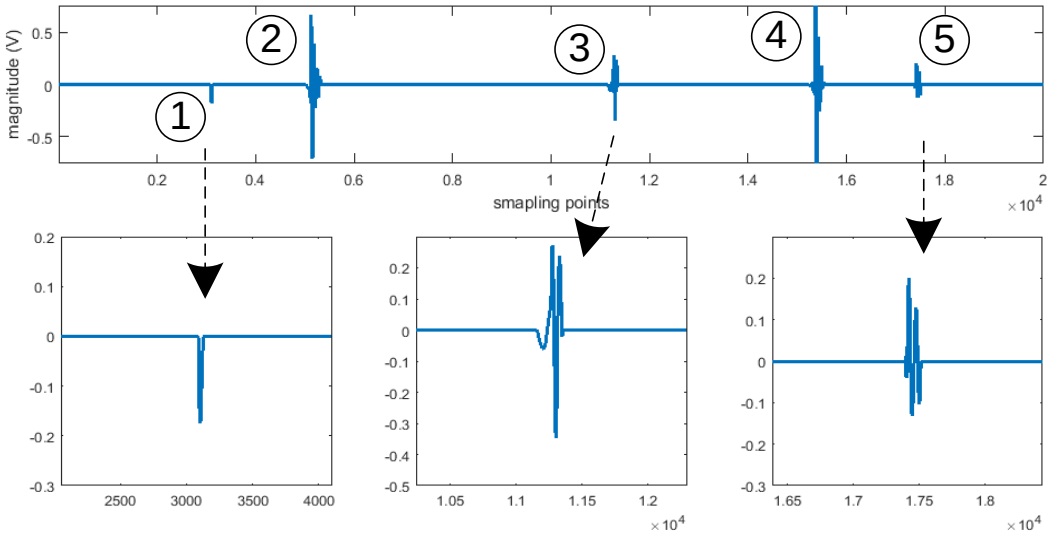

Figure 7-23 The 1st, the 3rd and the 5th PD pulse of denoised *spd2* using EMD-based denoising

Wavelet-based denoising has been demonstrated to be tolerant to DSI when its level in the

original PD signal is not very high. For EMD-based technique, it is also necessary to extend its application in PD signals with both white noise and DSI. EMD functions as a dyadic filter for white noise, and thus, lead to the power spectra of IMFs exhibiting self-similar characteristics. As such, the threshold for each IMF can be estimated through (7-3) and (7-9). When DSI is involved, (7-3) and (7-9) is not suitable for the threshold estimation as the IMFs of DSI cannot have that appealing feature. As a result, the residual of DSI in the denoised PD signals could be very noticeable. Figure 7-24 delineates the denoising results of PD signals corrupted by both white noise and DSI using EMD-based denoising. In both noisy s_1 and s_2, DSI is added with magnitude of 0.1, and white noise has a mean value of zero and standard deviation of 0.3. Figure 7-24 clearly shows that the main PD pulse have been distorted due to the remained DSI, together with small residuals of DSI in the denoised PD signals. This becomes more and more serious with the increase of DSI level in the noisy PD signals. Figure 7-25 illustrated the EMD-based denoising result of noisy s_1 and s_2, where the magnitude of DSI increases from 0.1 to 0.3.

The denoising results shown in Figure 7-24 and 7-25 support that the existing threshold estimation in EMD-based denoising is not suitable for PD signals corrupted by other noise sources, e.g., DSI, except white noise. From a practical point of view, wavelet-based technique, is better than EMD-based technique for PD denoising due to its relatively strong robustness. Low level of DSI in the noisy PD signals can be effectively reduced by wavelet-based denoising, as shown in Figure 6-13.

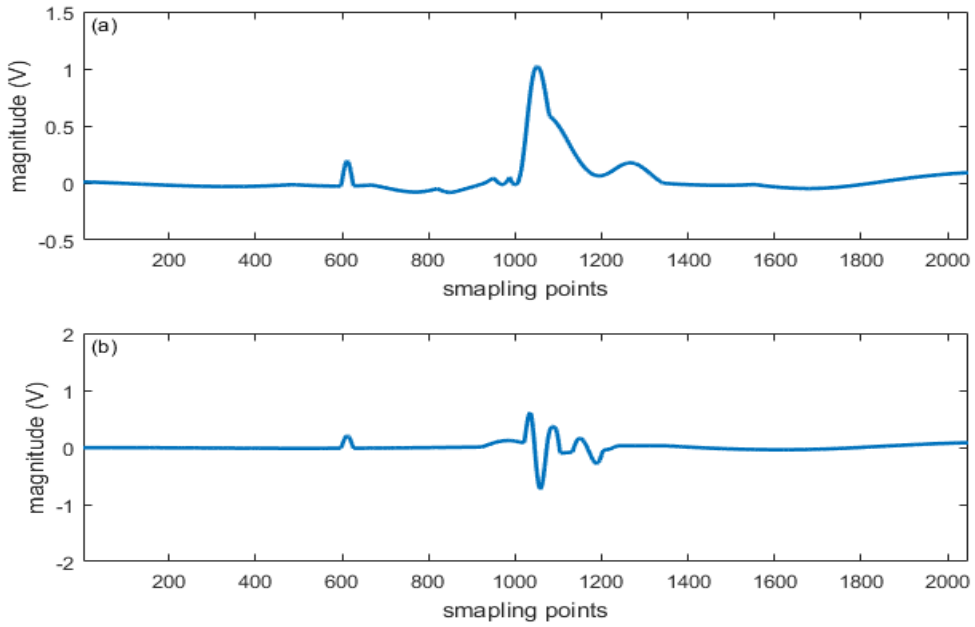

Figure 7-24 EMD-based denoising of (a) noisy s_1, (b) noisy s_2, corrupted by white noise and DSI with low level

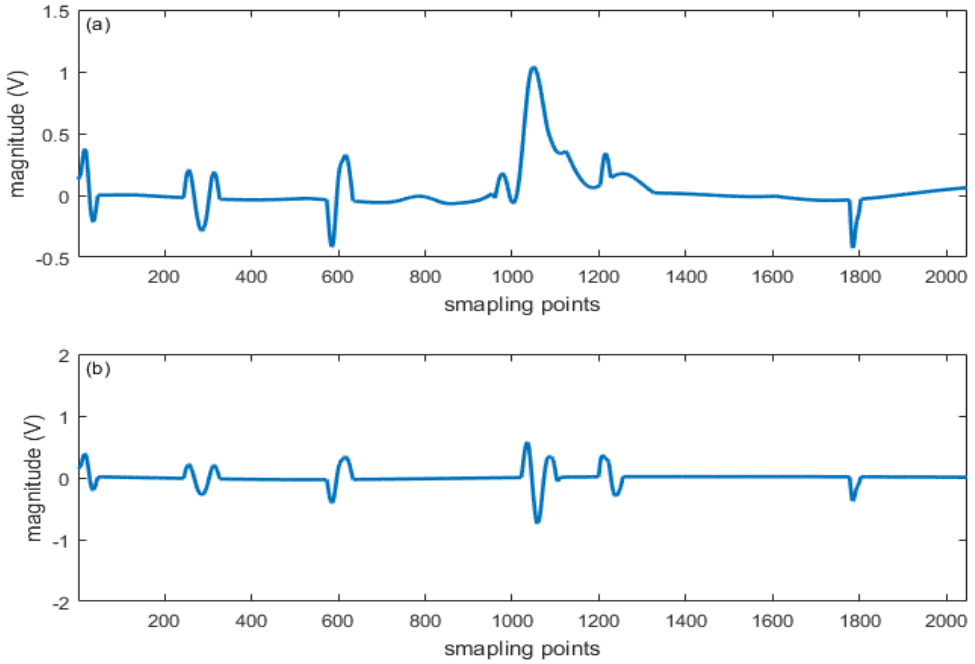

Figure 7-25 EMD-based denoising of (a) noisy s_1, (b) noisy s_2 corrupted by white noise and DSI with relatively high level

Possible threshold estimation that may be applied in EMD-based denoising is analyzed through the difference of working principle between EMD-based expansion and wavelet transform. In wavelet-based denoising, it is possible to detect the noise separately through

certain special layouts of power apparatus, e.g., noise detection from an offline power cable parallel to the cable under monitoring in [12]. Wavelet coefficients of a wavelet-based expansion is dependent on the correlation between the signal and the wavelet used. Given that separate wavelet transform is applied to the detected noise, and to a noisy PD signal corrupted by that noise, the noise can present the same characteristics at each decomposition scale if the same wavelet is used in these two separate transforms. As a result, the noise level at each decomposition scale of the noisy PD signal can be estimated through transformed noise. Also, the decomposition scale is determined by the length of the noise and the noisy PD signal. Normally, these two signals are sampled with equal length, and thus, decomposition scale can be set equal for these two wavelet-based expansions. However, this is not the case in EMD-based denoising. It is ascribed to the adaptivity of EMD in some sense. The number of IMFs of a signal under EMD-based expansion is dependent on the number of its extrema and the stopping criterion of sifting iterations. When EMD is applied to the noise or the PD signal contaminated by that noise, the number of extrema of them is not the same, and thus, the number of IMFs split from EMD is different. The difference in IMF numbers means that the noise at each IMF from these two EMD-based expansions cannot present the same characteristics. Based on this, the threshold estimation through 'noise-only' signal is not applicable to EMD-based denoising. To improve the effectiveness of EMD-based technique in PD denoising, more investigations need to be implemented through the analysis of the noise characteristics based on various noise sources.

7.4 Conclusion

EMD-based denoising in PD detection has been investigated in this chapter for the reason that wavelet-based denoising is often criticized for its non-adaptivity in denoising. Relative mode selection is claimed to be an important element in EMD-based denoising as it can differentiate the IMFs that contain noise only or not. Based on this, a new entropy-based criterion was proposed for relative mode selection. Results show that this new criterion does not outperform the existing ones. It was interesting to find that relative mode selection is actually not as important as claimed during the investigation, however. The output of this investigation

is that noise reduction applied to all IMFs can obtain the best denoising results. Also, it was found that the hard thresholding function in EMD-based denoising can have better performance than soft thresholding function.

Comparisons of performance between wavelet-based and EMD-based denoising were implemented through the denoising results of simulated PD signals. One serious issue of EMD-based denoising was unveiled: PD signals, especially those with small magnitudes, experienced severe distortion after denoising. It is also important to emphasize that the difference of decomposition mechanism between wavelet transform and EMD leads to the variations of denoised PD signals. As aforementioned, IMFs extracted from noisy PD signals are not Gaussian distributed with variance equal to the noise variance, while this is not the case for transformed signals decomposed by wavelet transform. That is, noise can be estimated through the variance of wavelet components irrespective of scale. Although a new threshold estimation was applied in EMD-based denoising, the noise energy model applied in this estimation may cause inappropriate thresholds used. Based on the discussion above, the improved wavelet-based denoising can provide better denoising results.

8 Validation of Proposed Denoising Algorithms with Real PD Data

8.1 Introduction

For a frequently used technique in PD signal denoising, wavelet-based technique has been improved through the analysis of statistical characteristics of noise and PD signals in previous chapters. The new algorithms proposed have been tested through simulated PD signals derived from (2-6) and (2-7). The results present the advances of these new algorithms in PD denoising. However, the performance evaluations were demonstrated from simulated PD signals. As known, simulated PD signals are derived from an optimal detecting environment, which cannot exactly reflect the characteristics of real PD signals in some sense. PD pulse shape, for example, may be distorted to some extent from a real detection system, not like the shapes shown in Figure 2-19. Based on this, the denoising results of real PD signals through the use of those newly proposed algorithms are under uncertainty. The advances of the improved wavelet-based technique in the reduction of noise from simulated PD signals need to be further investigated on real PD signals. This is also applied to EMD-based technique for PD denoising. As mentioned in last chapter, EMD-based technique can have its advantages in PD denoising when white noise is the only source for PD contamination. With the application into real PD signals, the denoising results may support that EMD-based technique is intolerant to other noise sources involved in PD contamination due to the threshold estimation used.

Details regarding the acquisition of real PD signals will be introduced in sub-sections. Real PD signals are mainly originated from experiments conducted in HV laboratory and on-site PD detection in one power substation in the UK. Experimental PD signals are obtained through PD tests performed on a dielectric insulation sample and power cable sample based on IEC60270. On-site PD data was detected from an in-service power cable in one power substation It is necessary to mention that the PD data from power cable sample and on-site measurement are shared by colleagues. The new algorithms proposed for wavelet-based

technique in pervious chapters, together with EMD-based technique, are applied to these real PD signals. The denoising results of real PD signals demonstrate that EMD-based technique is not as effective as expected, while the improved wavelet-based technique is a promising approach for PD denoising. When the improved wavelet-based technique is applied to cable diagnostics, it should have the potential to extend the range of PD detection from cables.

8.2 PD Data Acquisition through Dielectric Insulation Sample Test

The dielectric insulation sample is an epoxy resin disc with two brass plane electrodes fixed to its upper and lower faces, as shown in Figure 8-1. Two artificial voids with diameters 1.25 and 2.05 mm respectively are buried within the epoxy resin disc. The experiment setup for PD test of this epoxy resin disc is based on IEC60270 and delineated in Figure 8-2. The goal of the AC test is to obtain the real PD signal from laboratory experiment for algorithm validation. The test circuit is comprised of:

- AC test set
- A 100kV transformer
- A coupling capacitor, C_k
- The dielectric insulation sample, C_s
- HFCT
- An oscilloscope

In this test, a high frequency current transformer (HFCT) is clipped around the earth wire of the testing sample to collect PD signals. The specifications of the HFCT are listed in Table 8-1. This HFCT is also adopt for the experiment test on EPR cable sample as well as on-site PD detection in the power substation. Note that a pulse injection calibration unit needs to ensure the output from the HFCT could be quantified as an apparent charge on the IEC60270 measurement system. A high-performance digital oscilloscope with up to 10G/s is used to display and record PD signals detected by HFCT. The sampling rate of digital oscilloscope is set to 5G/s in the test. When the voltage applied on the epoxy resin disc reaches 7.2kV, sustaining PD pulses can occur in this testing sample. The PD signal recorded by the oscilloscope are illustrated in Figure 8-3.

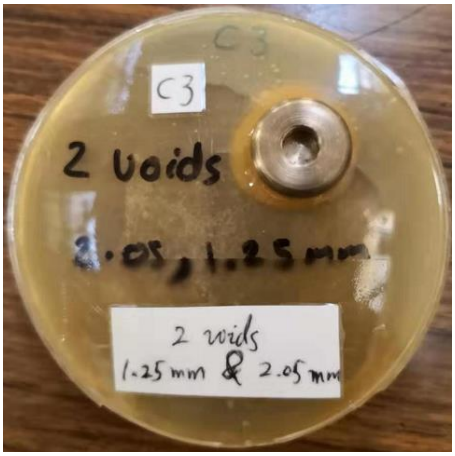

Figure 8-1 Dielectric insulation sample used for PD data acquisition: an epoxy resin disc with two artificial voids of diameters 1.25 and 2.05 mm respectively

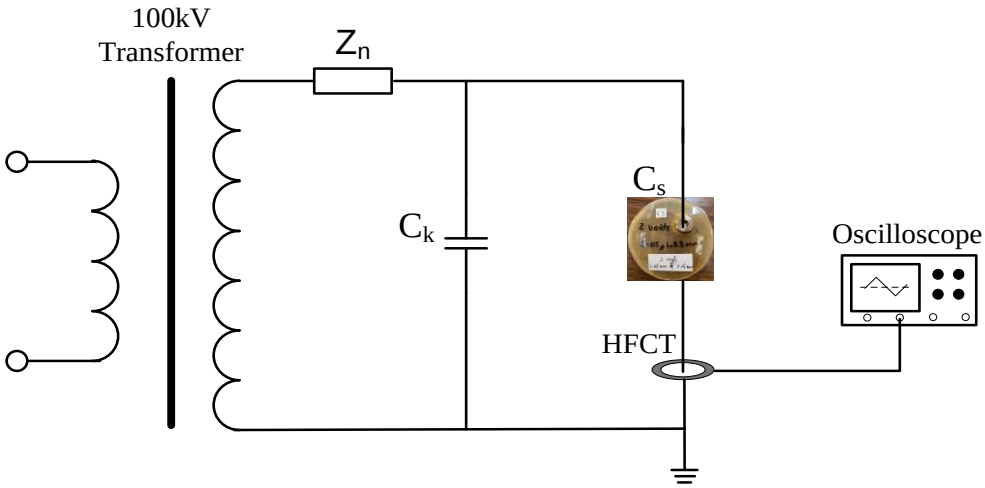

Figure 8-2 AC test circuit for dielectric insulation sample

Table 8-1 Specifications of the HFCT

Parameters	**HFCT**
Sensitivity	5 V/A
-3 dB bandwidth	90 kHz – 20 MHz
Internal diameter	50 mm
External diameter	110 mm
Load resistance	50 Ω
Output conductor	BNC
Manufacturer	IPEC

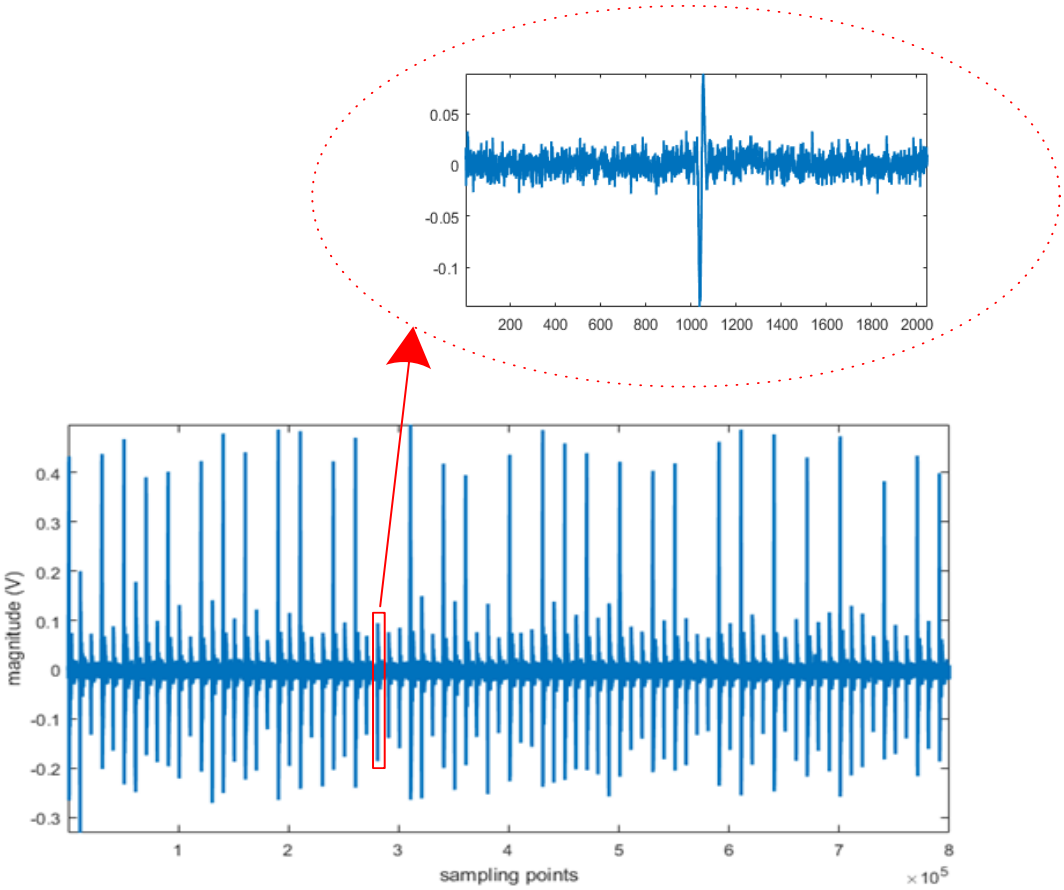

Figure 8-3 PD signal generated within the epoxy resin disc with applied voltage equal to 7.2kV （One single PD pulse is zoomed up to show the waveform）

In the attempt to test those new algorithms for wavelet selection, decomposition scale determination and threshold estimation, one of the PD pulses with 2048 samples, named s_3, is picked up from the recorded PD signal. It is important to note that the recorded PD signal shown in Figure 8-3 is slightly corrupted by ambient noise during experiment. To mitigate the effect of this noise on the denoising results, s_3 has been pre-processed using the method introduced in [88]. The smoothed and normalized s_3 is depicted in Figure 8-4. Difference between simulated and real PD signal can be seen from Figure 4-3 and Figure 8-4.

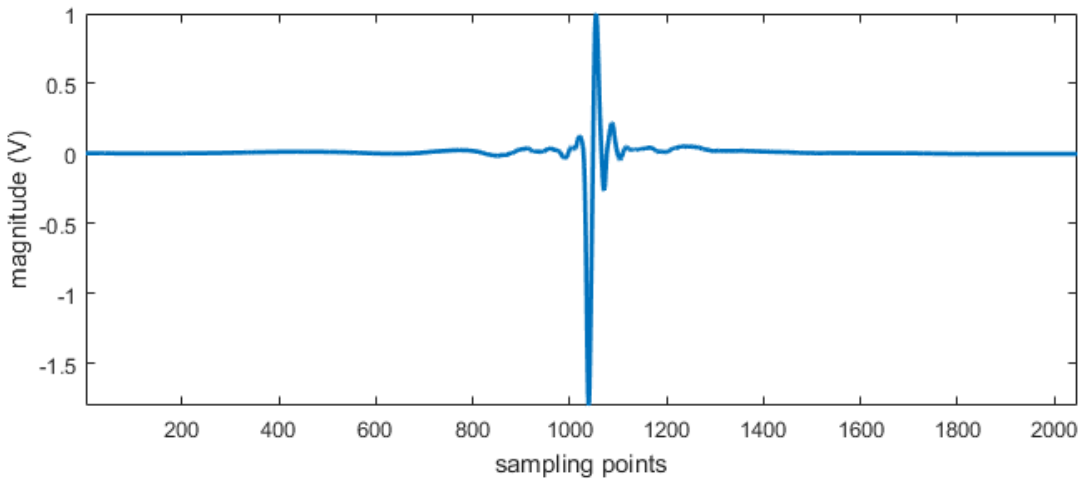

Figure 8-4 The smoothed and normalized s_3

8.2.1 WSS Validation

As did for s_1 and s_2, s_3 (smoothed and normalized) is contaminated by white noise generated from a Gaussian model in MATLAB with various SNRs. Then, different WSSs introduced in Chapter 4 are applied to the noisy s_3. Parameters, such as ME, MSE, and XCORR, are used to evaluate the performance of novel WSS proposed. All the related parameters are illustrated in Figure 8-5. For a real PD signal with single pulse, it can be seen that the novel WSS, i.e., WEBWSS, can present better denoising results than the existing ones, especially when the SNR is low. The improvement in magnitude error and mean square error indicates that the denoised PD signal using WEBWSS can suffer less peak and shape distortion through wavelet-based denoising. In other words, PD pulses with small magnitudes may be picked up and the accuracy for the PD location can be increased based on the novel WSS.

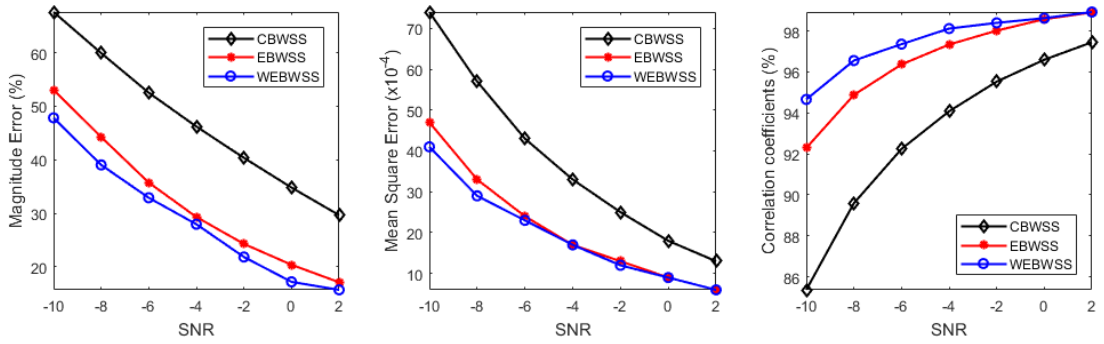

Figure 8-5 ME, MSE, and XOCRR between s_3 and denoised s_3 using various WSSs

The tolerance of the WEBWSS to DSI is also investigated, as did for s_1 and s_2. The issue raised in simulated PD signals is also involved in the denoising of s_3 contaminated by both white noise and DSI. When the level of DSI in the noisy s_3 is low, the residual of the DSI cannot be found in the denoised s_3. Instead, the residual of the DSI is clearly presented in the denoised s_3 with the increasing level of DSI. This can be seen from Figure 8-6, where the magnitude of the DSI increases from 0.1 to 0.3. The same conclusion drawn for the denoising of PD signals corrupted by white noise and DSI is the novel WSS can effectively reduce both white noise and DSI from noisy PD signals when the level of DSI is low. However, this effectiveness is gradually decreased with the increase of DSI level in noisy PD signals.

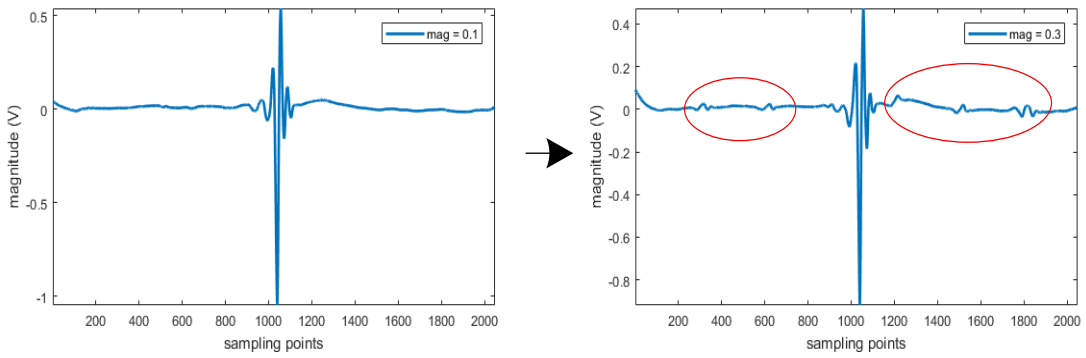

Figure 8-6 Denoising results of noisy s_3 with different DSI level: magnitude = 0.1 (left), magnitude = 0.3 (right)

8.2.2 Decomposition Scale Validation

For the validation of the sparsity-based method for decomposition scale selection, s_3 corrupted by white noise with SNR = -5 is used as an example for performance evaluation. Based on (5-3) and (5-7), the SCR of transformed noisy s_3 is shown in Figure 8-7. In terms of the selection criterion shown in Figure 5-3, 4 is the appropriate decomposition scale for wavelet-based expansion of noisy s_3. The detail coefficient sequences of a full-scale expansion of noisy s_3 is illustrated in Figure 8-8. On the top right corner of each diagram is the linked SCR of that scale. With a closer observation of Figure 8-8, the randomness or degree of disorder of the detail coefficient sequence at scale 5 is the minimum. However, the SCR of scale 5 is equal to 0.1, which means this detail coefficient sequence is comprised of signal component only with high possibility. As a result, 4-scale decomposition is a reasonable choice. To further explicate the new algorithm is a promising approach for decomposition scale determination, noisy s_3 has been denoised through wavelet expansion from 1 scale to full scale (11 scale in this case). The SNRs of these denoised s_3 are depicted in Figure 8-9. Obviously, 4-scale wavelet-based denoising of noisy s_3 can have the maximum SNR after denoising. In other words, the decomposition scale selected by this new algorithm for a real PD signal with single pulse can maximize the noise reduction from the PD signal. The investigation is also implemented on noisy s_3 corrupted by various levels of white noise. The decomposition scale selected by the new algorithm and the scale that has the best SNR after denoising are listed in Table 8-2. Difference between DS and Best SNR Scale can be seen when the SNR is equal to -3 and 3. It is necessary to mention that the scale selected by the sparsity-based method for these two SNRs are the scales close to the best one for wavelet-based denoising of noisy s_3, i.e., the SNRs of denoised s_3 in these two conditions are just lower than the best SNR scale listed in the table. Based on this, good denoising results can be obtained through the sparsity-based method for decomposition scale selection.

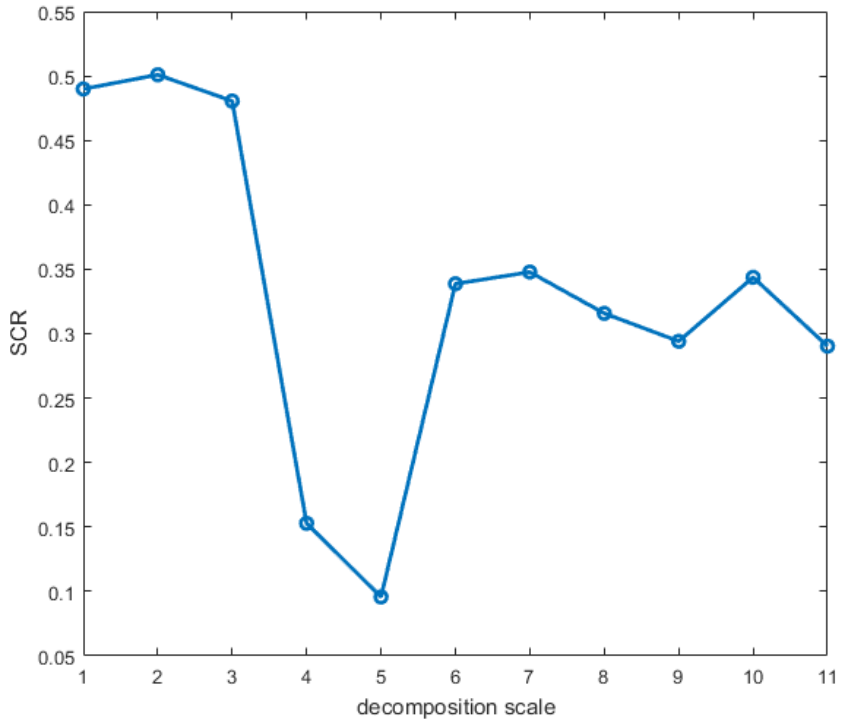

Figure 8-7 The SCR distribution of a full-scale transformed noisy s_3
(o represents the SCR value at each decomposition scale)

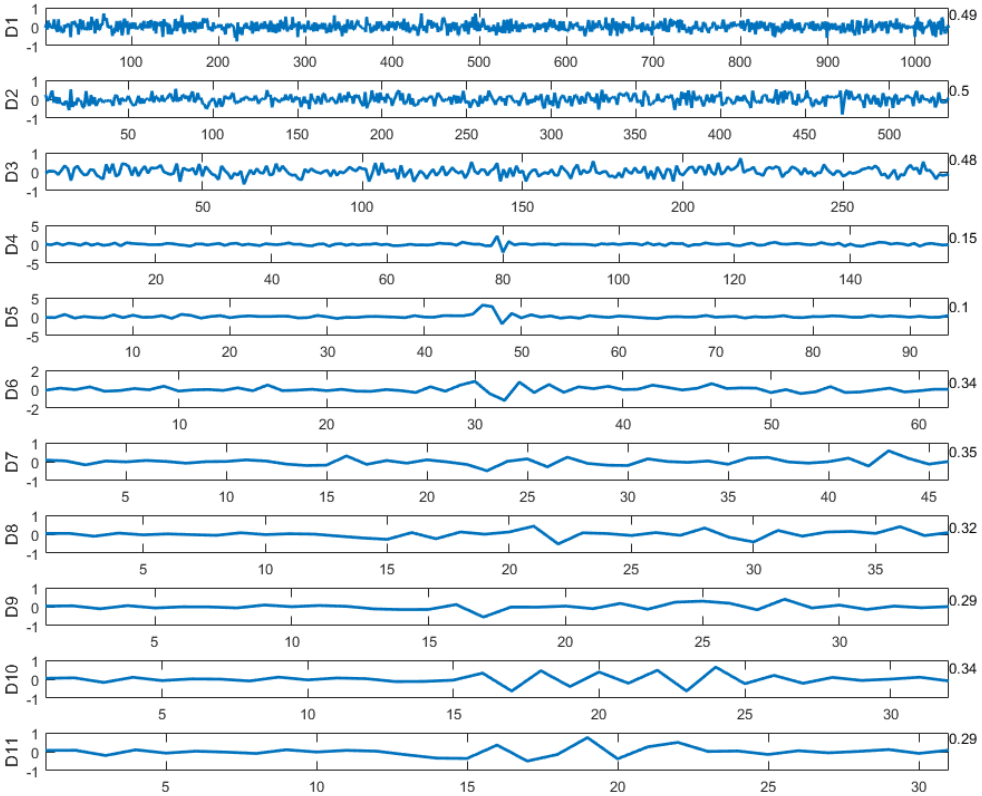

Figure 8-8 Detail coefficient sequence of a full-scale transformed noisy s_3

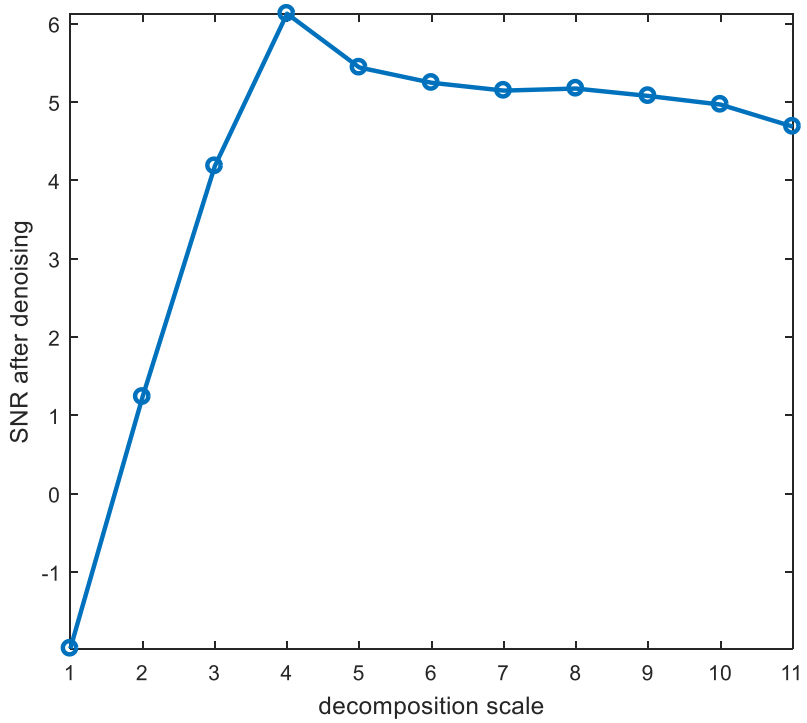

Figure 8-9 The SNRs of denoised s_3 through 1 scale to full-scale wavelet-based denoising (o represents the SNR value of denoised PD signal after each scale – dependent denoising)

Table 8-2 Decomposition scale selected by sparsity-based method for PD signals with various SNRs

ns_3	SNR			
	-3	**-1**	**1**	**3**
DS	4	4	4	3
Best SNR	5	4	4	4

As aforementioned in Chapter 5, a wavelet length method has been adopted for decomposition scale decomposition in wavelet-based PD denoising. As such, the comparison between wavelet length method (CBWSS-WL) and the sparsity-based method (WEBWSS-SP) has been implemented on noisy s_3 (SNR = -5). For noisy s_3, a fixed wavelet, 'db16', is selected by CBWSS for wavelet-based expansion. The decomposition scales selected by wavelet length method and sparsity-based method for noisy s_3 are 6 and 4 respectively. Denoising results are illustrated in Figure 8-10, which demonstrates that WEBWSS-SP outperforms CBWSS-WL in noise reduction for noisy s_3. The related distortions of denoised s_3 are evaluated through the parameters, e.g., ME and MSE, and listed in Table 8-3. The figures in the table clearly show that WEBWSS-SP can remarkably improve the denoising

results of noisy s_3.

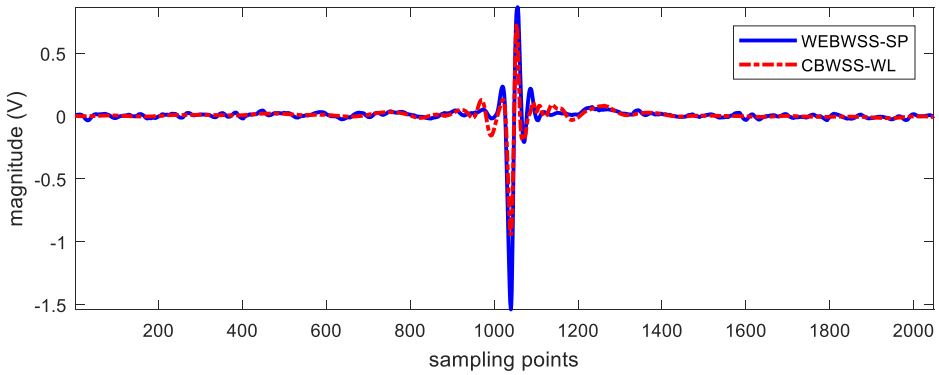

Figure 8-10 Denoising results of noisy s_3 using WEBWSS-SP and CBWSS-WL

Table 8-3 Parameters used for performance test of WEBWSS-SP and CBWSS-WL

s_3	Parameters		
	ME	**MSE**	**XCORR**
WEBWSS-SP	0.134	0.0007	0.978
CBWSS-WL	0.272	0.0035	0.934

The analysis of the sparsity-based method for decomposition scale determination through a real PD signal with single pulse demonstrates that the newly proposed algorithm can be applied as an alternative to the existing methods and provide good denoising results in PD denoising.

The combination of WEBWSS and sparsity-based method for scale selection is also applied to noisy s_3 contaminated by both white noise and DSI. The aim of this investigation is to test the tolerance of this improved method to DSI, and thus, DSI is set with low level to high level in noisy s_3. Similar results are obtained as the tolerance of WEBWSS to DSI. Denoising results using WEBWSS-SP are illustrated in Figure 8-11. It can be seen that the new method cannot effectively remove the DSI from noisy s_3 with the increase of DSI level. Figure 8-11 shows that when the magnitude of DSI is equal to or over 0.3, residuals of DSI can be directly observed in the denoised s_3.

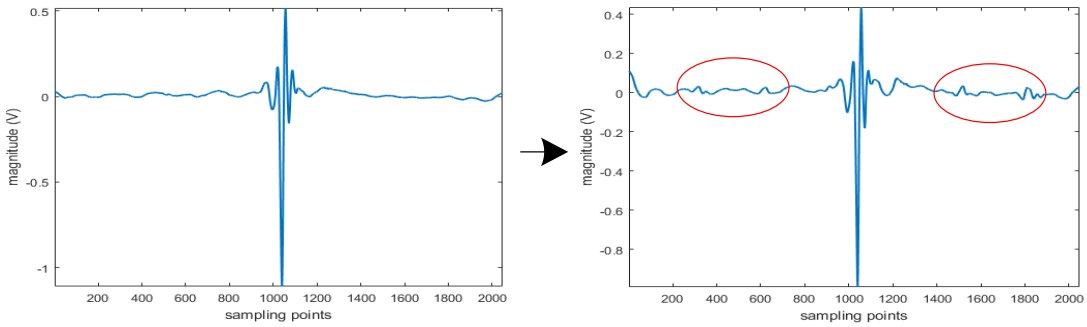

Figure 8-11 Denoising results of noisy s_3 using WEBWSS-SP with different DSI level: magnitude = 0.1 (left), magnitude = 0.3 (right)

8.2.3 Threshold Estimation Validation

A novel threshold estimation was proposed based on the SCR distribution of detail coefficient sequences of transformed PD signals. For simulated PD signals, it can improve the performance of wavelet-base denoising. As did for WSS and sparsity-based method for decomposition scale selection, the PD signal from experiment is also applied to test its performance for real PD signal. s_3 is corrupted by white noise with various levels. Evaluation of the denoising results using this novel threshold is illustrated in Figure 8-12. The parameters indicate that the use of new threshold can present better denoising results than the existing universal threshold for noisy s_3. It is highly desirable that less distortion of denoised s_3 can be obtained using new threshold.

White noise in noisy s_3 can be effectively reduced through the application of new threshold estimation. To investigate its application in complex noise environment, DSI is added into s_3, together with white noise. The denoising results of noisy s_3 with different DSI levels are illustrated in Figure 8-13. It has the same problem in the reduction of DSI, i.e., the effectiveness of the improved wavelet-based technique for PD denoising is inverse proportional to the level of DSI in the noisy s_3. As compared to the result in Figure 8-11, it can be seen that peak distortion of the denoised s_3 is reduced by the use of new threshold.

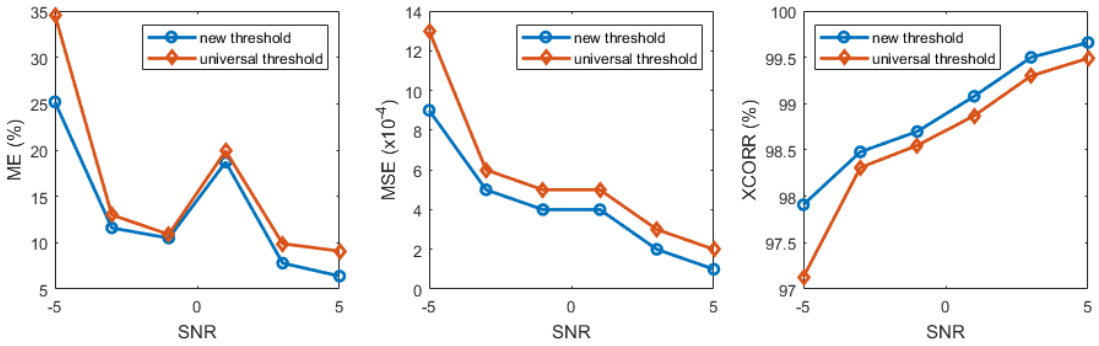

Figure 8-12 ME, MSE, and XCORR for performance test of soft threshold and new threshold

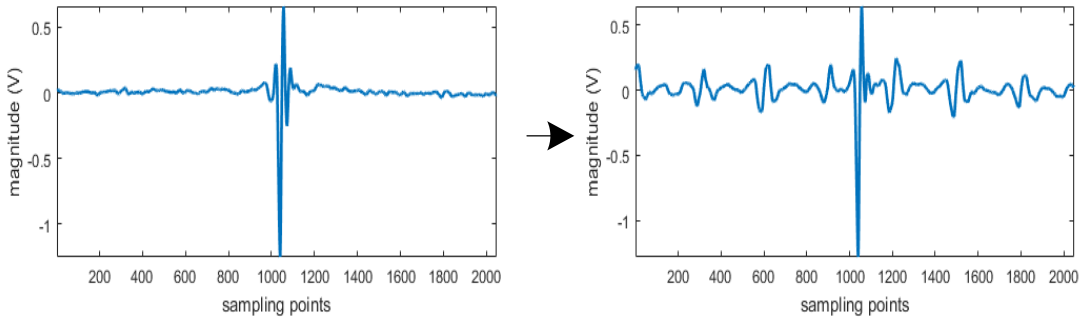

Figure 8-13 Denoising results of noisy s3 using new threshold with different DSI level: magnitude = 0.1 (left), magnitude = 0.3 (right)

In Chapter 6, a new H-S thresholding function was also introduced for wavelet-based PD denoising. s_3 contaminated by white noise with SNR = -5 is used as an example to embody the advance of this H-S thresholding function. The denoising result of noisy s_3 in Figure 8-14 demonstrates that the H-S thresholding function can provide better denoising result than traditional soft-threshold function. Parameters shown in Table 8-4 further supports the that the H-S thresholding function can remarkably improve the effectiveness of wavelet-based technique in PD denoising.

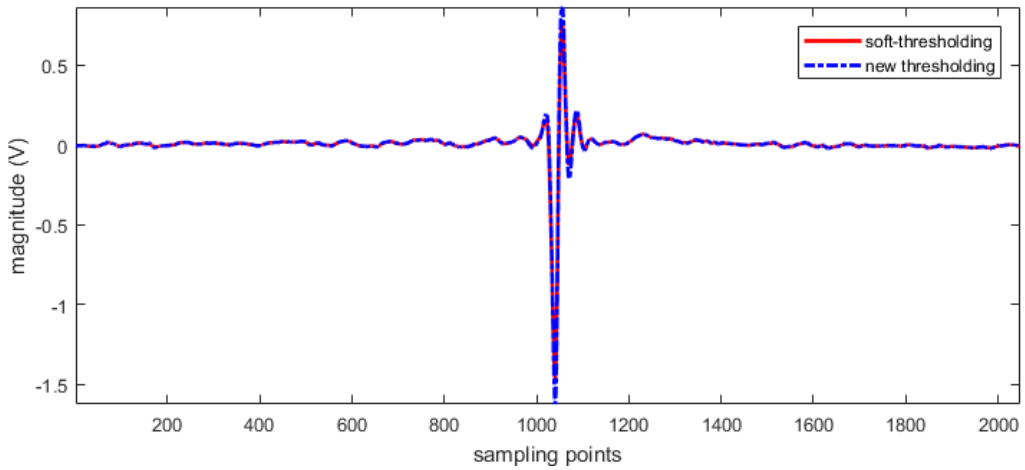

Figure 8-14 Denoising results of noisy s_3 with soft-thresholding function and H-S thresholding function

Table 8-4 Parameters used to evaluate the H-S thresholding function in wavelet-based PD denoising

s_3	Parameters			
	ME	**MSE**	**XCORR**	**SNR**
Soft-thresholding	0.252	0.0009	0.979	12.79
H-S thresholding	0.136	0.0004	0.988	15.89

In this subsection, a real PD signal with single pulse was extracted from a real PD signal from laboratory experiment. Note that the noise that contaminates the PD signal is simulated in MABLAB. The aim of this is to further demonstrate the newly proposed algorithms in wavelet-based technique are still effective for noise reduction from a real PD signal with single pulse. The denoising results through 3 aspects in the improved wavelet technique show that better denoising results can be obtained as compared to the exiting techniques in the area of PD detection.

Based on the discussion in this subsection and Chapter 4-6, it is necessary to detail the methods used in denoising of real PD signals (with multiple pulses) in the following subsections. For comparison, three methods, named M1, M2, and M3 respectively, vary from different combination of wavelet selection scheme, decomposition scale determination, threshold estimation and thresholding functions, and are described as follows:

- **M1**: CBWSS + Wavelet length method + Universal + Soft-thresholding function
- **M2**: EBWSS + Sparsity-based method + Universal + Soft-thresholding function

- **M3**: WEBWSS + Sparsity-based method + SCR-based threshold + H-S thresholding function

It is obvious that each method is comprised of techniques for wavelet selection, decomposition scale determination, threshold estimation and thresholding functions. The method M3 is the combination of all the newly proposed algorithms introduced in previous chapters.

8.3 PD Data Acquisition through EPR-insulated Power Cable Test

To develop the improved wavelet-based denoising for practical use, real PD signals were generated through an artificial defect of a $7mm \times 7mm$ breach in the outer conductor created in a 1.5m 11 kV EPR-insulated cable sample [190]. PD signals are detected using a HFCT, the specifications of which are listed in Table 8-1. Details regarding the experiment setups are depicted in Figure 8-15 [190].

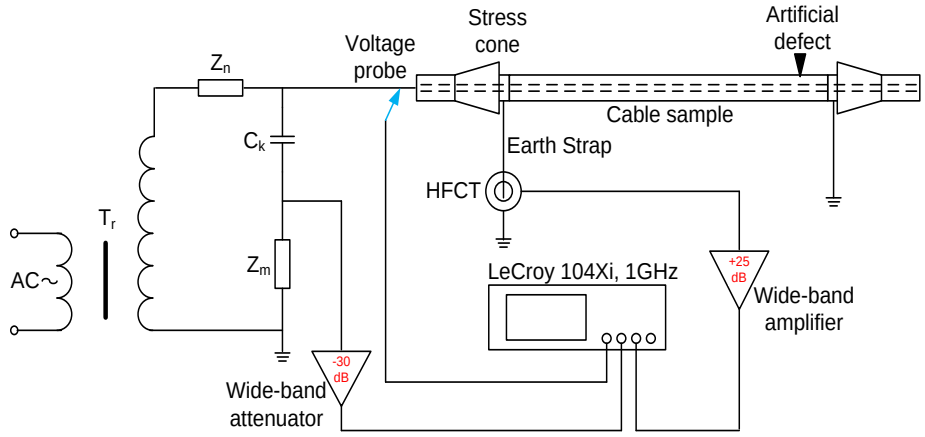

Figure 8-15 PD testing of a defective 11 kV EPR cable. HFCT was used to collect PD pulses (Ck and Zm represent the coupling capacitor and measuring impedance respectively)

Experiments are performed at various voltage levels. The PD pulses measured at 9kV are used as the real PD signal to demonstrate the improved wavelet-based denoising in this thesis. Figure 8-16 delineates the detected noisy PD signal from the EPR-insulated cable in the HV laboratory. The comparisons of the denoising results using M1, M2, and M3 introduced above

are illustrated in Figure 8-17.

As shown in Figure 8-17, the differences of denoising results among these 3 methods have been highlighted with red ovals. It is obvious that many more PD pulses with small magnitudes can be successfully extracted from noise using M3, the method comprised of novel algorithms proposed in each step of wavelet-based denoising. Several PD pulses from denoised EPR PD signal are taken as examples to show the advances of M3, which can be seen from Figure 8-18. In the meantime, it can be seen that the peak distortion of PD pulses extracted by M3 has been remarkably reduced as compared to M1 and M2. With a closer observation of the denoising results of M1 and M2, it is in agreement with the claim in [88] that energy-based wavelet selection scheme is better than correlation-based wavelet selection scheme in PD denoising.

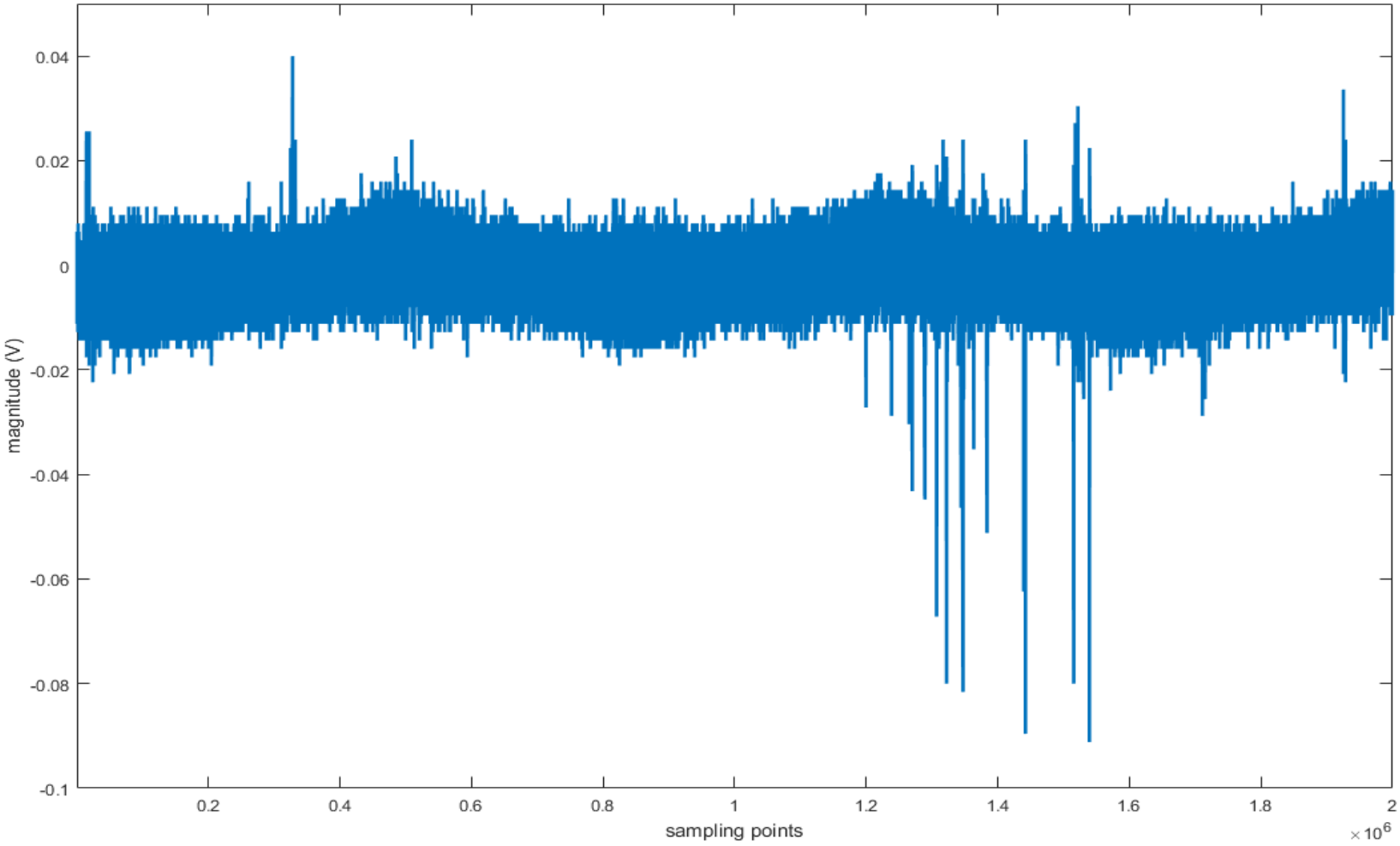

Figure 8-16 A noisy PD signal detected from a 11kV EPR-insulated power cable with an artificial defect under 9kV

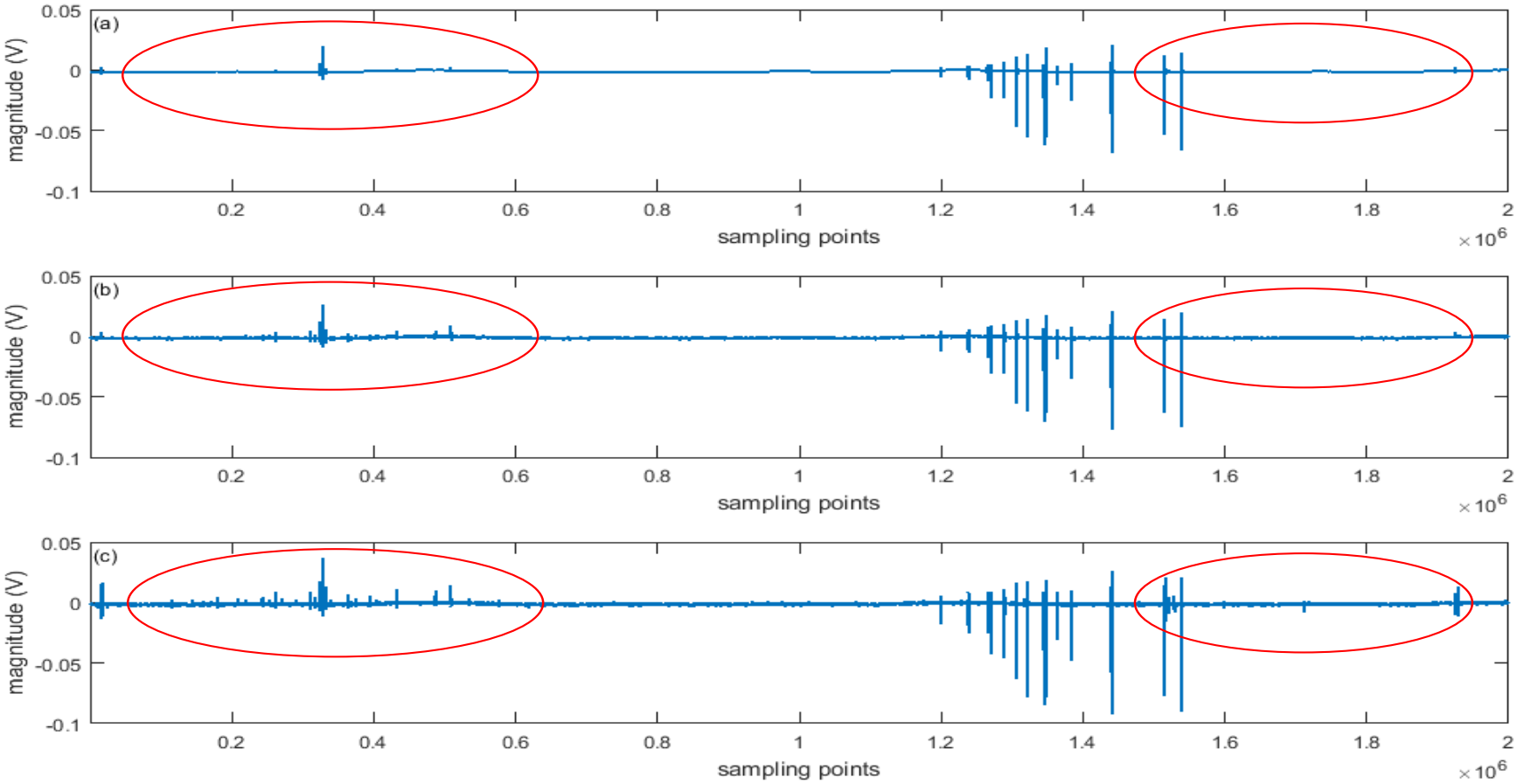

Figure 8-17 Denoising results of a noisy PD signal from an EPR cable using (a) M1, (b) M2, and (c) M3

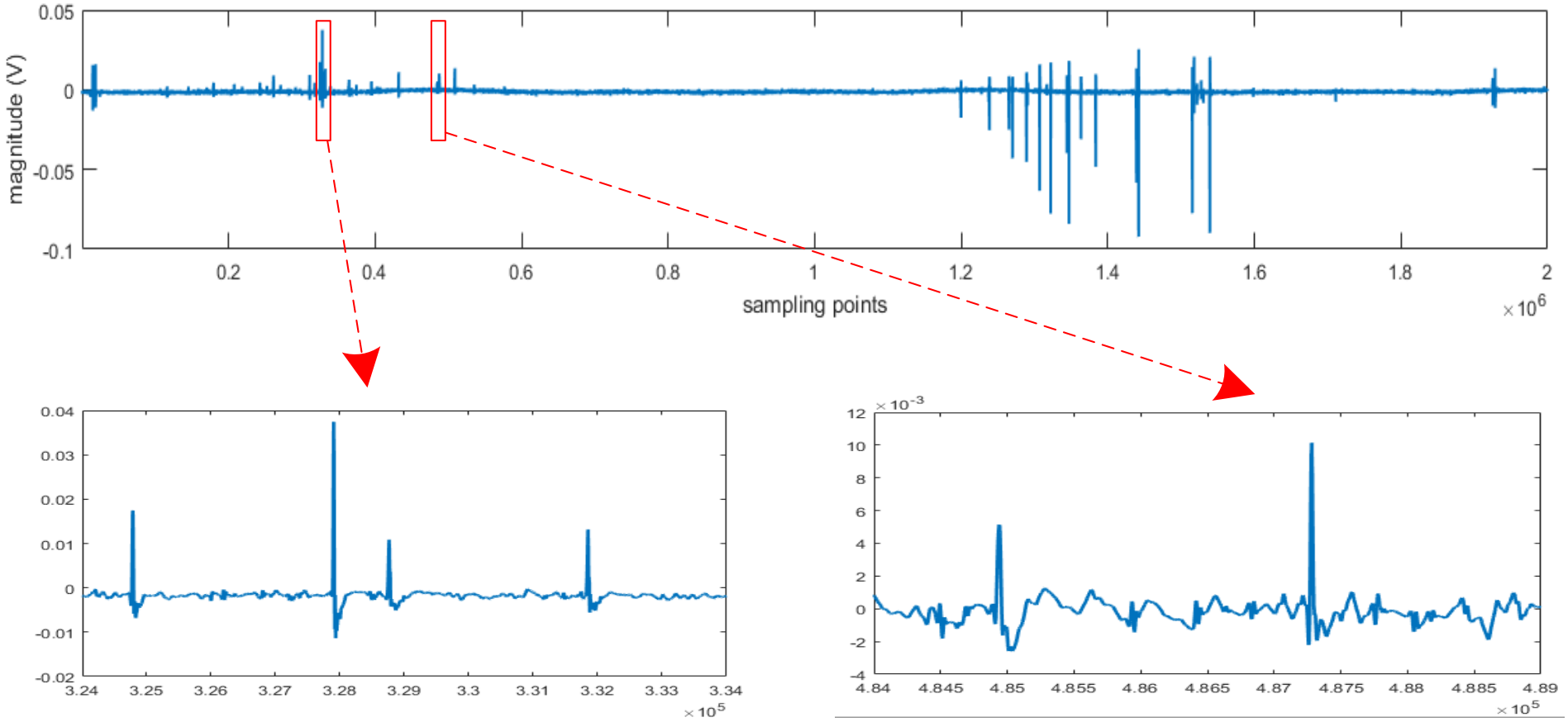

Figure 8-18 Highlight of some pulses from denoised EPR PD signal using M3

8.4 PD Data Acquisition through On-site Measurement

Real onsite PD data was collected from one power substation in the UK with a sample rate of 100MS/s. Figure 8-19 delineates the original on-site PD signal. The denoising results of this on-site measured PD signal using the methods, i.e., M1, M2, and M3, are depicted in Figure 8-20 for comparison. It can be seen that not only the PD pulses with magnitudes higher than the noise level has been extracted, but those ones with small magnitudes buried in the noise has been effectively extracted through the application of the improved wavelet-based technique (M3). The number of the PD pulses with small magnitude that have been extracted by M1 and M2, as shown in Figure 8-20 (a) and (b), is less than that by M3. The difference of the number of small-magnitude PD extraction among these three methods has been highlighted in red ovals in Figure 8-20. Also, the peak distortion of denoised PD pulses using M3 is substantially reduced as compared to M1 and M2. In Figure 8-21, two single PD pulse are selected from denoised PD signal using M3. It clearly shows that the PD pulse with small magnitude can be effectively extracted through M3 with lower distortion. The remarkable improvements of denoised PD signals by the use of newly proposed algorithms provide further support that the improved wavelet-based technique is more advantageous than the existing ones for denoising of PD detection of electrical apparatus in practice.

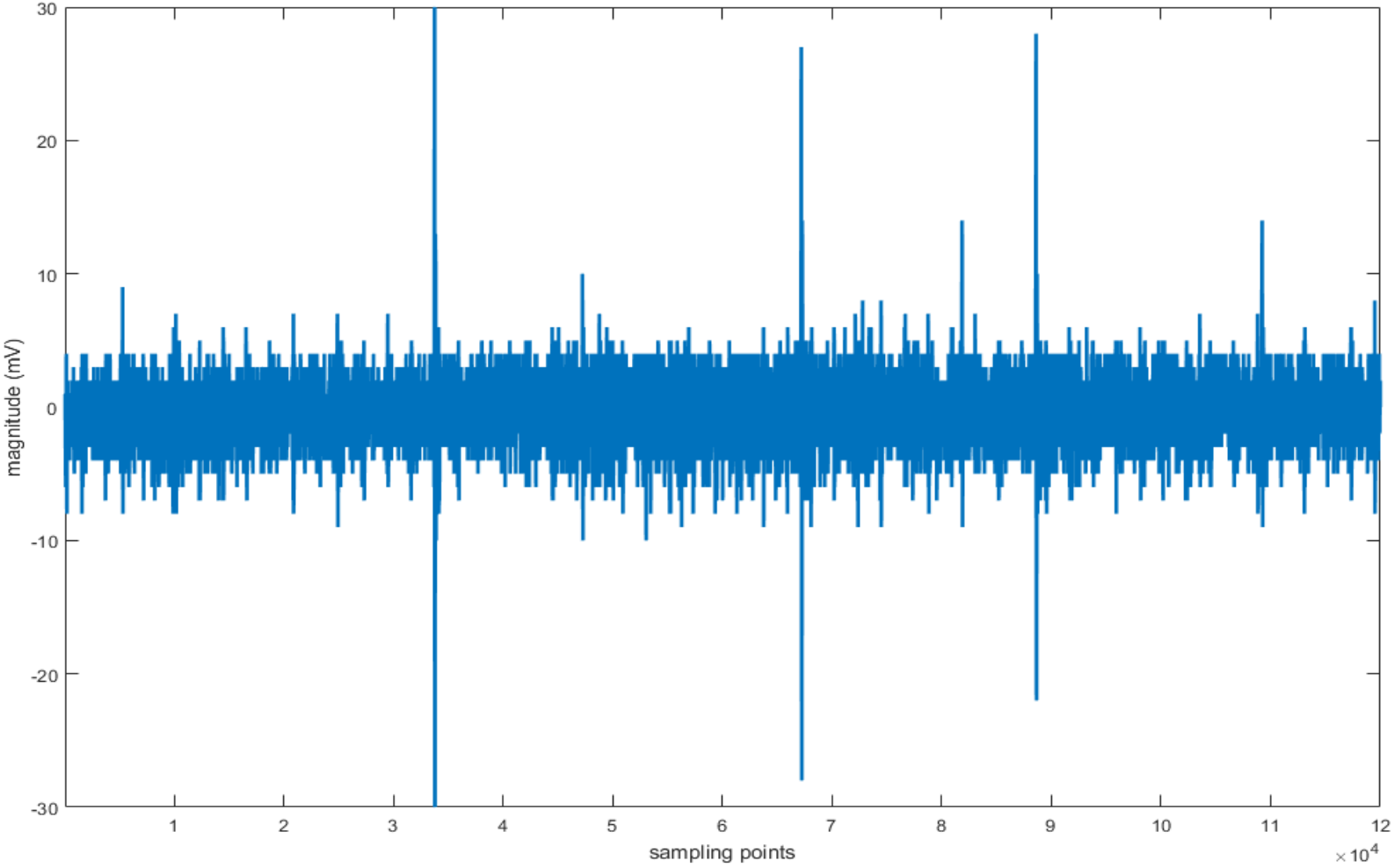

Figure 8-19 A noisy PD signal from on-site measurement in a power substation

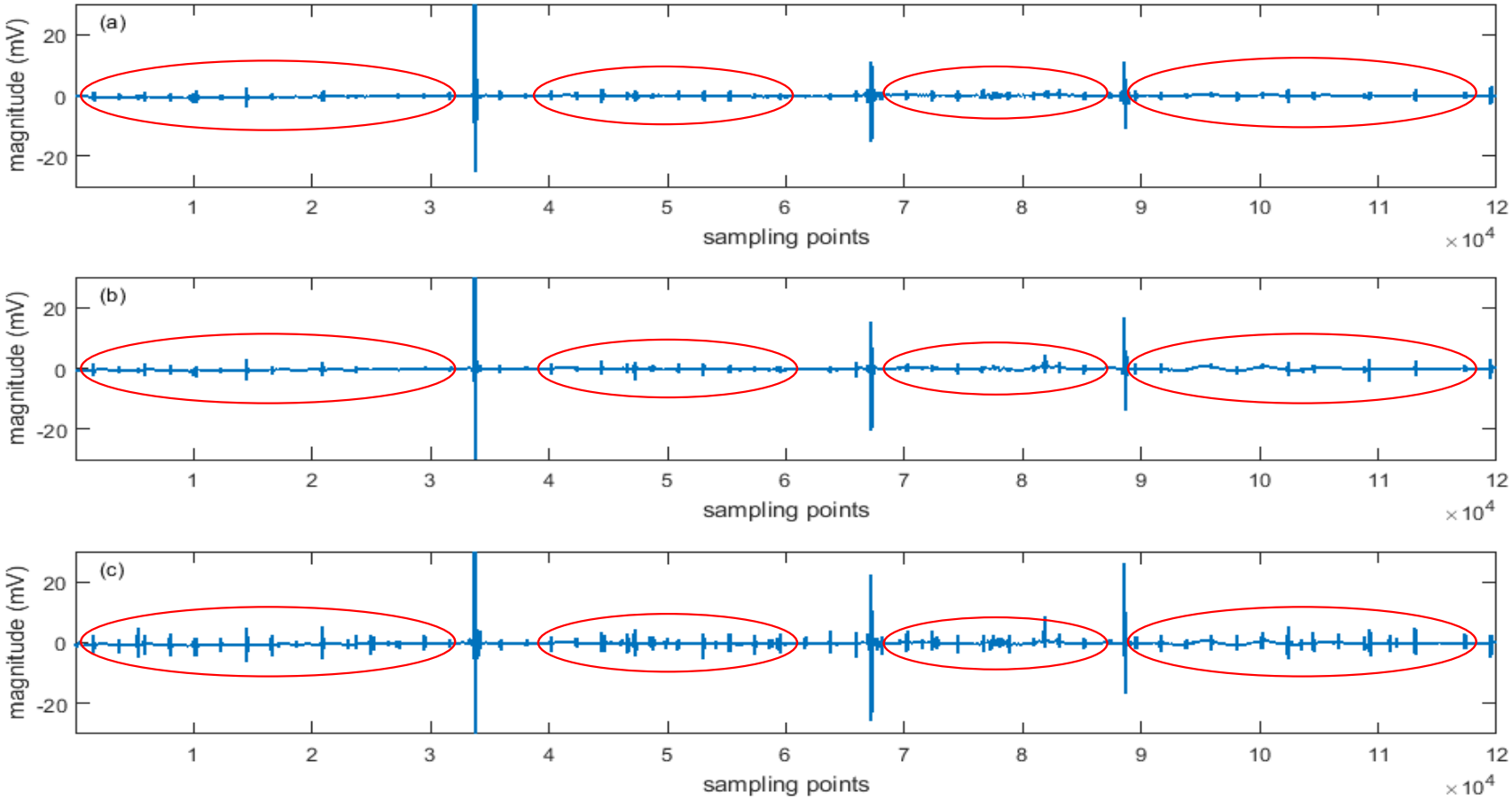

Figure 8-20 Denoising results of a noisy PD signal from on-site measurement using (a) M1, (b) M2, and (c) M3

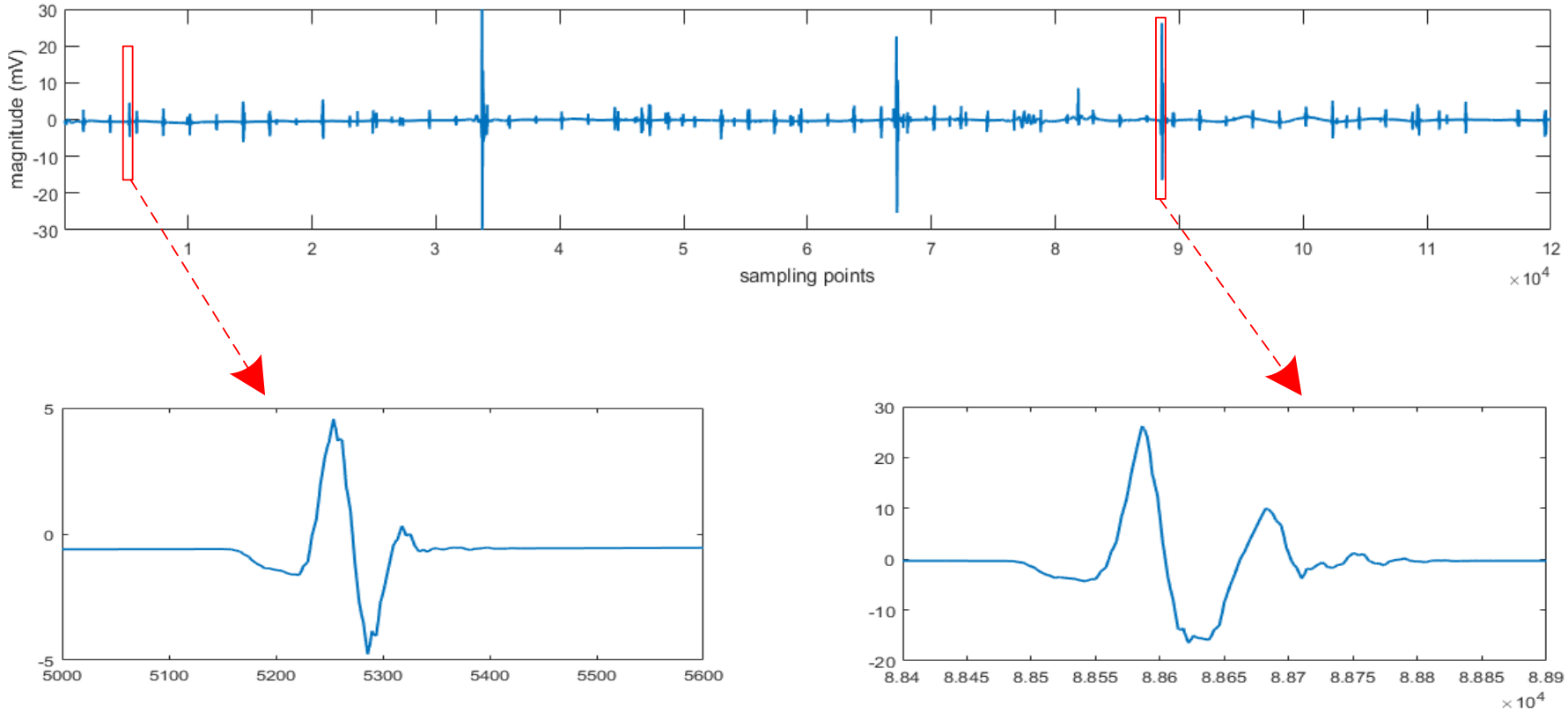

Figure 8-21 Highlight of some pulses from denoised on-site PD signal using M3

8.5 Denoising Real PD Data with EMD-based Technique

The real PD data used in last three sections to test the performance of improved wavelet-based technique is also applied to investigate the effectiveness of EMD-based technique for PD denoising. For s_3, the associated noise is still simulated with a Gaussian model in MATLAB. As such, noisy s_3 is a PD signal contaminated by white noise with SNR = -5. The denoising results of three real PD signals are illustrated in Figure 8-22. It can be seen that EMD-based denoising can extract PD pulses from PD data obtained through HV laboratory, as shown in Figure 8-22 (a) and (b). This may be due to the HV laboratory is highly screened, and thus, noise, such as DSI, has less or no effect on the detected PD signals. In other words, the detected PD signals in HV laboratory may be contaminated only by white noise. However, EMD-based denoising is ineffective in the extraction of PD pulses with small magnitudes from on-site measured PD data (see Figure 8-22 (c)).

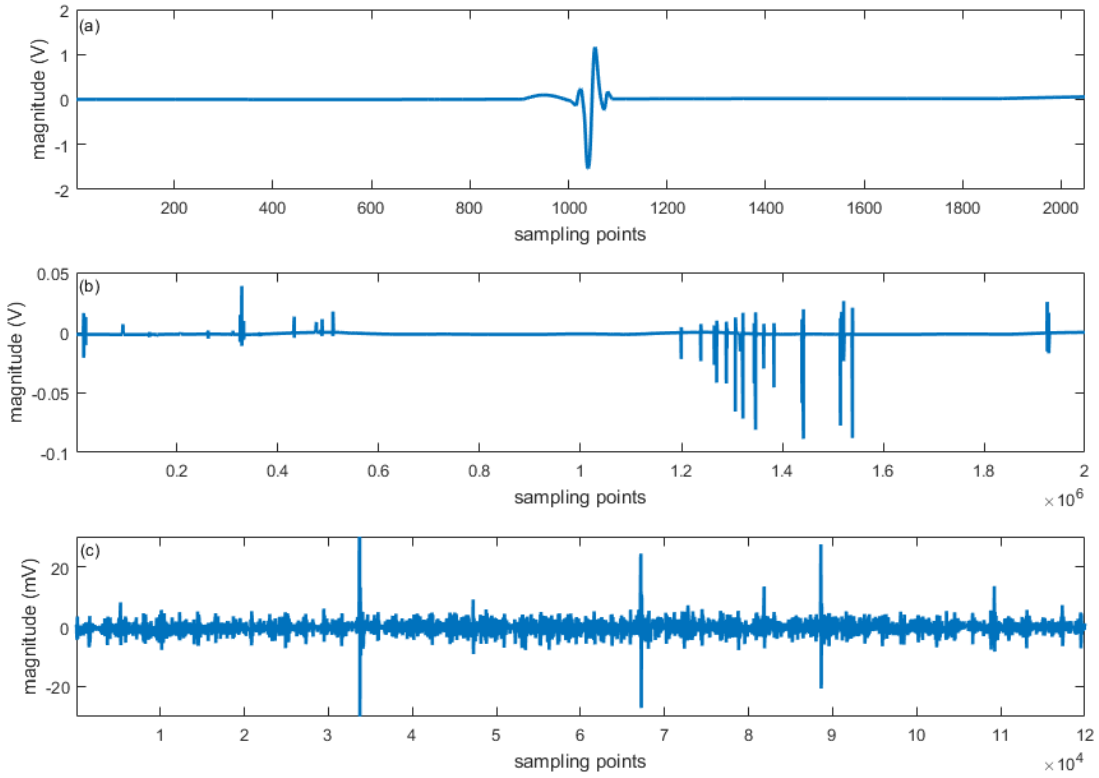

Figure 8-22 EMD-based denoising results of PD signal from (a) dilectric insulation sample, (b) EPR-insulated power cable, and (c) on-site measurement

Comparisons between EMD-based and wavelet-based techniques for PD denoising are depicted in Figure 8-23, in which the PD signal obtained through an EPR-insulated power cable is used as the original signal. As highlighted in red ovals in Figure 8-23(a) and (b), more PD pulses with small magnitudes that are successfully extracted by wavelet-based denoising as compared to EMD-based denoising. In this scenario, wavelet-based denoising is more preferable than EMD-based denoising as the effectiveness in extraction of small-magnitude PDs is highly desirable for the analysis of PD data, e.g., the degree of PD-induced deterioration of insulation system of electrical apparatus can be reflected by the number of PD events.

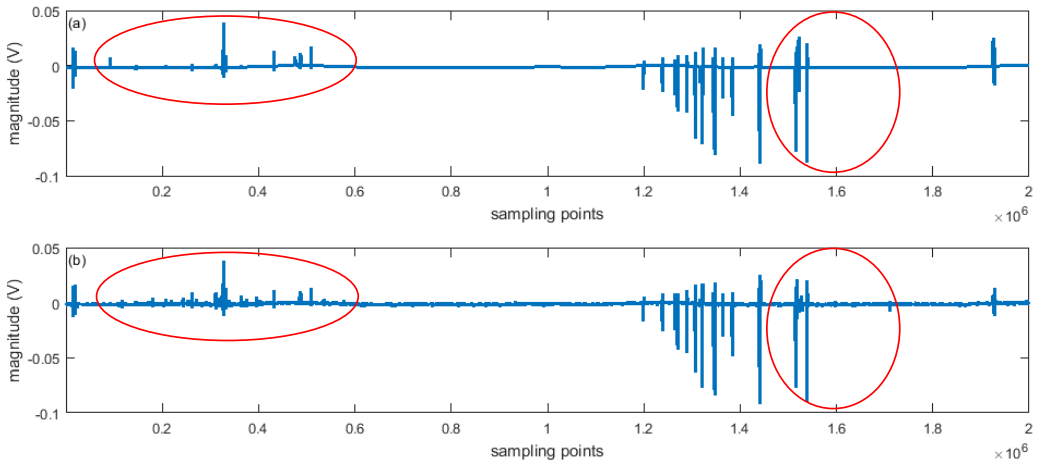

Figure 8-23 Denoising results of PD signal from EPR-insulated power cable using (a) EMD-based denoising, (b) wavelet-based denoising

In addition, Figure 8-24 delineates a closer observation of some pulses of the denoised PD signal from an EPR-insulated power cable using EMD-based technique. They are zoomed up with numbering ①-④, as shown in Figure 8-24. Number ④ is a real PD pulse generated in the EPR-insulated power cable when the applied voltage increased to 9kV. However, the denoised pulses, numbered as ①-③, are severely distorted under EMD-based denoising, and thus, cannot be recognized if they are real PD pulses in this case. This distortion of pulses with small magnitude has already been encountered in EMD-based technique for simulated PD signals. Based on the denoising results shown in Figure 7-22 and Figure 7-23, the

underlying reason leading to this distortion may be due to the low SNR or the involvement of noise types, e.g., DSI, in detected PD signals.

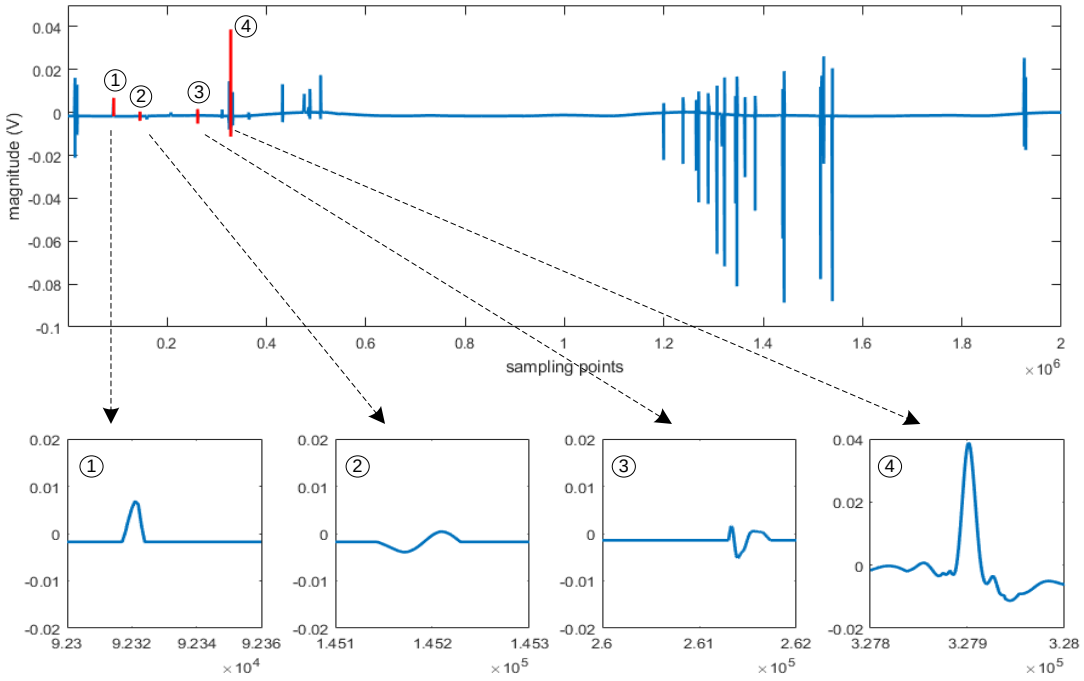

Figure 8-24 EMD-based denoising results of a PD signal from an EPR-insulated power cable in laboratory test

Denoising result of an on-site measured PD signal, as shown in Figure 8-22 (c), demonstrates that EMD-based denoising presents its ineffectiveness in extraction of PD pulses with small magnitudes. In other words, noise in on-site detected PD signals cannot be remarkably reduced through the application of EMD-based technique. The potential reasons for this ineffectiveness could be the underestimated threshold at each IMF mode or the involvement of other noise sources, e.g., DSI, which is equivalent to inappropriate threshold estimation based on (7-3) and (7-9). For underestimated threshold, Figure 8-23 illustrates the difference of $log_2(energy)$ between estimated noise energy and real noise energy of each IMF from an EMD-based expansion of white noise. Numerous simulations have been implemented to verify the relationship between real noise energy and estimated noise energy. This relationship shown in Figure 8-23 indicates that the estimated noise energy in the IMFs with relatively high order is smaller than the real noise energy of those IMFs. The difference increases with the increasing order of the IMF. As a result, the noise threshold estimated based on (7-3) and

(7-9) for high-order IMFs could be lower than real noise level, leading to the remains of noise in the denoised signals. For inappropriate noise threshold, the involvement of DSI, for example, can lead to serious distortion or noise residuals of denoised signals, as shown in Figure 7-24 and Figure 7-25.

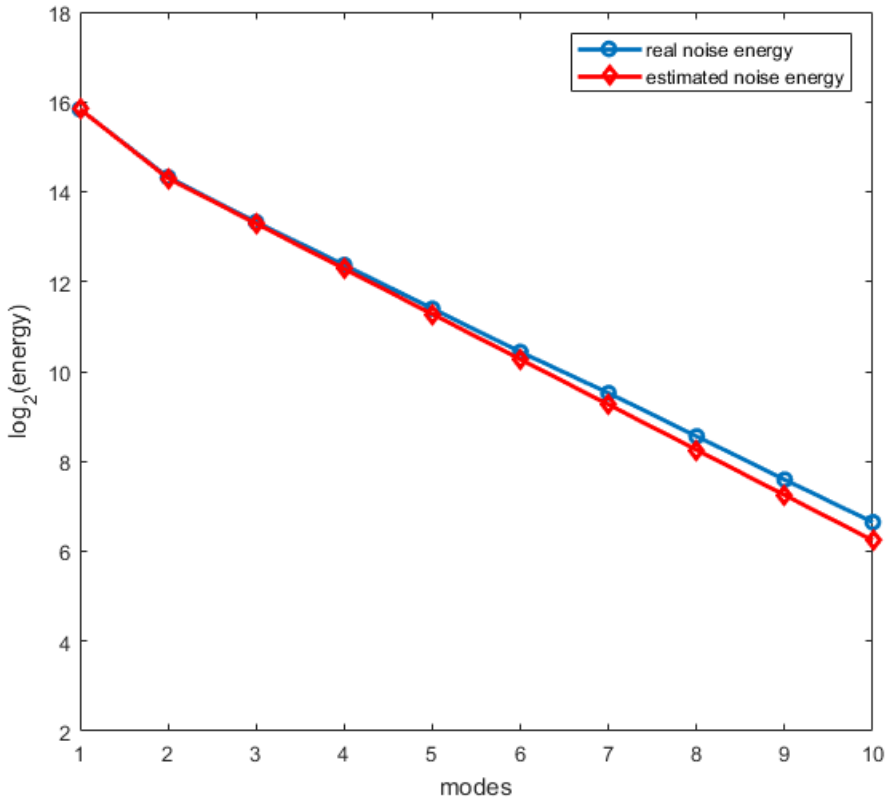

Figure 8-25 The *log₂(energy)* of each IMF from a noise model estimation and real noise energy

8.6 Recommendations on Denoising Techniques for PD Detection

Denoising strategies, i.e., wavelet-based technique and EMD-based technique, have been fully discussed through their applications in both simulated and real PD signals. Denoising results show that wavelet-based technique is more preferable than EMD-based technique as more PD pulses with small magnitudes can be successfully extracted using wavelet-based technique. A brief summary of the comparison between these two strategies are presented as follows:

- **Executing time**: The time for execution of the improved wavelet-based technique for PD denoising is substantially less than that of EMD-based technique under the same environment. For example, the executing time is 11.08s for wavelet-based denoising of the PD signal from the EPR insulated power cable while 41.72s for EMD-based denoising. The longer the signal, the more time is required for EMD-based denoising than wavelet-based denoising.
- **Distortion**: Distortion of denoised PD signals, especially for single PD pulses, using EMD-based technique is much more severe than the improved wavelet-based technique. This has been demonstrated in Figure 7-22, Figure 7-23 and Figure 8-24. The severe distortion of denoised PD signal is not expected as it can negatively affect the following analysis of PD data for, e.g., PD location and recognition.
- **Effectiveness**: More PD pulses with small magnitudes of real PD signals can be extracted through the application of the improved wavelet-based technique than EMD-based technique. An example can be seen from Figure 8-23. Also, more noise is remained in the denoised PD signals by the use of EMD-based technique, as shown in Figure 8-22 (c). Based on this, the effectiveness of EMD-based technique is highly reduced as compared to the improved wavelet-based technique.
- **Adaptivity**: The adaptivity is referred to as the decomposition of a signal using wavelet transform and EMD. Obviously, EMD is adaptive as the basis function used for signal expansion is derived from the data itself, while it is the limit for wavelet transform as wavelet transform is often criticized for its predefined wavelet for signal decomposition. However, the adaptivity of EMD sometimes may arise the issues for EMD-based denoising, as discussed in subsection 7.3.2.
- **Robustness**: The detected PD signal on site can be contaminated by white noise, DSI, pulse-type noise, and most of time are corrupted by the combination of these types of noise. Based on the investigations implemented in previous chapters, the improved wavelet-based technique is more robust than EMD-based technique as the improved wavelet-based technique can successfully reduce white noise and DSI with small levels. However, the EMD-based technique is not tolerant to DSI, even if it is involved with

small levels. Note that this intolerance of EMD-based technique may be due to the threshold estimation. However, the uncertainty of DSI information distributed in each IMF limits the development of new threshold estimation technique in EMD-based technique.

Apart from this, it is necessary to discuss more details regarding EMD-based technique in signal denoising. EMD expansion constitutes a critical element of EMD-based denoising. However, EMD is a simple, effective, and adaptive tool for signal decomposition but without any supportive theory for analysis. It is merely an algorithm that splits a signal into a few IMFs based on the number of extrema, cubic spline interpolation and a sifting iteration. As a result, the components of the IMFs can be easily affected by these factors. Also, it is claimed that the IMFs extracted from a signal are ranging from high frequency components with low order to low frequency components with high order. However, this is sometimes not the case due to the mode fixing problem of EMD. Based on this, it is difficult to interpret the information of the IMFs. It is not like wavelet transform that noise information may only concentrate on detail coefficient sequences, the noise and signal information are mixed together in the IMFs. The threshold estimation of each IMF for EMD-based denoising is highly dependent on the interpretation of the IMFs. The unexpected distortion of denoised signal can be easily encountered with an overestimated or underestimated threshold. This can be demonstrated through the denoising results shown in Figure 8-20 (b) and (d). Another issue raised by the IMFs in EMD-based denoising is for long-length signal processing. The analysis of long-length signals with EMD is extremely time-consuming, and sometimes even impossible in a reasonable time due to the fact cubic spline interpolation of a large number of points takes a lot of computer resources [191], [192]. For wavelet-based denoising, the long-length signals can be portioned into sections, and then, are processed separately to reduce the executing time. This segment-based processing of long-length signals, especially for those have fast changes in their frequency domain, is not suitable for EMD-based denoising. The number of IMFs extracted from each segment varies due to different number of extrema and sifting iterations, and thus, causes errors [192]. Generally, window functions are applied to address this issue, which increases the complexity of EMD-based technique for

PD denoising.

8.7 Conclusion

In previous chapters, the research output was demonstrated through simulated PD signals. To validate those newly proposed algorithms in real environment, PD data obtained from experiments in HV laboratory and onsite measurement were applied in this chapter. The results further support the findings in previous chapters, i.e., integrating those new algorithms into wavelet-based denoising can remarkably improve the effectiveness in PD detection. It can be seen that the PD pulses with small magnitudes can be successfully extracted using this improved wavelet-based denoising. In the meantime, EMD-based denoising was found ineffective in real PD denoising, which may be due to the noise estimation used. Comparisons of performance of the improved wavelet-based denoising and EMD-based denoising have been made through the aspects, i.e., executing time, distorting, effectiveness, adaptivity, and robustness, in PD signal processing. Results from real PD signal denoising indicate that the improved wavelet-based denoising is more effective than EMD-based denoising. As a result, the improved wavelet-based denoising was recommended for PD denoising in real situation.

9 Conclusion and Future Recommendations

9.1 Conclusions

Wavelet-based denoising has been intensively investigated for decades in the area of PD denoising but is still an area that faces challenges. Improvements of wavelet-based denoising can be achieved through wavelet selection, decomposition scale determination, and noise estimation. With further study of the statistical characteristics of PD signals and the associated noise, novel methods have been proposed in this research to improve the effectiveness of wavelet-based denoising through those three aspects mentioned above. In the meantime, EMD-based denoising, which does not require a predefined basis function for signal decomposition, has been discussed as an alternative to wavelet-based denoising in the context of PD signals. Comparisons between these two denoising strategies have been implemented through both simulated and real PD data. Details regarding these new methods and comparisons have been outlined in Chapter 4 – 8. A brief conclusion based on these contributions in this thesis are as follows:

1. A novel wavelet entropy-based wavelet selection scheme proposed in Chapter 4 was derived from the concept of Shannon entropy and the associated information cost function in information theory. This novel wavelet selection scheme selects the wavelet that can minimize the value of wavelet entropy of the approximation coefficient sequence at each decomposition scale, and thus, is adaptive and scale-dependent. It is one of the main contributions of this research project. Results of both simulated and real PD signals demonstrate that this new wavelet selection scheme can present better performance, i.e., the effectiveness of wavelet-based denoising can be improved by the use of this new method. As such, it can be a promising approach when applied to cable diagnostics. PD pulse variation during its propagation in the medium has been mentioned before. Although this new algorithm was demonstrated through simulated and real PD signals, the range of PD waveforms may still be far less than those in real situation. The following research work may continue based on the output of this as it sheds lights on the improvement of wavelet-based denoising for PD signals through the analysis of the

statistical characteristics.

2. Decomposition scale has not been received high attention in wavelet-based denoising, as it is normally selected by trial and error, or by experience, or by an empirical formula (termed wavelet length method in this thesis). Investigations in Chapter 5 demonstrated that the variation of decomposition scale did have a major effect on the denoising results. A sparsity-based method was then proposed as one major contribution based on the concepts of sparsity in compressive sensing and Shannon entropy in information theory. The sparsity of transformed PD signals can be reflected by the significant coefficient ratio, which is the ratio of the number of significant coefficients to the total length of that scale. Results from both simulated and real PD data show that it is an effective method for the decomposition scale determination in wavelet-based PD denoising. The new method for scale determination may have possible limit as same as that for the optimal wavelet selection, i.e., the involvement of PD waveforms in the investigation may not cover all PD waveforms from a practical point of view. However, the improvement of effectiveness in wavelet-based denoising indicates that it can be in a right track for PD denoising.
3. Chapter 6 firstly reviewed current existing threshold estimation techniques integrated in MATLAB, and it was found that the universal threshold estimation can provide the best threshold close to the noise level. The universal threshold is derived from noise estimation in noisy signals rather than the signal itself, however. Based on this, a novel SCR-based threshold estimation technique was then proposed based on the concept of significant coefficient ratio introduced in Chapter 5. The combination of the value of significant coefficient ratio and the number of significant coefficients can determine a reasonable threshold without the dependence on noise variance. The improvement of denoising results by the use of this new threshold has been demonstrated through both simulated and real PD signals. In the meantime, a new H-S thresholding function proposed in the processing of hydrologic series data was referenced in this thesis to overcome the discontinuous nature of traditional hard- and soft-thresholding functions. The denoising results of both simulated and real PD signals show that the effectiveness

of wavelet-based technique can be improved.

4. Chapter 7 reviewed the methods for RMS in EMD-based denoising. Based on the limits of these methods, the energy entropy of IMFs was proposed as a new method for RMS. The denoising results of simulated PD signals showed that it only works well when the SNR of original PD signals is relatively high. However, this investigation revealed that RMS is not as important as expected for EMD-based technique in PD denoising. Numerous simulations showed that all IMFs remained for further denoising can achieve better denoising results. One serious issue was raised from the application of EMD-based denoising in PD signals, i.e., denoised PD signals received severe distortion, especially those PD signals with small magnitude. Also, EMD-based denoising was found ineffective in real PD denoising. This may be highly related to the noise estimation used.
5. Chapter 8 tested the proposed algorithms in previous chapters by introducing real PD data. Experiment on testing samples conducted in HV laboratory was presented with details. Also, real PD data from an experiment on a power cable sample and onsite measurement was provided by other researchers for algorithms' tests. Comparisons between wavelet-based and EMD-based denoising was made through executing time, distortion, effectiveness, adaptivity and robustness. The denoising results of real PD signals showed that the improved wavelet-based technique is more preferable for the denoising of PD detection of electrical apparatus in practice due to its strength in effectiveness.

9.2 Recommendations for Future Work

Although the improved wavelet-based technique can present good denoising results of PD signals, some aspects as the subsequent effects of the contributions made may prompt future investigations.

1. The novel entropy-based wavelet selection scheme is an adaptive, automated, and scale-dependent approach. The appealing features of this new scheme is based on a preset wavelet library. As a result, the scope of the wavelet library has a major effect on the denoising results. In this thesis, the wavelets from Daubechies family were used to

build the wavelet library. For a future work, this can be extended to include wavelets from, e.g., Symlets, Coiflet, and Biorthogonal wavelets. It could be that the scale-dependent wavelet from these new wavelet families can provide better denoising results than current wavelet selected.

2. A SCR-based threshold estimation was proposed based on the number of significant coefficients and the significant coefficient ratio. Based on numerous simulation results, the tolerance of this improved threshold to noise, e.g., DSI, is not significantly enhanced. More efforts can be put into the investigations of a more robust threshold estimation in the future, and thus, increase the tolerance of wavelet-based denoising to DSI.
3. Current threshold estimation in EMD-based denoising is based on a white noise model, which cannot render EMD-based technique as a competitive alternative to wavelet-based technique in PD denoising. Based on simulation results, EMD-based denoising can present similar or better denoising results, but severe distortion occurs for PD pulses with small magnitude. Also, ineffectiveness can be seen from its application for real PD data. As a result, a series of future work can be done to improve the performance of EMD-based technique in PD denoising. A brief summary of these work is as follows.
    a) As aforementioned, the extraction of IMF in current algorithm is based on the number of extrema. When noise involved, this extraction cannot distinguish signal and noise components. It imposes the difficulty on IMF interpretation. Based on this, it is necessary to develop a method that can improve the accuracy of interpretations of the IMFs, and thus, reduce the inaccuracy for EMD-based denoising.
    b) With the accurate interpretation of the IMFs, a more accurate threshold estimation can be proposed to overcome the drawbacks of current existing one.
    c) Consider the effect of noise from other noise sources, except white noise, in EMD-based technique. The characteristics of noise, e.g., DSI, of the IMFs are not investigated. A more robust threshold estimation can be proposed with further investigation of noise characteristics of the IMFs.
    d) For long-length signals, a new method needs to be developed to enable EMD-based denoising to cope with segment-based processing, and thus, reduce the total

executing time.

4. During the investigation of the sparsity-based method for decomposition scale determination, it is known that PD signals can present sparse nature in some domains, e.g., wavelet domain. This attractive feature of PD signals can be further investigated with the combination of some popular sparsity-based denoising algorithms, e.g., dictionary learning algorithm (K-SVD) and compressive sensing.

References

[1] D. W. Gross and J. G. Herbig, “Partial discharge fault location and diagnosis on HV power cables,” *2000 Conf. Electr. Insul. Dielectr. Phenom.*, pp. 630–633, 2000.

[2] J. Veen, “On-line Signal Analysis of Partial Discharges in Medium-Voltage Power Cables,” Technische Universiteit Eindhoven, 2005.

[3] P. L. Lewin, “Continuous On-line Condition Monitoring of HV Cable Systems,” pp. 3–7, 2005.

[4] E. Gulski *et al.*, “On-site Testing and PD Diagnosis of High Voltage Power Cables,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 15, no. 6, pp. 1691–1700, 2008.

[5] P. Wagenaars, “Integration of Online Partial Discharge Monitoring and Defect Location in Medium-Voltage Cable Networks,” Technische Universiteit Eindhoven, 2010.

[6] P. Wagenaars, P. A. A. F. Wouters, P. C. J. M. Van Der Wielen, and E. F. Steennis, “Influence of ring main units and substations on online partial-discharge detection and location in medium-voltage cable networks,” *IEEE Trans. Power Deliv.*, vol. 26, no. 2, pp. 1064–1071, 2011.

[7] G. C. Montanari and A. Cavallini, “Partial discharge diagnostics: From apparatus monitoring to smart grid assessment,” *IEEE Electr. Insul. Mag.*, vol. 29, no. 3, pp. 8–17, 2013.

[8] S. Boggs, A. Pathak, and P. Walker, “Partial Discharge XXII: High Frequency Attenuation in Sheielded Solid Dielectric Power Cable and Implications Thereof for PD Location,” *IEEE Electr. Insul. Mag.*, vol. 12, no. 1, 1996.

[9] C. Zhou, X. Zhou, B. Stewart, A. Nesbitt, D. Hepburn, and D. Guo, “Comparisons of digital filter, matched filter and wavelet transform in PD detection,” *CIGRE*, 2006.

[10] F. P. Mohamed, W. H. Siew, J. J. Soraghan, S. M. Strachan, and J. McWilliam, “The Use of Power Frequency Current Transformers as Partial Discharge Sensors for Underground Cables,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 20, pp. 814–824, 2013.

[11] P. C. J. M. van der Wielen, J. Veen, P. A. A. F. Wouters, and F. Steennis, “On-line Partial Discharge Detection of MV Cables with Defect Localisation (PDOL) Based on Two Time Synchronised Sensors,” *18th Int. Conf. Electr. Distrib.*, 2005.

[12] 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 Trans. Dielectr. Electr. Insul.*, vol. 14, no. 1, pp. 3–14, 2007.

[13] Y. Tang, C. Tai, C. Su, C. Chen, and J. Chen, “A correlated empirical mode decomposition method for partial discharge signal denoising,” *Meas. Sci. Technol.*, 2010.

[14] H. Borsi, “Digital Location of Partial Discharges in HV Cables,” *IEEE Trans. Electr. Insul.*, vol. 27, no. 1, pp. 28–36, 1992.

[15] M. S. Mashikian, F. Palmieri, R. Bansal, and R. B. Northrop, “Location of Partial Discharges in Shielded Cables in the Presence of High Noise,” *IEEE Trans. Electr. Insul.*, vol. 27, no. 1, pp. 37–43, 1992.

[16] U. Kopf and K. Feser, “Rejection of Narrow-band Noise and Repetitive Pulses in

On-site P D Measurements," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 2, no. 3, 1995.

[17] J.Veen and P. C. J. M. van der Wielen, "The Application of Matched Filters to PD Detection and Localization," *IEEE Electr. Insul. Mag.*, pp. 20–26, 2003.

[18] P. Wagenaars, P. A. A. F. Wouters, P. C. J. M. Van Der Wielen, and E. F. Steennis, "Adaptive templates for matched filter bank for continuous online partial discharge monitoring," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 18, no. 5, pp. 1693–1701, 2011.

[19] X. Ma, C. Zhou, and I. J. Kemp, "Investigation into the use of wavelet theory for partial discharge pulse extraction in electrically noisy environments," *Dielectr. Mater. Meas. Appl. 2000. Eighth Int. Conf. (IEE Conf. Publ. No. 473)*, vol. 4, pp. 123–126, 2000.

[20] I. Shim, J. J. Soraghan, and W. H. Siew, "Detection of PD Utilizing Digital Signal Processing Method. Part 3 : Open-Loop Noise," *Electr. Insul. Mag. IEEE*, vol. 17, pp. 6–13, 2001.

[21] X. Ma, C. Zhou, and I. J. Kemp, "Interpretation of wavelet analysis and its application in partial discharge detection," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 9, no. 3, pp. 446–457, 2002.

[22] X. Ma, C. Zhou, and I. J. Kemp, "Automated wavelet selection and thresholding for PD detection," *IEEE Electr. Insul. Mag.*, vol. 18, no. 2, pp. 37–47, 2002.

[23] L. Satish and B. Nazneen, "Wavelet-based Denoising of Partial Discharge Signals Buried in Excessive Noise and Interference," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 10, no. 2, pp. 354–367, 2003.

[24] P. Ray, A. K. Maitra, and A. Basuray, "Extract Partial Discharge Signal using Wavelet for On-line Measurement," *Int. Conf. Commun. Signal Process.*, pp. 888–892, 2013.

[25] R. Hussein, K. B. Shaban, and A. H. El-Hag, "Wavelet Transform with Histogram-Based Threshold Estimation for Online Partial Discharge Signal Denoising," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 12, pp. 3601–3614, 2015.

[26] H. C.-J. Qian Yong Jiang Xiu-Chen, "Empirical mode decomposition based denoising of partial discharge signals ," *Proc. 5th WSEAS/IASME Int. Conferr. Electr. Power Syst. High Voltages, Electr. Mach. Tenerife, Spain*, vol. 2005, pp. 172–176, 2005.

[27] C. Herold, S. Wenig, and T. Leibfried, "Advanced de-noising of power cable Partial Discharge signals by Empirical Mode Decomposition," *20th Australas. Univ. Power Eng. Conf.*, 2010.

[28] M. Y. Lin, C. C. Tai, Y. W. Tang, and C. C. Su, "Partial discharge signal extracting using the empirical mode decomposition with wavelet transform," *Light. APL 2011 7th AsiaPacific Int. Conf.*, pp. 420–424, 2011.

[29] Y. Pei, Y. Wu, and D. Jia, "Research on PD Signals Denoising Based on EMD Method," *PRZEGLĄD ELEKTROTECHNICZNY (Electrical Rev.*, pp. 137–140, 2012.

[30] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction*. John Wiley & Sons, Inc., 2000.

[31] IEC60270, "High-voltage test techniques - Partial discharge measurements." 2000.

[32] BSEN60270, "High-voltage test techniques - Partial discharge measurements." 2001.

[33] IEEE Std 400.3, "IEEE Guide for Partial Discharge Testing of Shielded Power Cable Systems in a Field Environment," 2006.

[34] S. Boggs and J. Densley, “Fundamentals of partial discharge in the context of field cable testing,” *IEEE Electr. Insul. Mag.*, vol. 16, no. 5, pp. 13–18, 2000.

[35] P. H. F. Morshuis, “Partial discharge mechanisms in voids related to dielectric degradation,” *Sci. Meas. Technol. IEE Proc.*, vol. 142, no. 1, pp. 62–68, 1995.

[36] C. Mayoux and C. Laurent, “Contribution of Partial Discharges to Electrical Breakdown of Solid Insulating Materials,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 2, no. 4, pp. 641–652, 1995.

[37] S. A. Boggs, “Partial Discharge: Overview and Signal Generation,” *IEEE Electr. Insul. Mag.*, vol. 6, no. 4, pp. 33–39, 1990.

[38] E. Lemke *et al.*, “Guide for Electrical Partial Discharge Measurements in compliance to IEC 60270,” 2008.

[39] F.H. Kreuger, *Partial Discharge Detection in High-Voltage Equipment*. Butterworth-Heinemann Ltd, 1989.

[40] S. A. Boggs, “Partial Discharge - Part III: Cvity-Induced PD in Solid Dielectrics,” *IEEE Electr. Insul. Mag.*, vol. 6, no. 6, 1990.

[41] L. A. Dissado and J. C. Fothergill, *Electrical Degradation and Breakdown in Polymers*. London: Peter Peregrinus Ltd, 1992.

[42] P. H. F. Morshuis and J. J. Smit, “Partial discharges at dc voltage: Their mechanism, detection and analysis,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 12, no. 2, pp. 328–340, 2005.

[43] U. Fromm and F. H. Kreuger, “Statistical behaviour of internal partial discharges at DC voltage,” *Jpn. J. Appl. Phys.*, vol. 33, no. 12R, pp. 670–675, 1994.

[44] E. Kuffel, W. S. Zaengl, and J. Kuffel, *High Voltage Engineering Fundamentals*, Second Edi. Butterworth-Heinemann, 2000.

[45] R. Bartnikas, “Partial Discharges Their Mechanism, Detection and Measurement,” *IEEE Trans. Dielectr. Electr. Insualtion*, vol. 9, no. 5, pp. 763–808, 2002.

[46] R. J. Van Brunt, E. W. Cernyar, and P. Glahn, “Importance of Unraveling Memory Propagation Effects in Interpreting Data on Partial Discharge Statistics,” *IEEE Trans. Electr. Insul.*, vol. 28, no. 6, pp. 905–916, 1993.

[47] P. Von Glahn and R. J. Van Brunt, “Continuous Recording and Stochastic Analysis of PD,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 2, no. 4, pp. 590–601, 1995.

[48] D. F. M. Hui, “Online Partial Discharge Detection and Signal Analysis for High Voltage Cables,” University of Strathclyde, 2005.

[49] R. Bartnikas and E. J. McMahon, *Engineering Dielectrics Volume I Corona Measurement and Interpretation*. American Society for Testing and Materials, 1979.

[50] N. G. Trinh, “Partial Discharge XIX: Discharge in Air Part I: Physical Mechanisms,” *IEEE Electr. Insul. Mag.*, vol. 11, no. 2, pp. 23–29, 1995.

[51] L. A. Dissado, “Understanding electrical trees in solids: From experiment to theory,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 9, no. 4, pp. 483–497, 2002.

[52] S. Dodd, “Electrical tree grown in epoxy resin.” [Online]. Available: https://www2.le.ac.uk/departments/engineering/people/academic-staff/steve-dodd.

[53] F. P. Mohamed, “Parital Discharge Diagnostics in Medium Voltage Cables,” University of Strathclyde, 2013.

[54] P. C. J. M. Van Der Wielen, E. F. Steennis, and P. A. A. F. Wouters, “Fundamental

Aspects of Excitation and Propagation of On-line Partial Discharge Signals in Three-phase Medium Voltage Cable Systems," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 10, no. 4, pp. 678–688, 2003.

[55] A. R. Mor, P. H. F. Morshuis, and J. J. Smit, "Comparison of Charge Estimation Methods in Partial Discharge Cable Measurements," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 22, no. 2, pp. 657–664, 2015.

[56] B. Sheng *et al.*, "A novel on-line cable pd localisation method based on cable transfer function and detected pd pulse rise-time," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 22, no. 4, pp. 2087–2096, 2015.

[57] R. Bartnikas and J. P. Novak, "Effect of overvoltage on the risetime and amplitude of PD pulses," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 2, no. 4, pp. 557–566, 1995.

[58] J. Chan, H. Ma, T. Saha, and C. Ekanayake, "Self-adaptive partial discharge signal de-noising based on ensemble empirical mode decomposition and automatic morphological thresholding," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 21, no. 1, pp. 294–303, 2014.

[59] T. Okamoto and T. Tanaka, "Novel partial discharge measurement: computer-aided measurement systems," *IEEE Trans. Electr. Insul.*, vol. EI-21, no. 6, pp. 1015–1019, 1986.

[60] E. Gulski and F. H. Kreuger, "Computer-aided recognition of discharge sources," *IEEE Trans. Electr. Insul.*, vol. 27, no. 1, pp. 82–92, 1992.

[61] G. Chen, "Diagnosis of High Voltage Insulation Systems," 2014.

[62] P. BRINCKERHOFF, "Cavan-Tyrone and Meath-Cavan 400kV Transmission Circuits Comparison of High Voltage Transmission Options: Alternating Current Overhead and Underground, and Direct Current Underground," 2009.

[63] NationalGrid, "Undergrounding high voltage electricity transmission lines The technical issues Contents," 2015.

[64] E. Zaccone, "E-Highway2050 Talk plan : underground & subsea cables," 2014.

[65] R. Bhattacharyya, "ELECTRICAL , THERMAL AND THERMO-MECHANICAL DESIGN OF MV XLPE CABLES," *8th Int. Conf. Insul. Power Cables*, 2011.

[66] J. C. Fothergill, "The Coming of Age of HVDC Extruded Power Cables," *Electr. Insul. Conf.*, pp. 124–137, 2014.

[67] Nexans, "6-36kV Medium Voltage Underground Power Cables," 2010.

[68] "Cable Sizing & Selection." [Online]. Available: https://www.12voltplanet.co.uk/cable-sizing-selection.html.

[69] W. A. Thue, *Electrical Power Cable Engineering*. Marcel Dekker Inc, 1999.

[70] X. Qi and S. Boggs, "Thermal and Mechanical Properties of EPR and XLPE Cable Compounds," *IEEE Electr. Insul. Mag.*, vol. 22, no. 3, pp. 19–24, 2006.

[71] G. C. Stone, "Partial Discharge - Part VII: Practical Techniques for Measuring PD in Operating Equipment," *IEEE Electr. Insul. Mag.*, vol. 7, no. 4, pp. 9–19, 1991.

[72] D. A. Genutis, "On-Line Partial Discharge Testing Applications — An Overview," 2010. [Online]. Available: www.netaworld.org.

[73] Y. Tian, P. L. Lewin, A. E. Davies, S. G. Swingler, S. J. Sutton, and G. M. Hathaway, "Comparison of On-Line Partial Discharge Detection Methods for HV Cable Joints," *IEEE Trans. Dielectr. Electr. Insualtion*, vol. 9, no. 4, 2002.

[74] I. J. Kemp, “Partial discharge plant-monitoring technology: Present and future developments,” *IEE Proc. Sci. Meas. Technol.*, vol. 142, no. I, 1995.

[75] Y. Tian, P. L. Lewin, A. E. Davies, and G. Hathaway, “Acoustic Emission Techniques for Partial Discharge Detection within Cable Insulation,” *Eighth Int. Conf. Dielectr. Mater. Meas. Appl.*, pp. 503–508, 2000.

[76] R. Schwarz and M. Muhr, “Modern technologies in optical partial discharge detection,” *Annu. Rep. - Conf. Electr. Insul. Dielectr. Phenomena, CEIDP*, pp. 163–166, 2007.

[77] P. Barmann, S. Kroll, and A. Sunesson, “Spectroscopic measurements of streamer filaments in electric breakdown in a dielectric liquid,” *J. Phys. D. Appl. Phys.*, vol. 29, pp. 1188–1196, 1996.

[78] R. Schwarz, M. Muhr, and S. Pack, “Partial discharge detection in oil with optical methods,” *IEEE Int. Conf. Dielectr. Liq.*, 2005.

[79] M. Muhr and R. Schwarz, “Experience with optical partial discharge detection,” *Mater. Sci. - Pol.*, 2009.

[80] B. Song, M. Ren, J. Zhou, and M. Dong, “A New Optical Method of Partial Discharge distant Positioning in GIS,” *2018 IEEE Electr. Insul. Conf.*, 2018.

[81] A. Zargari and T. R. Blackburn, “Application of optical fibre sensor for partial discharge detection\nin high-voltage power equipment,” *Proc. Conf. Electr. Insul. Dielectr. Phenom. - CEIDP ’96*, 1996.

[82] Y. Tian, P. L. Lewin, J. S. Wilkinson, S. J. Sutton, and S. G. Swingler, “Optically based partial discharge continuous monitoring system for HV cable joints,” *Electr. Insul. Dielectr. Phenomena, 2004. CEIDP ’04. 2004 Annu. Rep. Conf.*, pp. 21–24, 2004.

[83] W. Wang, Y. Wu, and D. He, “Method of EFPI Fiber Sensor in Partial Discharge Detection,” *nternational Conf. Comput. Distrib. Control Intell. Environ. Monit.*, pp. 407–410, 2012.

[84] S. Karmakar, N. K. Roy, and P. Kumbhakar, “Partial discharge measurement of transformer with ICT facilities,” *2009 Int. Conf. Power Syst. ICPS ’09*, 2009.

[85] E. Mohammadi, M. Niroomand, M. Rezaeian, and Z. Amini, “Partial discharge localization and classification using acoustic emission analysis in power transformer,” *INTELEC, Int. Telecommun. Energy Conf.*, 2009.

[86] F. Y. Chu, “SF6 Decomposition in Gas-Insulated Equipment,” *IEEE Trans. Electr. Insul.*, vol. EI-21, no. 5, pp. 693–725, 1986.

[87] S. A. Boggs and G. C. Stone, “Fundamental Limitations in the Mvieasurement of Corona and Partial Discharge,” *IEEE Trans. Electr. Insul.*, vol. EI-17, no. 2, pp. 143–150, 1982.

[88] J. Li, T. Jiang, S. Grzybowski, and C. Cheng, “Scale dependent wavelet selection for de-noising of partial discharge detection,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 17, no. 6, pp. 1705–1714, 2010.

[89] IEEE, *IEEE Guide for Partial Discharge Testing of Shielded Power Cable Systems in a Field Environment*. 2007.

[90] P. Morshuis, G. C. Montanari, and L. Fornasari, “Partial Discharge Diagnostics — Critical Steps towards On-line Monitoring,” *2014 IEEE PES T&D Conf. Expo.*, 2014.

[91] E. Gulski, F. Wester, and J. Smit, “Advanced Partial Discharge Diagnostic of MV Power Cable System Using Oscillating Wave Test System,” *Electr. Insul. Mag.*, vol. 16,

no. 2, pp. 17–25, 2000.

[92] IEEE, *IEEE Std 400™-2012, IEEE Guide for Field Testing and Evaluation of the Insulation of Shielded Power Cable Systems Rated 5 kV and Above*. 2012.

[93] B. Oyegoke, P. Hyvonen, M. Aro, N. Gao, and M. Danikas, "Selectivity of Damped ac (DAC) and VLF Voltages in After-laying Tests of Extruded MV Cable Systems," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 10, no. 5, pp. 874–882, 2003.

[94] R. Bodega, P. H. F. Morshuis, M. Lazzaroni, and F. J. Wester, "PD Recurrence in Cavities at Different Energizing Methods," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 2, pp. 251–258, 2004.

[95] F. Wester, "Condition Assessment of Power Cables using Partial Discharge Diagnosis at Damped AC Voltages," Delft University of Technology, 2004.

[96] F. J. Webster, E. Gulski, and J. J. Smit, "Detection of Partial Discharges at Different AC Voltage Stress in Power Cables," *IEEE Electr. Insul. Mag.*, vol. 23, no. 4, 2007.

[97] IEEE, *IEEE Guide for Field Testing of Shielded Power Cable Systems Using Very Low Frequency (VLF) (less than 1 Hz)*. 2013.

[98] N. Chatlani, "Advanced Signal Enhancement Techniques with Application to Speech and Hearing," University of Strathclyde, 2011.

[99] S. Brettschneider, E. Lemke, J. L. Hinkle, and M. Schneider, "Recent Field Experience in PD Assessment of Power Cables Using Oscillating Voltage Waveforms," *Conf. Rec. 2002 IEEE Int. Symp. Electr. Insul.*, pp. 546–552, 2002.

[100] N.Ahmed and N.Srinivas, "On-line Partial Discharge Diagnostic System in Power Cable System," *IEEE/PES Transm. Distrib. Conf. Expo.*, pp. 853–858, 2001.

[101] P. C. J. M. Van Der Wielen and E. F. Steennis, "A Centralized Condition Monitoring System for MV Power Cables based on On-line Partial Discharge Detection and Location," *2008 Int. Conf. Cond. Monit. Diagnosis*, 2008.

[102] M. Seltzer-grant, R. Mackinlay, D. Clark, and R. Shuttleworth, "Experiences from over 15 years of On-line Partial Discharge (OLPD) Testing of In-Service MV and HV Cables, Switchgear, Transformers and Rotating Machines," *IX Lat. Am. Robot. Symp. IEEE Colomb. Conf. Autom. Control*, 2011.

[103] M. Seltzer-grant, D. Denissov, R. Mackinlay, F. Petzold, L. Renforth, and H. Schlapp, "On-line Continuous PD Monitoring for in Service Distribution Class Cables and Switchgears," *21st Int. Conf. Electr. Distrib.*, 2011.

[104] N. H. Ahmed and N. N. Srinivas, "On-line Partial Discharge Detection in Cables," *IEEE Trans. Dielectr. Electr. Insualtion*, vol. 5, no. 2, pp. 181–188, 1998.

[105] M. M. Yaacob, M. A. Alsaedi, J. R. Rashed, A. M. Dakhil, and S. F. Atyah, "Review on Partial Discharge Detection Techniques Related to High Voltage Power Equipment Using Different Sensors," *Photonic Sensors*, vol. 4, no. 4, pp. 325–337, 2014.

[106] P. C. J. M. Van Der Wielen, J.Veen, P. A. A. F. Wouters, and E.F.Steennis, "Sensors for On-Line PD Detection in MV Power Cables and their Locations in Substations," *Proc. 7th Intemational Conf. Prop. Appl. Dielectr. Mater.*, pp. 215–219, 2003.

[107] R. Mackinlay and C. Walton, "Diagnostics for MV Cables and Switchgear as a Tool for Effective Asset Management," *Proc. CIRED*, 2001.

[108] M. Seltzer-grant, D. Denissov, L. Renforth, R. Mackinlay, H. Schlapp, and F. Petzold, "On-Line PD Spot Testing and Continuous Monitoring for in Service Power Cables –

Techniques and Field Experiences," *8th Int. Conf. Insul. Power Cables*, no. June, pp. 1–5, 2011.

[109] F. Álvarez, F. Garnacho, J. Ortego, and M. Sánchez-Urán, "Application of HFCT and UHF Sensors in On-Line Partial Discharge Measurements for Insulation Diagnosis of High Voltage Equipment," *Sensors*, vol. 15, no. 4, pp. 7360–7387, 2015.

[110] P. Wagenaars, I. Tigchelaar, P. Wouters, and P. C. J. M. Van Der Wielen, "Partial Discharge Propagation through Cable Systems with Cross-bonding Joints," *Proc. 20th Nord. Insul. Symp.*, no. 1, pp. 119–121, 2007.

[111] S. Tumanski, "Induction coil sensors—a review," *Meas. Sci. Technol.*, pp. R31–R46, 2007.

[112] Z. S. Zhang, D. M. Xiao, and Y. Li, "Rogowski air coil sensor technique for on-line partial discharge measurement of power cables," *IET Sci. Meas. Technol.*, vol. 3, no. 3, pp. 187–196, 2009.

[113] E. Pultrum and M. van Riet, "HF Partial Discharge Detection of HV Extruded Cable Accessories," *Jicable*, pp. 662–665, 1995.

[114] E. Ildstad, H. Faremo, O. Lillevik, and D. Linhjell, "Detection and Location of Partial Discharges Sources in XLPE Cables," *Proc.10th ISH*, 1997.

[115] D. Pommerenke, I. Krage, W. Kalkner, E. Lemke, and P. Schmiegel, "On-site PD Measurement on High Voltage Cable Accessories Using Integrated Sensors," *Proc. 9th ISH*, 1995.

[116] E. Lemke, T. Strehl, and D. Rubwurm, "New Development in the Field of PD Detection and Location in Power Cables under On-site Condition," *Proc. 11th ISH*, 1999.

[117] P. Craaz, R. Plath, R. Heinrich, and W. Kalkner, "Sensitive On-site PD Measurement and Location Using Directional Coupler Sensors in 110kV Prefabricated Joints," *Proc. 11th ISH*, 1999.

[118] V. Nagesh and B. I. Gururaj, "Evaluation of Digital Filters for Rejecting Discrete Spectral Interference in On-site PD Mwasurements," *IEEE Trans. Electr. Insul.*, vol. 28, no. 1, pp. 73–85, 1993.

[119] X. Zhou, C. Zhou, and I. J. Kemp, "An Improved Methodology for Application of Wavelet Transform to Partial Discharge Measurement Denoising," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 12, no. 3, pp. 586–594, 2005.

[120] G. M. Hashmi, "Partial Discharge Detection for Condition Monitoring of Covered-conductor Overhead Distribution Networks using Rogowski Coil," Helsinki University of Technology, 2008.

[121] F. You and Y. Zhang, "Research of an Improved Wavelet Threshold Denoising Method for Transformer Partial Discharge Signal," *J. Multimed.*, vol. 8, no. 1, pp. 56–63, 2013.

[122] Y. H. Kim, Y. W. Youn, S. H. Yi, D. H. Hwang, J. H. Sun, and J. H. Lee, "High-Resolution Partial Discharge Location Estimation in Power Cables," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 21, no. 2, pp. 758–765, 2014.

[123] J. Li, C. Sun, and J. Yang, "Adaptive De-noising for PD Online Monitoring Based on Wavelet Transform," *IEEE Southeast Conf.*, pp. 71–74, 2006.

[124] Z. Du and M. S. Mashikian, "Self-learning Digital Filter for the Filed Location of Partial Discharge in Cables," *IEEE Int. Symp. Electr. Insul.*, pp. 245–248, 1994.

[125] Z. Chengke, M. M, D. M. Hepburn, and X. Song, “On-line partial discharge monitoring in medium voltage underground cables,” *IET Sci. Meas. Technol.*, vol. 3, pp. 354–363, 2009.
[126] W. Hauschild and E. Lemke, *High-Voltage Test and Measuring Techniques*. 2014.
[127] K. Feser, G. Konig, J. Ott, and P. Seitz, “An adaptive filter algorithm for on-site partial discharge measurements.pdf,” *IEEE Int. Symp. Electr. Insul.*, 1988.
[128] J. P. Steiner, P. H. Reynolds, and L. Weeks, “Estimating the Location Discharge in Cables,” *IEEE Trans. Electr. Insul.*, vol. 27, no. 1, pp. 44–59, 1992.
[129] C. Langton, “LTI Systems and Matched Filter,” 2004. [Online]. Available: www.complextoreal.com.
[130] M. H. Foo, “Online Partial Discharge Detection and Signal Analysis for High Voltage Cables,” PhD Thesis, University of Strathclyde, 2005.
[131] A. Schulz, E. A. B. da Silva, and L. Velho, “Compressive Sensing,” 2009. [Online]. Available: http://www.impa.br.
[132] R. Polikar, “The Wavelet Tutorial,” 1994. [Online]. Available: http://users.rowan.edu/~polikar/wavelets/wttutorial.html.
[133] F. Hlawatsch and G. F. Boudreaux-Bartels, “Linear and quadratic time-frequency signal representations,” *IEEE Signal Process. Mag.*, vol. 9, no. 2, pp. 21–67, 1992.
[134] Y. Sheng, “Wavelet Transform,” in *The Transforms and Applications Handbook: Second Edition*, 2000.
[135] C. Valens, *A Really Friendly Guide to Wavelets*. 1999.
[136] G. Kaiser, *A Friendly Guide to Wavelets*. Birkhauser, 1994.
[137] S. Alessio, *Digital Signal Processing and Spectral Analysis for Scientists: Concepts and Applications*. Springer, 2015.
[138] A. O. Boudraa, J. C. Cexus, and Z. Saidi, “EMD-based Signal Noise Reduction,” *Int. J. Signal Process*, no. October, pp. 33–37, 2004.
[139] Y. Kopsinis and S. McLaughlin, “Development of EMD-Based Denoising Methods Inspired by Wavelet Thresholding,” *Signal Process. IEEE Trans.*, vol. 57, no. 4, pp. 1351–1362, 2009.
[140] D. L. Donoho and I. M. Johnstone, “Ideal spatial adaptation by wavelet shrinkage,” *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
[141] N. E. Huang *et al.*, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proc. R. Soc. London A Math. Phys. Eng. Sci.*, pp. 903–995, 1998.
[142] G. Rilling, P. Flandrin, P. Gon, and D. Lyon, “On Empirical Mode Decomposition and its Algorithms,” *IEEE-EURASIP Work. Nonlinear Signal Image Process.*, 2003.
[143] N. Chatlani and J. J. Soraghan, “EMD-based Filtering (EMDF) of Low-Frequency Noise for Speech Enhancement,” *IEEE Trans. Audio. Speech. Lang. Processing*, vol. 20, no. 4, pp. 1158–1166, 2012.
[144] Y. Mao and P. Que, “Noise Suppression and Flaw Detection of Ultrasonic Signals via Empirical Mode Decomposition,” *Russ. J. Nondestruct. Test.*, vol. 43, no. 3, pp. 196–203, 2007.
[145] P. Flandrin, P. Gonçalvès, and Gabriel Rilling, “Detrending and Denoising With Empirical Mode Decomposition,” *2004 12th Eur. Signal Process. Conf.*, pp. 1581–

1584, 2004.

[146] K. Zeng and M. He, “A simple boundary process technique for empirical mode decomposition,” *IEEE Int. Geosci. Remote Sens. Symp. Proceedings. 2004*, pp. 4258–4261, 2004.

[147] R. Deering and J. F. Kaiser, “The use of a Masking Signal to Improve Empirical Mode Decomposition,” *ICASSP*, pp. 485–488, 2005.

[148] A. O. Boudraa and J. C. Cexus, “EMD-based signal filtering,” *IEEE Trans. Instrum. Meas.*, vol. 56, no. 6, pp. 2196–2202, 2007.

[149] N. Senroy, S. Suryanarayanan, and P. F. Ribeiro, “An Improved Hilbert–Huang Method for Analysis of Time-Varying Waveforms in Power Quality,” *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1843–1850, 2007.

[150] G. Rilling and P. Flandrin, “One or Two Frequencies? The Empirical Mode Decomposition Answers,” *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 85–95, 2008.

[151] A. Hill, B. G. Stewart, S. G. McMeekin, and G. Morison, “Partial Discharge Signal Denoising Using the Empirical Mode Decomposition,” *Int. J. Signal Process. Syst.*, vol. 1, no. 2, pp. 268–272, 2013.

[152] P. Flandrin, G. Rilling, and P. Goncalves, “Empirical mode decomposition as a filter bank,” *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 112–114, 2004.

[153] Z. Wu and N. E. Huang, “A study of the characteristics of white noise using the empirical mode decomposition method,” *Proc. R. Soc. Lond. A*, vol. 460, pp. 1597–1611, 2004.

[154] A. O. Boudraa and J. C. Cexus, “Denoising via Empirical Mode Decomposition,” *Proc. IEEE ISCCSP*, vol. 2, no. Ea 3876, pp. 1–4, 1999.

[155] T. Jing-tian, Z. Qing, T. Yan, L. Bin, and Z. Xiao-kai, “Hilbert-Huang transform for ECG de-noising,” *2007 1st Int. Conf. Bioinforma. Biomed. Eng. ICBBE*, pp. 664–667, 2007.

[156] R. E. Ziemer and W. H. Tranter, *Principles of Communications*. John Wiley & Sons, Inc., 2009.

[157] Y. F. Sang, D. Wang, J. C. Wu, Q. P. Zhu, and L. Wang, “The relation between periods’ identification and noises in hydrologic series data,” *J. Hydrol.*, vol. 368, pp. 165–177, 2009.

[158] B. Vigneshwaran, R. V. Maheswari, and P. Subburaj, “An improved threshold estimation technique for partial discharge signal denoising using Wavelet Transform,” *2013 Int. Conf. Circuits, Power Comput. Technol.*, pp. 300–305, 2013.

[159] C. E. Shannon, “A Mathematical Theory of Communication,” *Mob. Comput. Commun. Rev.*, vol. 5, no. 1, 1948.

[160] R. R. Coifman and M. V. Wickerhauser, “Entropy-based algorithms for best basis selection,” *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 713–718, 1992.

[161] M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*. IEEE Press, 1994.

[162] H. Šikić and M. V. Wickerhauser, “Information cost functions,” *Appl. Comput. Harmon. Anal.*, pp. 147–166, 2001.

[163] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelet Toolbox™ User’s Guide*.

2015.

[164] Steven W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*. San Diego: California Technical Publishing, 1999.

[165] DeCarlo L T, "On the Meaning and Use of Kurtosis," *Psychol. Methods*, vol. 2, no. 3, pp. 292–307, 1997.

[166] Y. Zhuang and J. S. Baras, "Existence and Construction of Optimal Wavelet Basis for Signal Representation," *IEEE Trans. Signal Process.*, 1994.

[167] Z. Y. He, X. Chen, and G. Luo, "Wavelet Entropy Measure Definition and Its Application for Transmission Line Fault Detection and Identification (Part I: Definition and Methodology)," *2006 Int. Conf. Power Syst. Technol.*, 2006.

[168] Y. F. Sang, D. Wang, and J. C. Wu, "Entropy-based method of choosing the decomposition level in wavelet threshold de-noising," *Entropy*, vol. 12, no. 6, pp. 1499–1513, 2010.

[169] L. Lei, C. Wang, and X. Liu, "Discrete Wavelet Transform Decomposition Level Determination Exploiting Sparseness Measurement," *Int. J. Electr. Comput. Energ. Electron. Commun. Eng.*, vol. 7, no. 9, pp. 691–694, 2013.

[170] C. F. F. C. Cunha, A. T. Carvalho, M. R. Petraglia, and A. C. S. Lima, "A new wavelet selection method for partial discharge denoising," *Electr. Power Syst. Res.*, vol. 125, pp. 184–195, 2015.

[171] A. M. R. Dixon, E. G. Allstot, D. Gangopadhyay, and D. J. Allstot, "Compressed sensing system considerations for ECG and EMG wireless biosensors," *IEEE Trans. Biomed. Circuits Syst.*, vol. 6, no. 2, pp. 156–166, 2012.

[172] J. E. Ruppert-Felsot, O. Praud, E. Sharon, and H. L. Swinney, "Extraction of coherent structures in a rotating turbulent flow experiment," *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, vol. 72, 2005.

[173] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. 2012.

[174] A. Antoniadis, J. Bigot, and T. Sapatinas, "Wavelet Estimators in Nonparametric Regression : A Comparative Simulation Study," *J. Stat. Softw.*, vol. 6, pp. 1–83, 2001.

[175] M. C. E. Rosas-Orea, M. Hernandez-Diaz, V. Alarcon-Aquino, and L. G. Guerrero-Ojeda, "A Comparative Simulation Study of Wavelet Based Denoising Algorithms," in *Proceedings of the 15th International Conference on Electronics,Communications and Computers*, 2005.

[176] B. Vidakovic, "Nonlinear wavelet shrinkage with bayes rules and bayes factors," *J. Am. Stat. Assoc.*, vol. 93, no. 441, pp. 173–179, 1998.

[177] I. Shim, J. J. Soraghan, and W. H. Siew, "A noise reduction technique for on-line detection and location of partial discharges in high voltage cable networks," *Meas. Sci. Technol.*, vol. 11, no. 12, pp. 1708–1713, 2000.

[178] A. M. Gaouda, A. El-Hag, T. K. Abdel-Galil, M. M. A. Salama, and R. Bartnikas, "On-line Detection and Measurement of Partial Discharge Signals in a Noisy Environment," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 15, no. 4, pp. 1162–1173, 2008.

[179] M. Jayakrishnan, B. N. Rao, K. P. Meena, and R. Arunjothi, "Optimum Threshold Estimator for De-noising Partial Discharge Signal using Wavelet Transform Technique," *2015 Int. Conf. Cond. Assess. Tech. Electr. Syst.*, pp. 76–82, 2015.

[180] M. Jansen, “Minimum Risk Thresholds for Data With Heavy Noise,” *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 296–299, 2006.

[181] G. . Nason, “Choice of the Threshold Parameter in Wavelet Function Estimation,” *Wavelets Stat. Lect. Notes*, vol. 103, pp. 261–280, 1995.

[182] L. B. Montefusco and S. Papi, “A Parameter Selection Method for Wavelet Shrikage Denoising,” *BIT Numer. Math.*, vol. 43, pp. 611–626, 2003.

[183] J. Li, C. Cheng, T. Jiang, and S. Grzybowski, “Wavelet De-noising of Partial Discharge Signals Based on Genetic Adaptive Threshold Estimation,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 19, no. 2, pp. 543–549, 2012.

[184] Y.-F. Sang, D. Wang, J.-C. Wu, Q.-P. Zhu, and L. Wang, “Entropy-Based Wavelet De-noising Method for Time Series Analysis,” *Entropy*, vol. 11, no. 4, pp. 1123–1148, 2009.

[185] P. Ray, A. K. Maitra, and A. Basuray, “A New Threshold Function for Signal Denoising Based on Wavelet Transform,” *Int. Conf. Signal Process. Image Process. Pattern Recognnition*, 2013.

[186] X. P. Zhang and M. D. Desai, “Adaptive denoising based on SURE risk,” *IEEE Signal Process. Lett.*, vol. 5, no. 10, pp. 265–267, 1998.

[187] P. Flandrin, P. Goncalves, and G. Rilling, “EMD Equivalent Filter Banks, From Interpretation to Applications,” in *Hilbert-Huang Transform and Its Applications*, 2005, pp. 67–87.

[188] G. Yang, Y. Liu, Y. Wang, and Z. Zhu, “EMD interval thresholding denoising based on similarity measure to select relevant modes,” *Signal Processing*, vol. 109, pp. 95–109, 2015.

[189] S. Y. Zhang, Y. Y. Liu, and G. L. Yang, “EMD interval thresholding denoising based on correlation coefficient to select relevant modes,” *Proceeding 34th Chinese Control Conf.*, 2015.

[190] X. Hu, W. H. Siew, M. D. Judd, and X. Peng, “Transfer function characterization for HFCTs used in partial discharge detection,” *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 2, pp. 1088–1096, 2017.

[191] P. Stepien, “Sliding Window Empirical Mode Decomposition -its performance and quality,” *EPJ Nonlinear Biomed. Phys.*, 2014.

[192] M. D. Ortigueira, R. T. Rato, and J. J. Trujillo, “An Improved Empirical Mode Decomposition for Long Signals,” *First Int. Conf. Adv. Signal, Image Video Process.*, pp. 58–62, 2016.