Partial Discharge (PD) Signal Detection and Isolation on High Voltage Equipment Using Improved Complete EEMD Method

Vu Cong Thuc 1,2 and Han Soo Lee 1,3,*

1 Transdisciplinary Science and Engineering Program, Graduate School of Advanced Science and Engineering, Hiroshima University, 1-5-1 Kagamiyama, Higashi-Hiroshima 739-8529, Japan
2 Hanoi Electrical Testing Company, Cau Giay, Ha Noi 100000, Vietnam
3 Center for the Planetary Health and Innovation Science (PHIS), The IDEC Institute, Hiroshima University
* Correspondence: leehs@hiroshima-u.ac.jp

Abstract: Electricity has a crucial function in contemporary civilization. The power grid must be stable to ensure the efficiency and dependability of electrical equipment. This implies that the high-voltage equipment at the substation must be reliably operated. As a result, the appropriate and dependable use of systems to monitor the operating status of high-voltage electrical equipment has recently gained attention. Partial discharge (PD) analysis is one of the most promising solutions for monitoring and diagnosing potential problems in insulation systems. Noise is a major challenge in diagnosing and detecting defects when using this measurement. This study aims to denoise PD signals using a data decomposition method, improved complete ensemble empirical mode decomposition with adaptive noise algorithm, combined with statistical significance test to increase noise reduction efficiency and to derive and visualize the Hilbert spectrum of the input signal in time-frequency domain after filtering the noise. In the PD signal analysis, both artificial and experimental signals were used as input signals in the decomposition method. For these signals, this study has yielded significant improvement in the denoising and the PD detecting process indicated by statistical measures. Thus, the signal decomposition by using the proposed method is proven to be a useful tool for diagnosing the PD on high voltage equipment.

Keywords: empirical mode decomposition; EMD; intrinsic mode function; white noise; denoising and filtering process

1. Introduction

The substation is a crucial part of the electrical system. It includes numerous expensive and complex device such as transformers, power cables, circuit breakers, and so on. Faulty primary equipment will result in serious losses, including damages in terms of repair and replacement costs of damaged objects as well as income loss due to power outages. A more critical problem is reduced power supply reliability due to frequent power interruptions [1]. Thus, maintenance and condition evaluation challenges are given much attention. The trend from time-based maintenance to condition-based maintenance is an inevitable trend to reduce unexpected system outages and improve power supply reliability [2].

Under the stress of electrical field strength, effects of environmental conditions such as temperature and humidity, and other problems during operation, the insulation system of primary equipment will degrade over time. Partial discharge (PD) measurement is an efficient and highly sensitive method that demands experienced-operator skill in data collection, processing, and analysis [3,4]. The collected signals in PD measurement are usually in the form of oscillating pulses with a wide frequency range. These signals can be represented as oscillating pulses in the time domain or as the spectrum produced by transient pulses in the frequency domain [5,6]. There are a number of studies and articles on frequency domain [7–9]. However, signal analysis in the frequency domain might result
in information loss in the time domain [10]. In addition, there are also other problems that cause difficulty and confusion such as selecting the appropriate parameter (midband frequency, bandwidth, PD threshold, trigger, and so on) for the adaptive filter, which may lead to misdiagnosis or unstable patterns if these parameters are set incorrectly [11]. In addition, PD pulses are a transient, nonlinear, and nonstationary signal [12,13] and when comparing the performance of artificial neural networks, fast Fourier transform (FFT), and discrete wavelet transform (DWT), FFT is unsuitable for non-stationary signal processing [14]. Even though there are many noise-filtering tools in frequency domain, such as three center frequencies of relation diagram, multi-channel synchronous of PD measurement [15–17], and time resolved partial discharge (TRPD) [18,19], there are many challenging to extract waveform characteristics of PD pulses.

To overcome those difficulties, recently several methods have been studied that can analyze and generate signal on time-frequency domain such as wavelet transformation (WT) [20,21] or other methods [22]. However, only the WT approach is insufficient to denoise a non-stationary and nonlinear PD signal [23]. Although these methods achieve certain results in removing noise from the signal, there are still many problems such that it is difficult and confusing to choose wavelet basis functions, wavelet thresholds, decomposition levels, and so on since it depends largely on human experience [12,24].

After the empirical mode decomposition (EMD) was introduced in 1998, there has been many variations of EMD to improve noise filtering [25]. EMD based on wavelet threshold denoising method may efficiently reduce noise with minimal signal distortion [26]. Wu and Huang proposed Ensemble EMD (EEMD) in 2009 to overcome the frequency aliasing problem [27]. On the other hand, the EEMD method has two drawbacks, which are the decomposition is incomplete and the different realizations of signal mixed with white noise could generate different numbers of intrinsic mode functions (IMFs), especially in the low frequency [28]. The complete EEMD with adaptive noise (CEEMDAN) method was introduced to effectively handle the EMD decomposition scale inconsistency [29]. To solve the two disadvantages of EEMD method, Colominas et al. (2014) further improved the CEEMDAN method and proposed the improved CEEMDAN (ICEEMDAN) method, which uses the original signal as the target of IMF sifting [30]. The advantages of the ICEEMDAN method has been verified for nonlinear and non-stationary signals by numerous publications in the field of fault diagnosis [31,32].

Noise is one of the unavoidable factors of PD measurement for both conventional and non-conventional methods. Because primary equipment is connected to each other in the substations, it is difficult to determine where PD originates. In addition, in substation environments especially with transmission substations, the electrical field surrounding test object also affects data collection and may lead to a large amount of noise mixed with the PD signal. Various forms of noise may present during PD measurement, including white noise, regular impulses, oscillatory. Because of its likeness to PD signal, white noise is the most hardest to eliminate [3].

Therefore, this study aims to denoise PD signals using a data decomposition approach, ICEEMDAN algorithm, and statistical significance test to enhance noise reduction efficiency, and to create the Hilbert spectrum of the input signal in 3D after filtering the noise for visual observation and improving diagnostic skills of operators. The results are also verified by some quantitative evaluation parameters.

2. Partial Discharge (PD) Measurement

PD phenomenon occurs with signals in the form of electrical pulses, electromagnetic wave, acoustic emission, and chemical reactions. As a result, various PD measurement methods have been developed corresponding to the accompanying signal types, including the conventional method (the electrical method) and non-conventional methods (the electromagnetic method, the acoustic method, the optical method, and so on) [33–35]. It is necessary to have some information and comprehension of the PD processes while analyzing and implementing measuring circuits, as well as when selecting appropriate
2.1. Conventional Method

Conventional PD measurement is the method using electrical signals mentioned in IEC 60270 standard [5], i.e., measuring the apparent charge in the measuring circuit. The apparent charge is the charge if injected a pulse between the test object’s terminals in a specific test circuit within a very short time, would provide the same value on measuring instrument as the PD current pulse itself, the unit of measurement is usually pC. Conventional measurement test circuit includes coupling capacitor connected in series with coupling capacitor or test object, balanced circuit arrangement, polarity discrimination test circuit. The most popular circuit measurement in the field is to connect coupling device to coupling capacitor in series. The conventional method according to IEC 60270 is considered as a proven effective method for performing PD measurements by many professionals, applied engineers and in the field of academic research. This method can be applied on all primary equipment (power transformer, circuit breaker, instrument transformer, power cable and so on). The limitation of this method is that it requires to separate the test subject from the grid and use an external source to generate high voltages to apply in the insulation system of the test subject. Furthermore, this method is affected by high noise levels. However, the outstanding advantage is the ability to quantify the amplitude of the partial discharge quantity, determine the partial discharge inception voltage (PDIV), partial discharge extinction voltage (PDEV) values to quickly distinguish the tested high voltage equipment pass or no pass. Therefore, the conventional method is widely used as an application for testing new electrical equipment after installation, acceptance factory test, or laboratory applications.

2.2. Non-Conventional Method

The most notable advantage of the non-conventional method is that there is no need to separate the test subject from the grid, and it does not affect the test object as well as the operating system. There are many ways to collect PD signals by non-conventional method such as using high frequency current transformer (HFCT) sensor [6], transient earth voltage (TEV) sensor [36], acoustic emission (AE) sensor [37]. The common way is to use HFCT to measure the PD signal through the grounding or neutral system of high-voltage electrical equipment. When performing partial discharge measurements in the field, there are many factors that affect the measurement and data collection process, thereby indirectly affecting the ability to diagnose and evaluate the insulation system of primary equipment. The disadvantage is that there are no standard evaluation thresholds to confirm whether the test object passes or not.

3. Materials and Methods

3.1. Materials

Typical PD pulse shapes from RLC detector circuit is an oscillating pulse with a damping amplitude [38]. To demonstrate the efficiency to filter noise with a signal shaped like a PD signal, this paper uses a damped oscillator signal generated from the arbitrary function generator (AFG) with eight pulses blended with white noise. In addition, to consider the adaptation of the proposed method to a real signal that have nonlinear and nonstationary properties, a case study was conducted to collect the experimental signal. A defect was created by cutting a 3-mm deep hold on cross linked polyethylene (XLPE) insulation of medium voltage cable ($U_o = 12.7$ kV), then applying a test voltage of 12.7 kV to cable insulation and acquiring the PD signal with an HFCT sensor. The PD signal was then mixed with white noise to disrupt the signal. The white noise magnitude added to the artificial signal and the experimental signal is 0.5 (normalized) and 0.005 (V), respectively.
3.2. Methods

3.2.1. Empirical Mode Decomposition (EMD)

A decomposition is a separation of a signal into different components, then analyze the newly obtained components to gain new insight into the features inherent in the whole data. There are several methods to perform such a decomposition and the EMD is one of them. The most popular decomposition method is the Fourier transform that is a coordinate transformation projecting a data onto an orthogonal basis system with sines and cosines. It covers all frequencies that could be contained in the data [39]. In a strict sense, the Fourier transform is only valid for linear and stationary data, however, nearly all of the practical data are non-linear and non-stationary [40]. The wavelet transform is an extension of the Fourier transform [41]. In addition to indicating which frequencies are significant, it also displays the time the event happened; nonetheless, its limitations have been evaluated [23]. Huang et al. were the first to propose the EMD approach in 1998 [42]. In contrast to most other decompositions, the EMD is empirical based on a concise mathematical foundation but the data itself dictate the decomposition. However, this method was first proved mathematically [43] and so far there have been many variations of EMD.

The modes produced by the EMD are called intrinsic mode function (IMF). The IMF is described as follows: “An intrinsic mode function is a function that satisfies two conditions: (1) In the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one; and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.” [43]. After decomposing a signal into a set of IMFs in ascending order, the IMFs will show the corresponding descending frequency and each IMF will display the specific information of the signal in the time domain. Thus, the decomposed signal includes both time and frequency domain information. This method is performed according to the following steps [30,42]:

Step 1. Set \( k = 0 \) and find all extrema of \( r_0 = x \).

Step 2. Interpolate between minima (maxima) of \( r_k \) to obtain the lower (upper) envelope \( e_{\text{min}} (e_{\text{max}}) \).

Step 3. Compute the mean envelope \( m = (e_{\text{min}} + e_{\text{max}}) / 2 \).

Step 4. Compute the IMF candidate \( d_{k+1} = r_k - m \). IMF candidate is the result of the first sifting process.

Step 5. Is \( d_{k+1} \) an IMF? An IMF is an IMF candidate that satisfy two conditions of an IMF.

Yes. Save \( d_{k+1} \), compute the residue \( r_{k+1} = x - \sum_{i=1}^{k} d_i \), do \( k = k + 1 \), and treat \( r_k \) as input data in step 2. No. Treat \( d_{k+1} \) as input data in step 2.

Step 6. Continue until the final residue \( r_k \) satisfies the predefined stopping criterion.

The predefined stopping criterion is based on two thresholds \( \theta_1 \) and \( \theta_2 \). Stop condition will be implemented as follows: Calculate the mode amplitude \( a(t) = (e_{\text{min}} - e_{\text{max}})/2 \) and the evaluation function \( \sigma(t) = |m/a(t)| \) sifting process will be repeated if \( \sigma(t) < \theta_1 \) for some prescribed fraction \((1-\alpha)\) of the total duration, and \( \sigma(t) < \theta_2 \) for the remaining fraction. In this study, set \( \alpha = 0.05 \), \( \theta_1 = 0.05 \), and \( \theta_2 = 0.5 \).

Steps 2 to 5 are the steps to find out the IMF modes, this process is called the sifting process.

3.2.2. Improved CEEMDAN (ICEEMDAN)

The ICEEMDAN algorithm is presented as follows [30].

\( E_k(\cdot) \) is the EMD operator which create the \( k \)th mode. \( w^{(i)} \) is the white noise. \( M(\cdot) \) is the local mean operator.

Step 1. Determine the local means of I realization using EMD \( x^{(i)} = x + \beta_0 E_1(w^{(i)}) \) to get the first residual \( r_1 = \langle M_1(w^{(i)}) \rangle \), where \( I \) is the number of realization in the ensemble and the magnitude of additional noise \( \beta > 0 \).

Step 2. In the first phase (\( k = 1 \)) compute the first mode: \( \tilde{d}_1 = x - r_1 \).
Step 3. Estimate the second residue as the average of local means of the realizations 
\( r_1 + \beta_1 E_2(w(i)) \) and define the second mode: 
\[ d_2 = r_1 - r_2 = r_1 - \langle M(r_1 + \beta_1 E_2(w(i))) \rangle. \]
Step 4. For \( k = 3, \ldots, K \) calculate the \( k \)th residue \( r_k = \langle M(r_{k-1} + \beta_k E_k(w(i))) \rangle. \)
Step 5. Compute the \( k \)th mode \( d_k = r_{k-1} - r_k \)
Step 6. Go to step 4 for next \( k \).

Constants \( \beta_k = \varepsilon_k \text{std}(r_k) \) are selected to achieve a desired signal-to-noise ration (SNR) 
between the additional noise and the residue to which the noise is added. Figure 1 illustrates 
the ICEEMDAN algorithm flowchart.

![Flowchart of ICEEMDAN Algorithm](image)

**Figure 1.** The flowchart of ICEEMDAN algorithm.

### 3.2.3. Statistical Significance Test (SST)

After decomposing the original signal into IMFs using ICEEMDAN, these IMFs will 
contain both significant information and less significant information. IMFs containing less 
significant information are considered noise and should be removed. To determine which 
IMFs provide significant information to keep and which IMFs contain less information 
to remove, statistical significance test was introduced by Wu and Huang (2004). First, 
decompose the dataset of white noise which targeted into IMFs. Second, choose the confidence 
level and determine the upper and lower bounds. Finally, compare the IMF’s energy 
density from the data with the spread functions; the IMFs that have their energy located 
above the upper bound and below the lower bound should be providing information at 
that confidence level. After applying the statistical significant test, remove the IMFs which 
have their energy located inside the spread lines (non-significant IMFs) and reconstruct a 
signal by combining those significant IMFs only above and below the confidence level [28].
The spread lines may be adjusted based on the properties of the signal and the experience of user; for further information, see Appendix A.2 of Appendix A.

4. Results
4.1. Artificial Signal

To check the advantages of the mentioned method, this study will analyze the artificial signal to simulate a PD phenomenon. This artificial signal is a damped oscillatory pulse which is generated by an arbitrary function generator then it is mixed with white noise as Figures 2a and 2b, respectively. The unit of the artificial signals has been normalized. The sample rate of the artificial signal is 1024 Hz. The intrinsic mode functions of the mixed signal in Figure 2b are presented in Figure 3. As shown in Figure 3, the PD signal fundamentally appears in the IMF4–IMF7 and the white noise signal is primarily localized in the high frequency IMF1–IMF3. The IMF8–IM11 contain the interference signals of different frequencies at low energy.

![The original artificial signal](image1)

![Artificial signal mixed white noise](image2)

**Figure 2.** (a) The original artificial damped oscillation signal from the arbitrary function generator (AFG), and (b) the artificial signal mixed with white noise.

In the statistical significance test, we choose the confidence limit of 99% and determine the upper and lower bounds. These bounds are the energy-density spread function of the IMF components which are derived from the probability distribution. As shown in Figure 4a, there are some IMFs that have their energy located above the upper bound should be containing information, the others locate between two spread lines or locate very close these bounds may contain noisy information or contain insignificant information and should be removed. To enhance noise filtering capability, these boundaries are extended by 0.5 axial units. Therefore, the IMF1, IMF2, IMF3, IMF10, IMF11 are eliminated as in Figure 4b.
Figure 3. The artificial signal with white noise and the resulting intrinsic mode functions (IMFs), IMF1-IMF12, obtained after decomposing with ICEEMDAN algorithm.

In the statistical significance test, we choose the confidence limit of 99% and determine the upper and lower bounds. These bounds are the energy-density spread function of the IMF components which are derived from the probability distribution. As shown in Figure 4a, there are some IMFs that have their energy located above the upper bound should be containing information, the others locate between two spread lines or locate very close these bounds may contain noisy information or contain insignificant information and should be removed. To enhance noise filtering capability, these boundaries are extended by 0.5 axial units. Therefore, the IMF1, IMF2, IMF3, IMF10, IMF11 are eliminated as in Figure 4b.

Figure 4. (a) The statistical significance test (SST) outcomes for the IMFs (IMF1–11 from left to right) of the artificial signal, and (b) the result after removing insignificant IMFs.

As seen in the reconstructed signal after denoising in Figure 5a, the waveform with the proposed method is closer to the original signal. Both noise reduction and preserving the original signal’s properties yield good results. Figure 5b depicts the Hilbert spectrum of the de-noised signal in the time-frequency domain to assist with visualization. The extensive details about the artificial PD signal’s frequency, timing, and pattern, can be notified.
As seen in the reconstructed signal after denoising in Figure 5a, the waveform with the proposed method is closer to the original signal. Both noise reduction and preserving the original signal’s properties yield good results. Figure 5b depicts the Hilbert spectrum of the de-noised signal in the time-frequency domain to assist with visualization. The extensive details about the artificial PD signal’s frequency, timing, and pattern, can be noticed.

**Figure 5.** The results of the reconstructed artificial signal after denoising: (a) in the time domain, and (b) the Hilbert spectrum in the time-frequency domain.

Some common evaluation parameters are used to estimate the effectiveness of the suggested de-noising method [44,45], including:

1. The SNR of the input and output were determined using (1) and (2). The difference of SNR values was also computed using (3).

\[
\text{SNR}_{\text{input}} = 10 \log \frac{\sum_{i=1}^{n} s^2(i)}{\sum_{i=1}^{n} (N(n))^2} \tag{1}
\]

\[
\text{SNR}_{\text{output}} = 10 \log \frac{\sum_{i=1}^{n} \hat{s}^2(i)}{\sum_{i=1}^{n} (s(i) - \hat{s}(i))^2} \tag{2}
\]

\[
\Delta \text{SNR} = \text{SNR}_{\text{output}} - \text{SNR}_{\text{input}} \tag{3}
\]

where \(s(i)\) is the original PD signal, \(\hat{s}(i)\) is the de-noised signal and \(N(n)\) is the white noise.

2. Mean square error (MSE) is applied to compare the consistency between the original signal and the de-noised signal. The lower the MSE, the closer the original and the de-noised signals are;

\[
\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} |s(i) - \hat{s}(i)|^2 \tag{4}
\]

3. Normalized correlation coefficient (NCC) is a widely used criteria for determining signal similarity, having a value range of 0 to 1. The higher the NCC number, the more similar the two signals are;

\[
\text{NCC} = \frac{\sum_{i=1}^{n} s(i) \hat{s}(i)}{\sqrt{\sum_{i=1}^{n} s^2(i) \sum_{i=1}^{n} \hat{s}^2(i)}} \tag{5}
\]

Figure 6 provides a quantitative overview of the input and output SNR with respect to the magnitude of white noise added to the original signal. Figure 6a indicates that the maximum \(\Delta\text{SNR}\) was around 9.6 dB, and the minimum was approximately 4.8 dB, with an average of 8.3 dB. This demonstrates that, even when diverse noise amplitudes are used to disrupt the PD signal, the presented approach still has significant noise filtering capabilities, even when the SNR_{input} is negative. Figure 6b,c shows how the assessment parameters for
the proposed algorithms’ de-noising performance change with SNR. With varying SNR, the NCC results are almost equal to one and the MSE values are low, suggesting that the signal after denoising is nearly similar to the original signal.

Figure 6 provides a quantitative overview of the input and output SNR with respect to the magnitude of white noise added to the original signal. Figure 6a indicates that the maximum $\Delta$SNR was around 9.6 dB, and the minimum was approximately 4.8 dB, with an average of 8.3 dB. This demonstrates that, even when diverse noise amplitudes are used to disrupt the PD signal, the presented approach still has significant noise filtering capabilities, even when the SNR input is negative. Figure 6b,c shows how the assessment parameters for the proposed algorithms’ de-noising performance change with SNR. With varying SNR, the NCC results are almost equal to one and the MSE values are low, suggesting that the signal after denoising is nearly similar to the original signal.

4.2. Experimental Signal

The acquisition diagram for the experimental PD signal is described in Appendix A.3 of Appendix A. The sample rate of the simulate signal is $10^8$ Hz. The experimental PD signal demonstrated in Figure 7a is a typical signal when measuring PD on power cables due to the reflective properties of the power cable. If PD activity occurs in a power cable, two current pulses are generated. One signal travels to the beginning of the cable, while the other travels to the end and is reflected backward with diminished amplitude [46]. Due to the low level of noise interference in the laboratory condition and to keep the test results as close as possible to the PD in the field, the PD signal acquired by the experimental test is processed by adding the white noise as indicated in Figure 7b. Figure 8 displays the IMFs, where the noise signal is substantially concentrated in the IMF1–IMF2 and the PD
signal is mostly present in the IMF4–IMF6. IMF3 obtains the original pulse’s properties but not the reflected pulse’s characteristic, and it also includes a considerable quantity of noise. The remaining IMFs with higher frequencies have low energies and unclear signal properties. The statistical significance test result in Figure 9a depicts the energy density of 12 IMFs, except the residual (IMF13). Consequently, the denoise technique eliminates the IMF1–IMF3, IMF9, and IMF 12 as shown in Figure 9b.

Figure 7. (a) The original experimental signal, and (b) the experimental signal mixed with white noise.

Figure 8. The experimental signal with white noise and the resulting IMFs from ICEEMDAN.
Figure 8. The experimental signal with white noise and the resulting IMFs from ICEEMDAN.

Figure 9. (a) The statistical significance test results for the IMFs (IMF1–12 from left to right) of the experimental signal, and (b) the result after removing insignificant IMFs.

Figure 10 illustrates the outcome of reconstructing the signal from the significant IMFs containing the PD information. The white noise is effectively suppressed. In addition, the signal after denoising still has the PD pulse properties, especially the low amplitude reflected pulses at the end of signal. The PD pattern is clearly visible on the Hilbert spectrum.

Figure 10. The results of the reconstructed experimental signal after denoising: (a) in the time domain, and (b) the Hilbert spectrum in the time-frequency domain.

As may be observed, the signal in the time domain after denoising with the suggested approach is almost identical to the waveform of the original signal. Furthermore, the Hilbert spectrum provides us with comprehensive information on the PD phenomena in both the time-frequency domain. These results demonstrate that the proposed method is highly efficient in terms of detecting and extracting PD signals. Figure 11 displays the quantitative results of the described technique for the experimental signal in the laboratory. The quantitative computation results further indicate that the high efficacy of the applied method with the average of $\Delta$SNR is 5.8 dB. Other evaluation indices, as shown in Figure 11, also provide good results.
The proposed method effectively removes white noise while keeping and isolating respectively, for the experimental signal. 

Many studies have applied decomposition algorithms for investigating nonlinear and non-stationary signals, and it can be stated that the ICEEMDAN algorithm has a clear advantage [28,30,44,47]. This study combines the ICEEMDAN algorithm with the statistical significance test to depict their efficacy in denoising for PD measurements. It can be inferred that the mentioned method successfully filters the white noise from the PD signal and provides the following substantial benefits.

1. The proposed method effectively removes white noise while keeping and isolating the characteristics of the PD signal.
2. The ICEEMDAN algorithm, in conjunction with the statistically significant test approach, reduces the difficulty of picking significant IMFs and discarding insignificant IMFs.
3. The high SNR, delta SNR, and NCC parameters show that this method is very effective even when the signal amplitude is very low and the SNR_input is negative.

**5. Conclusions**

Data used in this study is available from authors upon requests.

**Figure 11.** (a) The difference of input and output SNR, (b) NCC and (c) MSE with varying SNR, respectively, for the experimental signal.
The Hilbert spectrum in time-frequency domain of denoised signal is very useful for quick fault categorization.

To accurately diagnose PD phenomenon in high voltage devices, it is necessary to collect signals of such defect types which will be addressed in future work. The findings described in this research are meant to lay the framework for future experimental studies. In future works, typical types of PD in the laboratory such as corona discharge, void discharge, internal discharge, and so on will be conducted and PD signals will be collected using conventional methods based on IEC 60270 standard to form a database for PD pattern analysis using machine learning algorithm.

**Author Contributions:** Conceptualization, software, formal analysis, investigation, resources, data curation, visualization, project administration, funding acquisition, H.S.L. and V.C.T.; methodology, H.S.L.; validation, H.S.L.; writing—original draft preparation, V.C.T.; writing—review and editing, H.S.L.; supervision, H.S.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** V.C.T. received a support from Japanese Grant-Aid Program for Human Resource Development Scholarship (JDS).

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data used in this study is available from authors upon requests.

**Acknowledgments:** The first author thanks Hiroshima University colleagues, particularly Chisale Sylvester William, for helpful discussion and comments. The first author also thanks his Vietnamese colleagues, Dang Tran Chuyen and Le Viet Cuong, for providing input data and useful discussions.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A**

**Appendix A.1. Partial Discharge Mechanism**

PD is the breakdown of a small part of the insulation of an insulating system under high voltage stress, which over time gradually evolves into a fault. Definition from IEC 60270 standard, PD is “localized electrical discharge that only partially bridges the insulation between conductors, and which can or cannot occur adjacent to a conductor” [5]. Partial discharge is a complex and unpredictable phenomenon because the primary equipment can be badly affected by thermal, mechanical, electrical, and environmental stresses due to overload, moisture, heat, and other critical operational conditions [48]. To simplify the operation of PDs, Figure A1 shows the dielectric of a capacitor includes a gas void and illustrates the sequence of breakdown under sinusoidal alternating voltage [49]. If the electric field strength in the insulation becomes higher than the dielectric strength of the gas inside the void, the total breakdown will appear inside the void. In this moment, the voltage of Cc drops to zero, current pulse appears, then the discharge is extinguished. The process is repeated when the electric field strength in the insulation becomes again higher than the dielectric strength of the gas inside the void. This process appears at the zero crosses of Va and depends on the voltage gradient (around the peaks the voltage gradient tends to zero). The equations and formulas for calculating the apparent charge value and how to choose the capacitance of the coupling capacitor are detailed in [49,50].
Appendix A.2. Spread of Energy

Figure A2 plots the spread lines for the first and 99th of white noise samples. The groups of the dots from upper left to the lower right are the energy density as a function of the averaged period for IMFs 1–9 for all 1000 samples with an identical length of 1000 data points. The superimposed black dots are the energy density as a function of the averaged period for IMFs 2–9 for a single sample with $10^6$ data points. The bold blue dashed lines and the thin black dashed lines are the first and 99th percentiles calculated from Equations (3.6) and (3.7), respectively in [52].

As shown in Figure A2, there are still certain cases when the energy density of some IMFs is very near to its spread function. For example, a short distance below or beyond the bounds, as seen by the red circle. These spread lines can be adjusted by changing the confidence level, which is decided by signal characteristics and operator experience. During denoising for artificial signals, to enhance noise filtering capability, these boundaries are extended by 0.5 axial units.

Figure A2. The spread functions and the energy density of white noise used to determine the statistical significance of IMFs.

Figure A1. Partial discharge mechanism: (a) Scheme of an insulation system comprising a cavity; (b) sequence of cavity breakdown under alternating voltages [49,51].
Appendix A.3. Experimental PD Signal Diagram;

In the laboratory, a real power cable PD signal is used to demonstrate the ability to filter noise and get the Hilbert spectrum. PD measurement diagram is shown in Figure A3. The PD test circuit at the HV laboratory is comprised of the following components: a high voltage source, a PD signal acquisition device using HFCT sensor, and a defected power cable. The defect is generated by cutting the XLPE insulation of the cable to a depth of 3 mm. The voltage from the high voltage source would be increased until PD occurred, and PD signals appear on PD signal acquisition device. In this test, it cannot be ruled out that the PD source may originate from the power supply. However, the signal’s properties indicate that it is propagated in the cable insulation environment.

Test object is a medium voltage cable with cable type 12.7/22(24) kV—single core, nominal area is 1 × 70 mm², and dimension conductor is 9.7 mm.

Figure A3. Wiring diagram of the PD measurement on power cable used in this study.

References


