

Article

A Novel Hybrid Approach for Partial Discharge Signal Detection Based on Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and Approximate Entropy

Haikun Shang ^{1,*}, Yucai Li ¹, Junyan Xu ¹, Bing Qi ¹ and Jinliang Yin ²

¹ Key Laboratory of Modern Power System Simulation and Control and Renewable Energy Technology, Ministry of Education, Northeast Electric Power University, Jilin 132012, China; 2201900107@neepu.edu.cn (Y.L.); 2201900164@neepu.edu.cn (J.X.); qibing@neepu.edu.cn (B.Q.)

² School of Electrical and Electronic Engineering, Tianjin University of Technology, Tianjin 300384, China; searchforever@tjut.edu.cn

* Correspondence: shanghk@neepu.edu.cn; Tel.: +86-432-6480-6691

Received: 15 August 2020; Accepted: 15 September 2020; Published: 17 September 2020



Abstract: To eliminate the influence of white noise in partial discharge (PD) detection, we propose a novel method based on complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and approximate entropy (ApEn). By introducing adaptive noise into the decomposition process, CEEMDAN can effectively separate the original signal into different intrinsic mode functions (IMFs) with distinctive frequency scales. Afterward, the approximate entropy value of each IMF is calculated to eliminate noisy IMFs. Then, correlation coefficient analysis is employed to select useful IMFs that represent dominant PD features. Finally, real IMFs are extracted for PD signal reconstruction. On the basis of EEMD, CEEMDAN can further improve reconstruction accuracy and reduce iteration numbers to solve mode mixing problems. The results on both simulated and on-site PD signals show that the proposed method can be effectively employed for noise suppression and successfully extract PD pulses. The fusion algorithm combines the CEEMDAN algorithm and the ApEn algorithm with their respective advantages and has a better de-noising effect than EMD and EEMD.

Keywords: partial discharge; de-noising; complete ensemble empirical mode decomposition; approximate entropy; correlation coefficient analysis

1. Introduction

Insulation deterioration is one of the most critical faults in the power system. Partial discharge (PD) is an essential symptom of insulation deterioration. Effective PD detection plays an irreplaceable role in the evaluation of insulation conditions [1,2]. The electrical equipment's operating environment is very complex, often surrounded by various kinds of noise. PD signal is very weak, and on-site interference will cause significant difficulties in signal detection and extraction [3]. To extract real and reliable PD signals, some necessary de-noising methods are needed in the on-site environment [4].

Recently, various kinds of de-noising methods have been applied in PD signal extraction. Wavelet transform is suitable for processing non-stationary signal with better time-frequency resolution performance [5–7]. It has been widely researched in PD signal de-noising and achieved excellent application effectiveness [8–10]. Dai et al. [8] present a denoising method based on improved portogram and wavelet transform for UHF PD signals. Simulation and practical tests show that this novel method can effectively suppress the periodic narrowband noise and random white noise in PD signals and has

better performance than the traditional wavelet method. Ghorat et al. [9] propose a novel denoising algorithm for PD signals based on adaptive dual-tree complex wavelet transform. Simulation and experimental tests show that this new algorithm has superior performance in PD noise reduction in comparison with other methods. Sun et al. [10] employ a new signal processing method based on wavelet packet to detect random PD signals in power cables. The wavelet packet threshold method was used to filter out residual white noise in the reconstructed PD signal. The encouraging performance demonstrates the effectiveness of the proposed algorithm. However, owing to the manual selection of mother wavelets and decomposition levels in the above studies, the wavelet method has some inherent limitations in practical applications. Empirical mode decomposition (EMD) provides a fully automatic decomposition technique for non-stationary and nonlinear signals [11–13]. EMD decomposes the original signal into several different intrinsic mode functions (IMFs). EMD does not need a predefined basis function and decomposition level. It has been widely applied in PD signal analysis and achieved a good effect [14–16]. Chen et al. [14] propose a data compression denoising method based on EMD to extract real PD signals under large disturbances. Simulation and actual signal results demonstrate that the proposed algorithm does not lose the original signal energy and can achieve better performance than other methods. Zhang et al. [15] employ EMD for UHF PD signal denoising. The results show that, compared with the conventional wavelet method, this method achieves a higher signal-to-noise ratio and slighter waveform distortion. Wang et al. [16] present a PD diagnostic method based on neural network and EMD. The actual measurements demonstrate that the proposed method is applicable to power capacitor PD detection. However, EMD has an inherent limitation of mode mixing problems that limits its wider use. To deal with the above issues, ensemble empirical mode decomposition (EEMD) was proposed by Huang based on EMD [17–19]. White noise components are added manually in EEMD and eliminated through repetitive averaging. EEMD could effectively overcome the mode mixing in EMD by decomposing the original noise-corrupted signals into IMFs. Jin et al. [20] present a novel adaptive EEMD method for switchgear PD signal denoising. The simulation and experimental results show the effectiveness and superiority of the proposed method compared with the conventional wavelet algorithm and EMD-based method. Chan et al. [21] propose a self-adaptive technique for PD signal denoising with automatic threshold determination based on EEMD and mathematical morphology. The simulated and real PD results show that the proposed denoising algorithm is superior to wavelet transform and EMD-based PD de-noising methods. However, in the above research, the white noise added to the decomposing process of EEMD is not easily eliminated, which will cause reconstruction errors.

To improve the de-noising ability of EEMD, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) was introduced [22–25]. In CEEMDAN, adaptive noise is added into the signal decomposition. The residual noise in IMFs can be reduced by changing the decomposition process. This method can overcome the shortcomings of EEMD as well as improve the operational efficiency and reconstruction accuracy. Approximate entropy (ApEn) is a non-negative number that can measure the complexity of time series [26,27]. It represents the complexity degree in different scales and frequencies of signals. It is suitable for non-stationary signal analysis. The more complex the time series, the larger the ApEn value. ApEn has been studied in nonlinear analysis of various kinds of complex signals and achieved a good effect [28–30].

The proposed signal de-noising technique combines the advantages of CEEMDAN and ApEn. In this paper, CEEMDAN is utilized for PD signal decomposition to eliminate the mode mixing phenomenon in conventional EMD and improve the reconstruction accuracy in EEMD. ApEn of each IMF is then calculated to measure the complexity of PD signals. To select useful IMFs representing prominent PD features, correlation coefficient analysis (CCA) between IMFs and original signals is introduced [31]. Real IMFs can then be extracted according to ApEn and CCA for PD signal reconstruction. To verify the de-noising effect and superiority of the proposed method, EMD and EEMD are introduced into this paper. The simulation and on-site PD results demonstrate its effectiveness and superiority.

The rest of this paper is organized as follows. Sections 2 and 3 introduce the conception of CEEMDAN and approximate entropy. Section 4 presents the PD de-noising approach based on CEEMDAN-ApEn. Section 5 describes PD signal simulation and gives de-noising results with simulated signals. Section 6 evaluates the proposed method's performance with experimental and on-site signals and compares it with different de-noising methods. Section 7 concludes this paper.

2. Review of CEEMDAN

EEMD represents an extension of EMD. During EEMD decomposition, the noise added into the original signals cannot be eliminated, which may cause reconstruction errors [32]. On the basis of EEMD, CEEMDAN was proposed by Colominas et al. for further elimination of the mode mixing phenomenon. It is suitable for non-stationary signals thanks to its small iteration numbers and high convergence performance.

In CEEMDAN decomposition, adaptive white noise is introduced into each residual component [33]. The algorithm procedure will briefly be discussed in the following steps.

Step 1: Add a noise $w^i(t)$ to the original signal $x(t)$:

$$X^i(t) = x(t) + w^i(t) \quad (1)$$

in which $w^i(t) (i = 1, 2, \dots, N)$ meets the Gauss distribution and N is the number of samples.

Step 2: Decompose $X^i(t)$ using EMD into IMF F_1^i . By averaging F_1^i , the first component of IMF can be obtained in Formula (2).

$$\tilde{F}_1(t) = \frac{1}{N} \sum_{i=1}^N F_1^i \quad (2)$$

Step 3: Calculate the first residual component $r_1(t)$.

$$r_1(t) = x(t) - \tilde{F}_1(t) \quad (3)$$

Step 4: Calculate the second IMF.

$$\tilde{F}_2(t) = \frac{1}{N} \sum_{i=1}^N E_1(r_1(t) + \varepsilon_1 E_1[w^i(t)]) \quad (4)$$

where $E_j(\cdot)$ represents the j th IMF of the signal, and ε_j is the parameter of white noise power.

Step 5: Calculate the k th residual component as follows:

$$r_k(t) = r_{k-1}(t) - \tilde{F}_k(t) \quad (5)$$

where $k = 2, 3, \dots, K$ and K represents the highest order of IMF.

Step 6: Calculate the $(k + 1)$ th IMF component as follows:

$$\tilde{F}_{k+1}(t) = \frac{1}{N} \sum_{i=1}^N E_1(r_k(t) + \varepsilon_k E_k[w^i(t)]) \quad (6)$$

Step 7: Repeat Stage (5) and (6) until the residual component cannot be subdivided. The final residual component can be described as follows:

$$R(t) = x(t) - \sum_{k=1}^K \tilde{F}_k(t) \quad (7)$$

Step 8: The final signal can be decomposed as follows [22]:

$$x(t) = \sum_{k=1}^K \tilde{F}_k(t) + R(t) \quad (8)$$

3. Review of Approximate Entropy

Approximate entropy (ApEn) is a non-negative number used for the complexity measurement of time series [34]. It is widely used in the analysis of nonlinear characteristics of complex signals. The calculation of ApEn can be realized through the following steps.

Suppose the original time series $u(i), i = 0, 1, \dots, N$, in which N is the number of data sets.

Step 1: The time series $\{u(i)\}$ can be extended to m^{th} vector $X(i)$, which can be defined as follows:

$$X(i) = [u(i), u(i+1), \dots, u(i+m-1)] \quad (9)$$

where $x = 1, 2, \dots, N - m + 1$ and m is the pattern dimension.

Step 2: Calculate the distance between $X(i)$ and $X(j)$.

$$d[X(i), X(j)] = \max_{k=0,1,\dots,m-1} |u(i+k) - u(j+k)| \quad (10)$$

where $j = 1, 2, \dots, N - m + 1$.

Step 3: Calculate the ratio of $n(d)$ to $n(t)$, which is defined as $C_i^m(r)$.

$$C_i^m(r) = \frac{n(d)}{n(t)} \quad (11)$$

where $n(d)$ means the number of $d[X(i), X(j)] < r$, $n(t) = N - m + 1$ is the total number of vectors, and $r > 0$ is the preset threshold.

Step 4: Calculate the logarithm to $C_i^m(r)$ and the mean value can be obtained as follows:

$$\Phi^m(r) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \quad (12)$$

Step 5: The ApEn can be defined as follows [26]:

$$ApEn(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r) \quad (13)$$

in which $\Phi^{m+1}(r)$ can be obtained through Steps (1)–(4).

4. PD Signal De-Noise Based on CEEMDAN and ApEn

4.1. Algorithm Principle

Aimed at the non-stationary and non-linearity of PD signals, CEEMDAN is employed to analyze original signals. Compared with traditional EEMD, this approach has a higher extraordinary ability to discriminate different frequency parts. Firstly, CEEMDAN decomposes noisy PD signals into IMFs, which contain both noise parts and real signals. Secondly, the ApEn values of each IMF are calculated. These values can describe the irregularity and complexity of PD signals. The higher the complexity, the larger the value. According to the principle of ApEn, those IMFs that represent dominant signal features can be selected. Then, the correlation coefficient (CC) between each IMF and the original PD signal is calculated. CC values can effectively measure the similarity degree between two signals. After that, a threshold is set to eliminate those IMFs with low similarity. Finally, real IMFs are obtained for PD signal reconstruction.

4.2. Algorithm Procedure

1. Extract the original PD signals. Because of the complexity of power transformers' field environments, PD signals are always interfered with by various noises.
2. Decompose the original PD signal into different IMFs with CEEMDAN. These IMFs may consist of real PD information and complex noise interference.
3. Calculate ApEn values of IMFs extracted from CEEMDAN decomposition. ApEn represents the complexity of signals in different scales and frequencies. Owing to the non-stationary of original PD signals, the ApEn values may be different from each other.
4. Remove those IMFs that represent noise components. According to ApEn theory, noise and PD signals can be distinguished by different ApEn values. If the ApEn value is above a certain threshold, then the IMF is regarded as noise and abandoned. Otherwise, the IMF contains PD information and will be kept.
5. Calculate the correlation coefficient between the original PD signal and each selected IMF. As known, the CC value of the IMF that contains little PD information will be small. Therefore, real IMFs similar to the original signal will be selected as final parts through a certain threshold.
6. Reconstruct clean PD signal with final IMFs. These IMFs contain dominant PD features and show a strong correlation with original signals. This reconstruction method can efficiently recover the clean PD pulses from noisy signals.

The main de-noising procedure based on CEEMDAN and ApEn is shown in Figure 1.

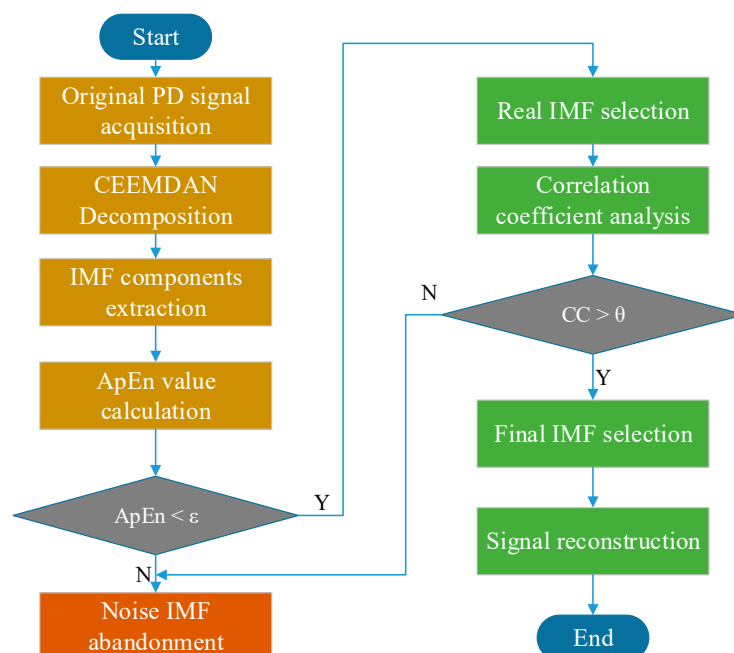


Figure 1. De-noising procedure based on complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and approximate entropy (ApEn). PD, partial discharge; IMF, intrinsic mode function; CC, correlation coefficient.

5. Simulation Analysis

5.1. Simulated Signal

A substantial amount of studies indicate that partial discharge is a non-stationary and high-frequency signal. High-frequency PD pulses can be expressed by mathematical models as follows [35]:

$$s_1(t) = Ae^{-t/\tau} \sin 2\pi f_c t \quad (14)$$

$$s_2(t) = A(e^{-1.3t/\tau} - e^{-2.2t/\tau}) \sin 2\pi f_c t \quad (15)$$

where A is the signal amplitude, τ is the attenuation coefficient, and f_c is the oscillation frequency.

The sampling frequency is 150 MHz and the number of samples is 1024. The sampling parameters of four different PD pulses are shown in Table 1. The simulated signal is shown in Figure 2a.

Table 1. Simulation parameters.

| | Amplitude/mV | Attenuation Coefficient/us | Oscillation Frequency/MHz |
|---------|--------------|----------------------------|---------------------------|
| Pulse 1 | 1.5 | 0.05 | 20 |
| Pulse 2 | 3 | 0.1 | 15 |
| Pulse 3 | 2 | 0.05 | 15 |
| Pulse 4 | 1 | 0.1 | 20 |

Owing to the poor operational environment of electrical equipment, the on-line detection of PD signals is usually influenced by environmental inference, mostly the white noise. The white noise is added to simulate the real PD signal, which satisfies Gauss distribution $N(0, 0.02^2)$. A simulated noisy signal is shown in Figure 2b. The spectrograms of signals are shown in Figure 2c–d.

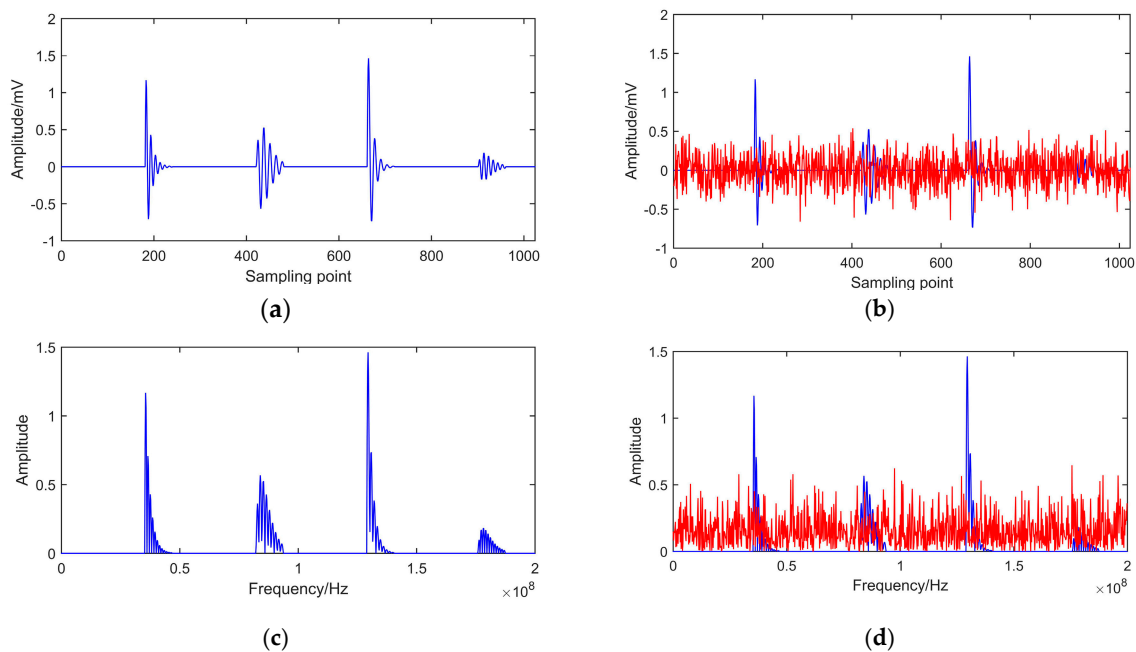


Figure 2. Simulated PD signal. (a) Original PD signal. (b) Noisy PD signal. (c) Spectrogram of original signal. (d) Spectrogram of noisy signal.

From Figure 2, it can be seen that the added noise considerably corrupts clean PD pulses. Four pulses are completely immersed in random noise. The clean signal cannot be recognized for the second and fourth pulse because of the small amplitude. To extract effective real PD signals, we consider in this paper a few specific de-noising methods.

5.2. Signal Decomposition

To verify the proposed algorithm's effectiveness and superiority, EMD and EEMD algorithms are employed to analyze the PD signal. Figures 3–5 show the signal decomposition using different methods, including EMD, EEMD, and CEEMDAN. For EEMD and CEEMDAN, 100 groups of Gaussian white noise with standard deviation of 0.2 were added into the original signal.

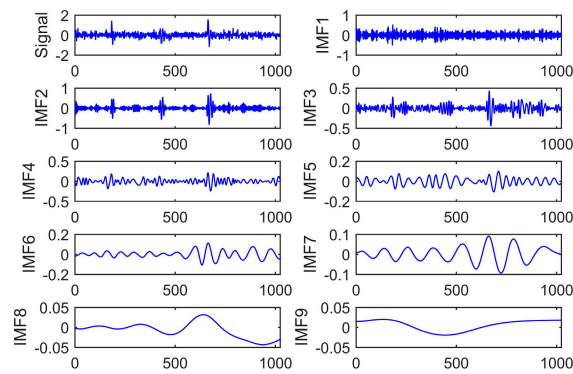


Figure 3. PD signal decomposition based on EMD.

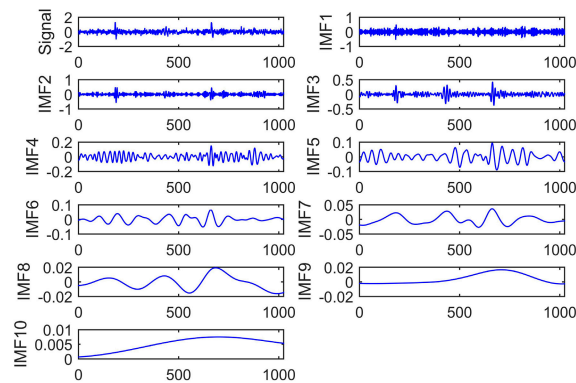


Figure 4. PD signal decomposition based on EEMD.

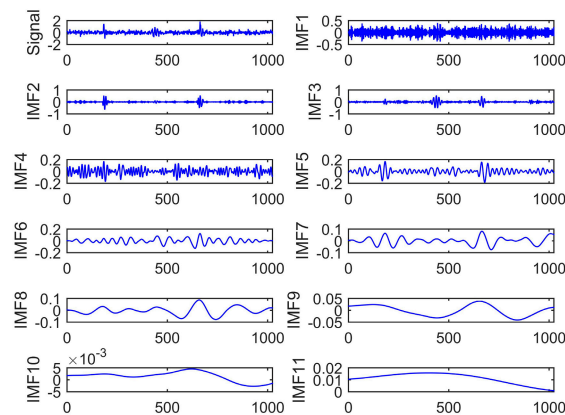


Figure 5. PD signal decomposition based on CEEMDAN.

It can be seen from Figure 3 that eight IMF components and a single residual component were obtained through EMD decomposition. IMF1 has the highest component frequency. The nature of signals cannot be analyzed accurately. Besides, the substantial similarity between IMF6 and IMF7 indicates that the mode mixing phenomenon exists in EMD decomposition.

The noisy PD signal decomposition result based on EEMD is presented in Figure 4. The standard deviation of white Gaussian noise is 0.2, and the repetitive number is 200. Figure 4 shows that the number of IMFs decomposed by EEMD is ten, which is greater than that of EMD. This means more details of signals can be found out through EEMD. The white noise makes each IMF maintain the continuity in the time domain. The EEMD decomposition method could obtain frequency components of the original PD signal. However, IMF 1–3 show that a particular oscillation phenomenon occurs during signal component analysis. This suggests that the white noise added in EEMD causes unfavorable

influence on signal decomposition. Some necessary steps need to be taken to control noise’s impact to ensure the accuracy of detection.

Figure 5 presents the IMFs’ components decomposed by CEEMDAN. It is clear that ten IMF components and a single residual component were obtained. This decomposition method makes it more uniform in the distribution of the IMFs. Besides, the frequency changes between different IMF components have become more apparent. The detailed decomposition helps CEEMDAN solve the problems of mode mixing even further.

The boxplots of computation numbers of IMFs using different decomposition methods are shown in Figure 6. It shows that, for each IMF selection, the number of calculations with CEEMDAN is smaller than that of EEMD. This means the CEEMDAN method can reduce the operation time and improve the efficiency of signal decomposition.

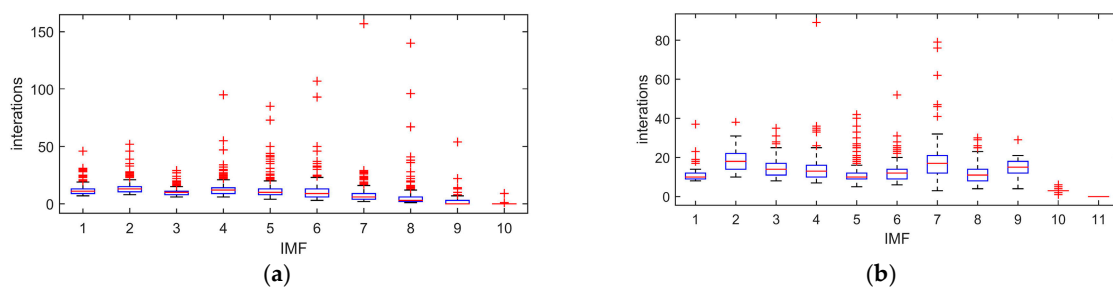


Figure 6. Iterations of each IMF. (a) EEMD. (b) CEEMDAN.

5.3. Approximate Entropy Calculation

It is shown from Section 5.2 that eleven IMFs are obtained by CEEMDAN decomposition. However, the number of IMF components of PD signals may vary with different trials. Therefore, it is inaccurate to estimate the boundary between noise and PD signals by subjective judgment. To analyze the complexity of PD signals and reduce the reconstruction error, ApEn is introduced to distinguish between noise and real PD pulses. Before calculation, the pattern dimension m and similar tolerance boundary r should be predefined. On the basis of the researchers’ previous study experience [36], the parameters are defined as follows.

$$\begin{cases} m = 2 \\ r = 0.2E_{SD} \end{cases} \tag{16}$$

where E_{SD} is the standard deviation of original signals.

The value of approach entropy of each IMF is calculated. Through multiple calculations, the mean value of ApEn is shown in Figure 7.

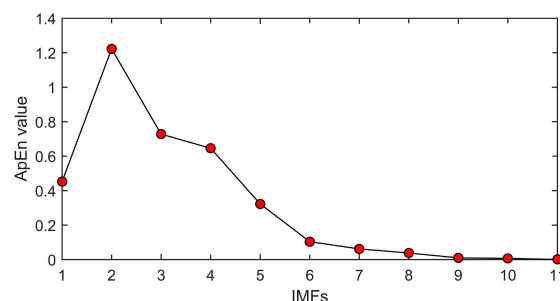


Figure 7. ApEn values of IMFs.

Figure 7 shows that different IMFs possess different ApEn values, which means diverse complexity exists in the decomposition levels. It can be concluded that, from A2 to A11, the ApEn values decrease gradually. This illustrates that the complexity of each IMF is gradually reduced. According to the

theory of ApEn, more irregular time series obtain greater ApEn values. Owing to the randomness and irregularity of white noise, ApEn values of noisy IMFs will be greater than those of clean ones. In this paper, a threshold ε is set to 0.5 to eliminate noise IMFs. If the ApEn value is greater than ε , the IMF will be abandoned as a noisy part. Otherwise, the IMF will be kept as a clean part. After a comparison of different ApEn values, IMF 1 and 5–11 are retained.

5.4. Correlation Coefficient Analysis

Through Section 5.3, those IMFs that represent noise parts are effectively eliminated. However, some over-decomposition phenomenon may occur during CEEMDAN, which can produce pseudo-component in IMFs. It will be even worse when the signal gets more complicated. This will also cause reconstruction errors. In this paper, to extract useful and effective IMF components, correlation coefficient analysis is employed for IMF selection. The CC is defined as follows [37].

$$CC = \frac{\sum_{i=1}^k (x_i - \bar{x})(imf_i - \overline{imf})}{\sqrt{\sum_{i=1}^k (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^k (imf_i - \overline{imf})^2}} \quad (17)$$

where x is the original signal, \bar{x} is the mean value of x , imf is the IMF component, \overline{imf} is the mean value of IMF, and k is the number of IMF components.

Firstly, the CC value of each IMF is calculated using Formula (17). Eight CC values are obtained shown in Table 2. The CC value can effectively quantify the similarity between two different time series. Table 2 shows that different IMFs have a certain degree of similarity with the original PD signal.

Table 2. Correlation coefficient (CC) values. IMF, IMF, intrinsic mode function.

| IMF1 | IMF5 | IMF6 | IMF7 | IMF8 | IMF9 | IMF10 | IMF11 |
|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.552 | 0.823 | 0.906 | 0.932 | 0.773 | 0.521 | 0.328 | 0.195 |

To eliminate those IMFs that have low similarity with original PD signals, a threshold θ is preset. If the CC value is greater than θ , the IMF will be kept as an effective component. Otherwise, the IMF will be abandoned as a useless part. In this paper, the threshold is defined as follows.

$$\theta = \sqrt{\frac{\sum_{i=1}^k (CC_i - \overline{CC})^2}{k}} \quad (18)$$

After multiple trials, θ is set to 0.6. According to CC's principle, IMF 1, IMF 9, IMF 10, and IMF 11 are removed as a result of CC's small value. Those IMFs that have less similarity with original PD signals are abandoned. The remaining IMFs have a strong correlation with original signals, which means they possess prominent information on PD features.

5.5. De-Noising Results Analysis

To compare the performance of different de-noising methods, EMD, EEMD, and CEEMDAN-ApEn are employed to de-noise the simulated PD signal. Three evaluation indexes are used for quantitative analysis of the quality of signal de-noising, including the signal to noise ratio (SNR), mean square error (MSE), and normalized correlation coefficient (NCC). The higher the SNR and NCC, the more effective

the de-noising result. The smaller the MSE, the more similar the original and the de-noised signal. SNR, MSE, and NCC are defined as follows:

$$SNR = 10 \log \frac{\sum_{i=1}^n s^2(i)}{\sum_{i=1}^n (s(i) - \hat{s}(i))^2} \tag{19}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (|s(i) - \hat{s}(i)|^2) \tag{20}$$

$$NCC = \frac{\sum_{i=1}^n s \cdot \hat{s}}{\sqrt{\sum_{i=1}^n s^2 \sum_{i=1}^n \hat{s}^2}} \tag{21}$$

where $s(i)$ is the original PD signal and $\hat{s}(i)$ is the de-noised signal.

De-noising results with different methods are shown in Figure 8a–f.

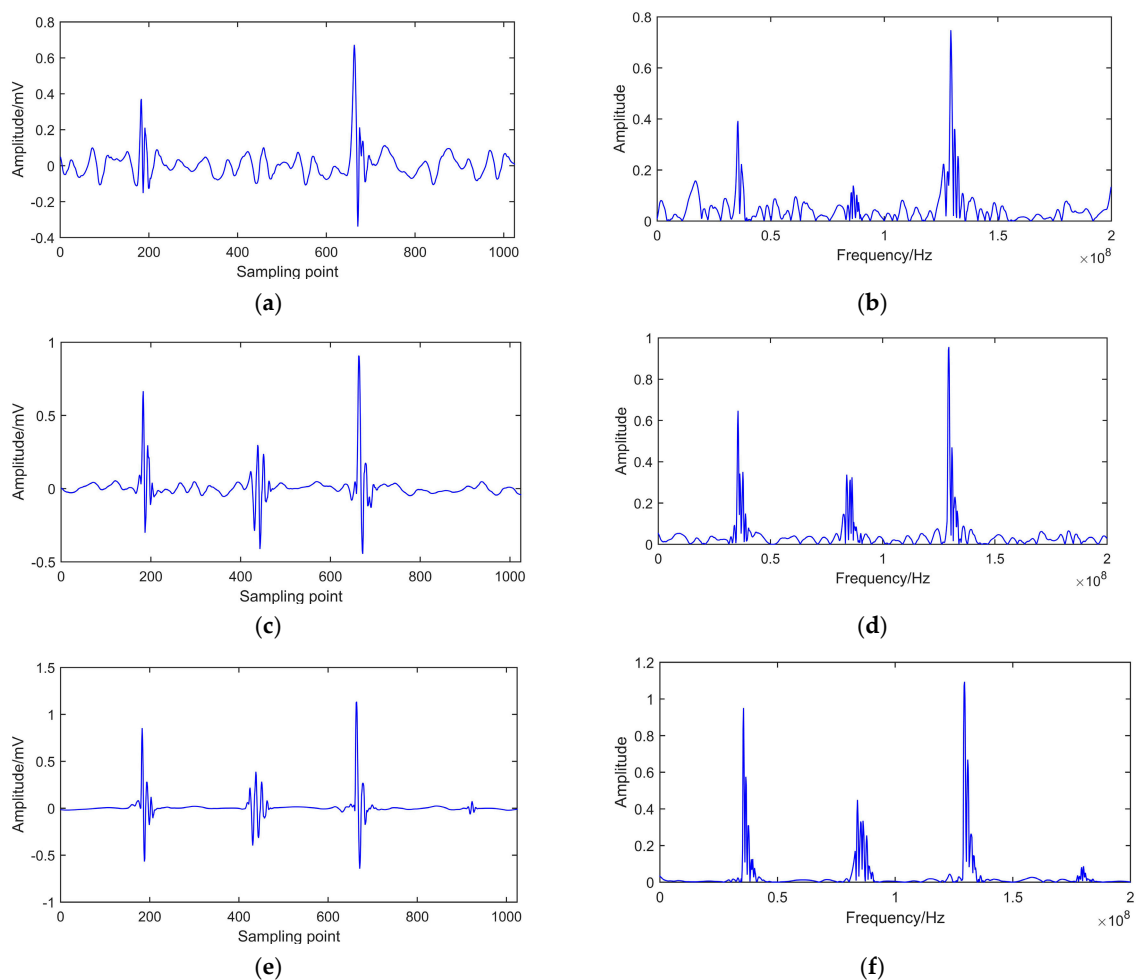


Figure 8. De-noised PD signal in time and frequency domain. (a) EMD de-noising. (b) Spectrum of EMD de-noising. (c) EEMD de-noising. (d) Spectrum of EEMD de-noising. (e) CEEMDAN-ApEn de-noising. (f) Spectrum of CEEMDAN-ApEn de-noising.

Figure 8a,b shows that the fourth pulse, which has a smaller amplitude, cannot be detected after EMD de-noising. Furthermore, there is an apparent oscillating phenomenon after de-noising. Figure 8c,d indicates that the oscillating phenomenon can be restrained to a certain extent after EEMD. However, the fourth pulse still cannot be detected. It can be seen from Figure 8e,f that all of the PD pulses could be detected, and a smooth signal is achieved using the CEEMDAN-ApEn method.

It can be concluded from the above results that the oscillating phenomenon was produced as a result of mode mixing after EMD decomposition. The unidentified PD pulse was caused by considering detailed information in the IMF during the de-noising process. Part of critical information was lost after EMD. Adding the white Gaussian noise during EEMD could ensure the continuity of each IMF in the time domain. It can eliminate the mode mixing phenomenon to a certain extent. However, because of the influence of added white noise during EEMD decomposition, some oscillation distortion still cannot be suppressed entirely. Consequently, the fourth pulse still cannot be detected clearly, and some useful frequency spectral components are missing. In CEEMDAN-ApEn, adaptive white noise was added into each residual component, which leads to complete decomposition. It can solve the problem of mode mixing further. ApEn can measure the non-linearity and complexity of time series. ApEn values can easily recognize noisy parts in IMF. Combined with CEEMDAN and ApEn, useful IMFs can be selected for PD signal reconstruction. After CEEMDAN-ApEn, all PD pulses were successfully identified. Moreover, the de-noised signal was very smooth, and clear frequency components were also obtained.

SNR, MSE, and NCC results with different de-noising methods are shown in Table 3.

Table 3. De-noising index comparison. CEEMDAN, complete ensemble empirical mode decomposition with adaptive noise; ApEn, approximate entropy; SNR, signal to noise ratio; MSE, mean square error; NCC, normalized correlation coefficient.

| | SNR/dB | MSE/dB | NCC |
|-----------------|---------|--------|-------|
| Original Signal | -5.3024 | 0.9356 | 1 |
| EMD | 15.8726 | 0.4067 | 0.772 |
| EEMD | 21.6218 | 0.1068 | 0.985 |
| CEEMDAN-ApEn | 28.2298 | 0.0326 | 0.993 |

It can be seen from Table 3 that the SNR and NCC are the largest, and the MSE is the smallest using the CEEMDAN-ApEn method, which means the de-noising effect is the best. Compared with EMD and EEMD, CEEMDAN-ApEn can eliminate the mode mixing phenomenon in EMD and reduce the influence of added noise in EEMD decomposition. The de-noised signal could well represent the characteristics of the original PD signal.

To verify the effectiveness of the de-noising method in different noise environments, PD simulation results under different SNR conditions are discussed in detail. Figure 9 presents the results of different de-noising algorithms varying with diverse SNR.

Figure 9 shows that the three algorithms' de-noising performance is different from the increase of SNR. It can be seen from Figure 9a that the NCC value of CEEMDAN-ApEn is larger in each SNR than that of EMD and EEMD. This means that the proposed method has the highest similarity between the reconstructed and the original signal. Moreover, the performance of CEEMDAN-ApEn is relatively stable. Figure 9b illustrates that, compared with EMD and EEMD, CEEMDAN-ApEn obtained smaller MSE. It indicates that the de-noising effect of the proposed method is better than the other two traditional ones. It also shows the stable performance varying with SNR.

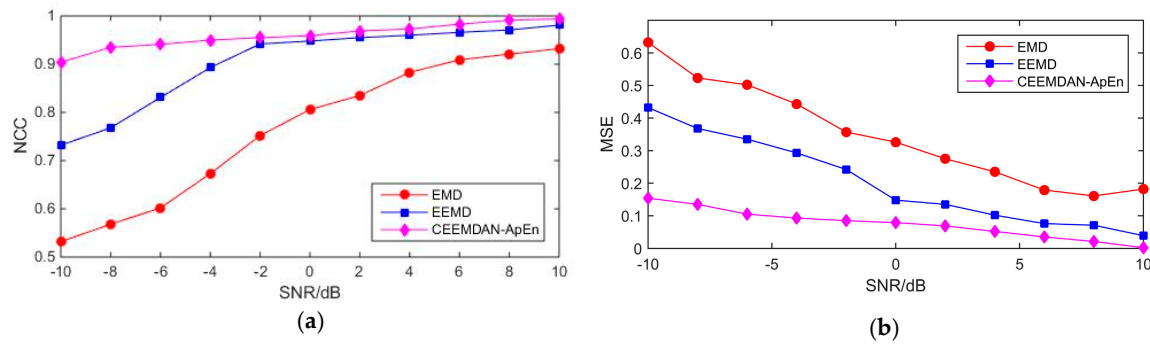


Figure 9. De-noised results of different methods. SNR, signal to noise ratio; MSE, mean square error; NCC, normalized correlation coefficient. (a) NCC vary with different SNR. (b) MSE vary with different SNR.

6. Experimental and On-Site PD Signal Analysis

6.1. Experimental PD Signal

Figure 10 shows the circuit for PD simulation experiments. To verify the proposed de-noising method’s effectiveness, two PD signals are extracted under the noisy experimental environment. Measured signals are shown in Figure 11. It indicates that PD signals are disturbed by observed interference, and real PD pulses cannot be detected effectively. To remove the experimental interference mixed in PD pulses, certain denoising measures are taken as follows. Because of the randomness and unknown characteristics of noises, in this paper, the noise rejection ratio (NRR) is utilized to measure de-noising quality [38]. NRR is defined as follows:

$$NRR = 10(\lg\sigma_1^2 - \lg\sigma_2^2) \tag{22}$$

where σ_1 , σ_2 represent the noise deviation of pre-treatment and post-treatment, respectively. The deviation can be defined as follows:

$$\sigma = \sqrt{\frac{1}{Q} \sum_{d=1}^Q (S_d - \mu)^2} \tag{23}$$

where Q is the number of samples, S_d represents the d th sampling signal, and μ is the mean of signal.

NRR reflects the prominent level of de-noised signal. The signals de-noised by EEMD and CEEMDAN-ApEn are shown in Figures 12 and 13. NRR results are shown in Table 4.

Figure 12 presents the time and frequency domain of de-noised results by EEMD. It shows that, after EEMD denoising, PD pulses could be detected. However, noticeable oscillation distortion exists in the de-noised signal. It can be seen from Figure 13 that all PD pulses were effectively extracted. Moreover, the oscillation almost disappeared after denoising. This means CEEMDAN-ApEn effectively reduces the added noise’s influence and improves the signal’s smoothness. Table 4 shows that a higher NRR is obtained after CEEMDAN-ApEn compared with EEMD. Above all, CEEMDAN-ApEn has an obvious advantage over EEMD on PD de-noising.

Table 4. Noise rejection ratio (NRR) comparison of de-noised PD signal.

| | NRR/dB |
|--------------|---------|
| EEMD | 16.5825 |
| CEEMDAN-ApEn | 28.3012 |

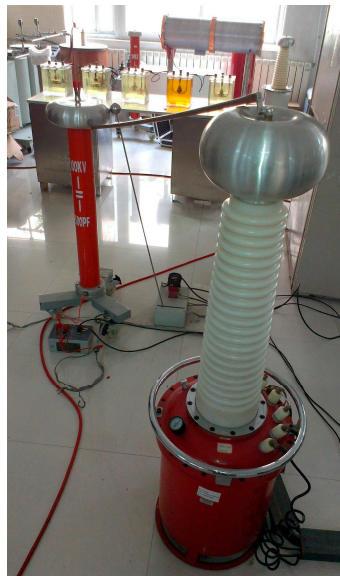
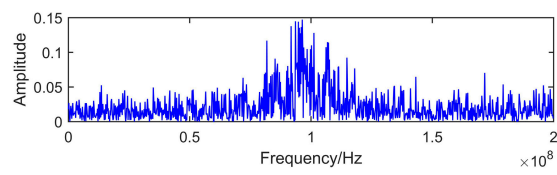
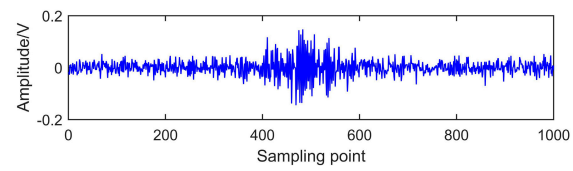
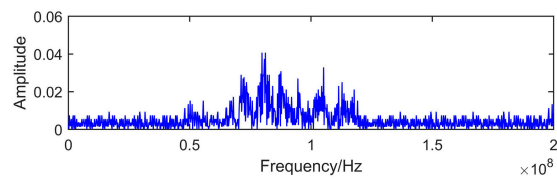
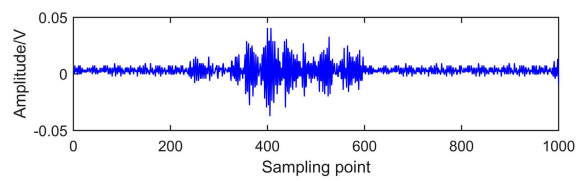


Figure 10. Experimental setup.



(a)



(b)

Figure 11. Measured PD signals in laboratory. (a) Measured signal 1. (b) Measured signal 2.

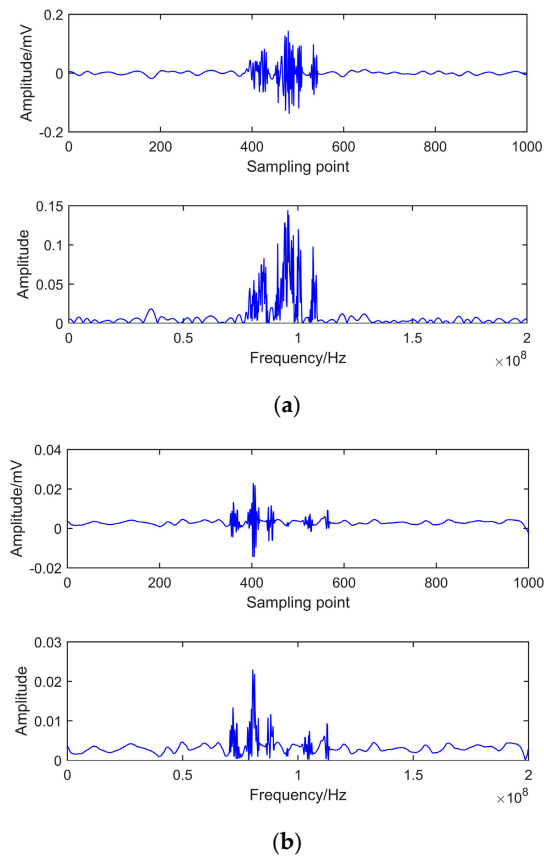


Figure 12. De-noising with EEMD. (a) Measured signal 1. (b) Measured signal 2.

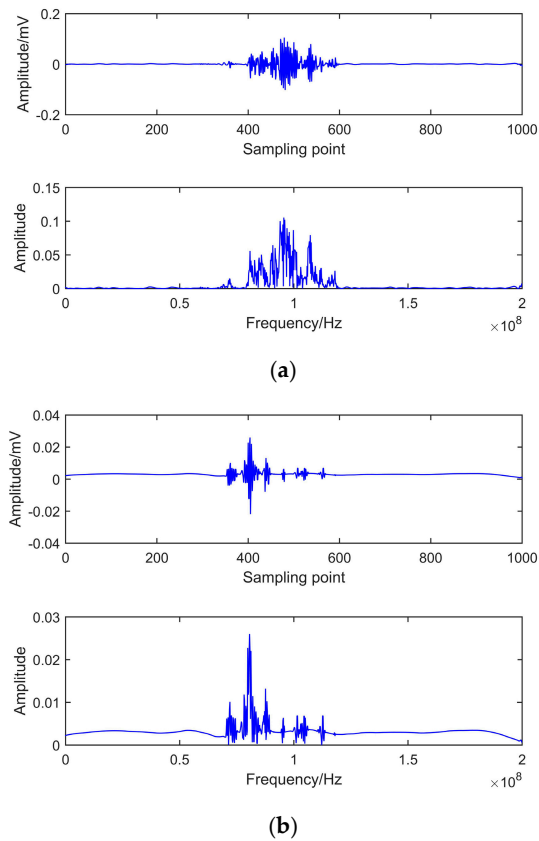


Figure 13. De-noising of CEEMDAN-ApEn. (a) Measured signal 1. (b) Measured signal 2.

6.2. On-Site PD Signal

Figure 14a presents the on-site PD signal from the on-line monitoring system of one substation in Anhui Province. Because of the complex practical environment, PD signals are disturbed by environmental interference, and real PD signals cannot be detected effectively. To eliminate the environmental disturbance and extract effective PD pulses, CEEMDAN-ApEn is employed for signal analysis. The de-noised signal is shown in Figure 14b. It shows that PD pulses can be effectively extracted after de-noising using the proposed method.

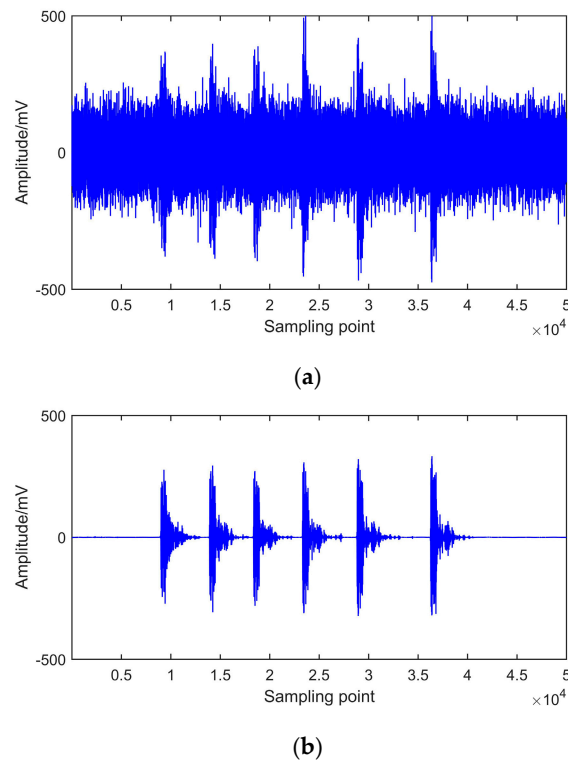


Figure 14. Analysis of on-site PD signal. (a) Field PD signal. (b) CEEMDAN-ApEn de-noising.

7. Conclusions

Noise interference is a big problem in PD signal extraction. This paper presents an approach based on CEEMDAN-ApEn for PD de-noising. CEEMDAN can effectively separate the original signal into different intrinsic mode functions. ApEn of each IMF is calculated to measure the complexity of the signal. Those IMFs that contain noise components can be removed by ApEn analysis. Correlation coefficient analysis is then employed between IMFs and the original PD signal. Finally, effective IMFs are obtained for signal reconstruction by analyzing the correlation coefficient.

To verify the proposed approach's effectiveness and superiority, EMD and EEMD are applied for PD signal analysis in this paper. EMD can decompose signals into several IMFs and restrain the noise to a certain extent. However, mode mixing and oscillation distortion restricted its application. EEMD can eliminate the mode mixing phenomenon by adding manual white noise into signal decomposition. However, the added noise cannot be easily eliminated, and the residual noise still exists, which may cause signal reconstruction errors. In CEEMDAN, adaptive white noise is added into signal decomposition, and its unique decomposition procedure realizes residual noise isolation. ApEn can measure the complexity and non-stationarity of the PD signal. Using the ApEn analysis of decomposed IMFs, those noisy components can be effectively removed from PD signals. Useful components that represent prominent PD features can be further selected for signal reconstruction according to the correlation coefficient analysis.

Simulated and on-site PD signals indicate that CEEMDAN can successfully eliminate the mode mixing problem and improve the reconstruction efficiency. The ApEn extracted from IMFs can represent the prominent components of PD features. Combining the advantages of CEEMDAN and ApEn, the proposed de-noising method can effectively suppress the inference in PD detection. Compared with EMD and EEMD, CEEMDAN-ApEn has a better de-noising effect through different denoising index analysis. It can effectively extract the weak PD pulse from strong background noise. In brief, the proposed approach has a good signal denoising effect and provides a new practical tool for PD signal detection.

It is noteworthy that the operation environment of field equipment is quite complex. PD signals extracted from sensors are usually contaminated by various kinds of noise, such as white noise, periodic narrowband noise, impulse noise, and so on. Moreover, different types of noise may cause various difficulties in real PD extraction. For further consideration, different types of noise could be added into PD simulation to verify the effectiveness of the proposed algorithm. Further, more complex on-site PD signals could be collected for deeper analysis.

Author Contributions: Conceptualization, H.S.; Data curation, Y.L.; Software, J.Y.; Validation, J.X.; Writing—original draft, H.S.; Writing—review & editing, B.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Science and Technology Project of the State Grid Corporation of China (SGLNDK00KJJS1500008) and the Doctoral Scientific Research Foundation of Northeast Electric Power University (No. BSJXM-201406), China.

Conflicts of Interest: The author declares that there is no conflict of interest regarding the publication of this paper.

References

- Alexander, P.; Uwe, S. Partial discharge measurement at DC voltage—Evaluation and characterization by NoDi* pattern. *IEEE Trans. Dielectr. Electr. Insul.* **2018**, *25*, 883–891.
- Wang, J.Y.; Liang, H.J.; Kang, A.M. Researches of Power Transformer Partial Discharge Location Using Ultrasonic Method in China. *J. Northeast Dianli Univ.* **2019**, *39*, 10–14.
- Darwish, A.; Refaat, S.S.; Toliyat, H.A.; Abu-Rub, H. On the Electromagnetic Wave Behavior Due to Partial Discharge in Gas Insulated Switchgears: State-of-Art Review. *IEEE Access* **2019**, *7*, 75822–75836. [[CrossRef](#)]
- Torii, H.; Hayase, Y.; Yamashiro, K.; Matsumoto, S. Noise discrimination method for partial discharge current focused on damped oscillation waveform. *IEEE Trans. Fundam. Mater.* **2018**, *138*, 64–70. [[CrossRef](#)]
- Daubechies, I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. Theory* **1990**, *36*, 961–1005. [[CrossRef](#)]
- Lu, W.; Lu, X.; Han, J.; Zhao, Z.; Du, X. Online Estimation of ESR for DC-Link Capacitor of Boost PFC Converter Using Wavelet Transform Based Time–Frequency Analysis Method. *IEEE Trans. Power Electron.* **2020**, *35*, 7755–7764. [[CrossRef](#)]
- Liaw, J.-J.; Lu, C.-P.; Huang, Y.-F.; Liao, Y.-H.; Huang, S.-C. Improving Census Transform by High-Pass with Haar Wavelet Transform and Edge Detection. *Sensors* **2020**, *20*, 2537. [[CrossRef](#)]
- Dangdang, D.A.I.; Xianpei, W.A.N.G.; Jiachuan, L.O.N.G. De-noising Method of Ultra-high Frequency Partial Discharge Signal Based on Improved Protrugram and Wavelet Transform. *High Volt. Eng.* **2018**, *44*, 3577–3586.
- Ghorat, M.; Gharehpetian, G.B.; Latifi, H.; Hejazi, M.A. A New Partial Discharge Signal Denoising Algorithm Based on Adaptive Dual-Tree Complex Wavelet Transform. *IEEE Trans. Instrum. Meas.* **2018**, *67*, 2262–2272. [[CrossRef](#)]
- Sun, K.; Zhang, J.; Shi, W.; Guo, J. Extraction of Partial Discharge Pulses from the Complex Noisy Signals of Power Cables Based on CEEMDAN and Wavelet Packet. *Energies* **2019**, *12*, 3242. [[CrossRef](#)]
- Boudraa, A.O.; Cexus, J.C. EMD-Based Signal Filtering. *IEEE Trans. Instrum. Meas.* **2007**, *56*, 2196–2202. [[CrossRef](#)]
- Ziani, R.; Hammami, A.; Chaari, F.; Felkaoui, A.; Haddar, M. Gear fault diagnosis under non-stationary operating mode based on EMD, TKEO, and Shock Detector. *Comptes Rendus Mécanique* **2019**, *347*, 663–675. [[CrossRef](#)]

13. Das, K.; Nath, D.; Pradhan, S. FPGA and ASIC Realization of EMD Algorithm for Real-Time Signal Processing. *IET Circuits Devices Syst.* **2019**. [[CrossRef](#)]
14. Chen, X.J.; Yang, Y. Analysis of the partial discharge of ultrasonic signals in large motor based on Hilbert-Huang transform. *Appl. Acoust.* **2018**, *131*, 165–173. [[CrossRef](#)]
15. Zhang, Y.; Wang, X.; Guo, L.; Zhang, S. Partial Discharge Detection for GIS Based on Normalized Autocorrelation Function and Similar Wavelet Soft Threshold. *High Volt. Apparatus* **2018**, *54*, 17–24.
16. Wang, M.-H.; Lu, S.-D.; Lin, K.-D.; Pan, C.-Y. Application of Extension Neural Network Algorithm and Empirical Mode Decomposition Method to Partial Discharge Diagnosis of Power Capacitors. *Sens. Mater.* **2019**, *31*, 3959–3972. [[CrossRef](#)]
17. Wang, T.; Zhang, M.; Yu, Q.; Zhang, H. Comparing the applications of EMD and EEMD on time–frequency analysis of seismic signal. *J. Appl. Geophys.* **2012**, *83*, 29–34. [[CrossRef](#)]
18. Tan, Q.-F.; Lei, X.; Wang, X.; Wang, H.; Wen, X.; Ji, Y.; Kang, A.-Q. An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach. *J. Hydrol.* **2018**, *567*, 767–780. [[CrossRef](#)]
19. Yu, Y.H.; Zhang, H.; Singh, V. Forward Prediction of Runoff Data in Data-Scarce Basins with an Improved Ensemble Empirical Mode Decomposition (EEMD) Model. *Water* **2018**, *10*, 388. [[CrossRef](#)]
20. Jin, T.; Li, Q.; Mohamed, M.A. A Novel Adaptive EEMD Method for Switchgear Partial Discharge Signal Denoising. *IEEE Access* **2019**, *7*, 58139–58147. [[CrossRef](#)]
21. Chan, J.C.; Ma, H.; Saha, T.K.; Ekanayake, C. Self-adaptive partial discharge signal de-noising based on ensemble empirical mode decomposition and automatic morphological thresholding. *IEEE Trans. Dielectr. Electr. Insul.* **2014**, *21*, 294–303. [[CrossRef](#)]
22. Torres, M.; Colominas, M.A.; Schlotthauer, G.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; pp. 4144–4147.
23. Zhang, W.; Qu, Z.; Zhang, K.; Mao, W.; Ma, Y.; Fan, X. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. *Energy Convers. Manag.* **2017**, *136*, 439–451. [[CrossRef](#)]
24. Kuai, M.; Cheng, G.; Pang, Y.; Li, Y. Research of Planetary Gear Fault Diagnosis Based on Permutation Entropy of CEEMDAN and ANFIS. *Sensors* **2018**, *18*, 782. [[CrossRef](#)] [[PubMed](#)]
25. Mumtaz, A.; Prasad, R. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. *Renew. Sustain. Energy Rev.* **2019**, *104*, 281–295.
26. Pincus, S. Approximate entropy (ApEn) as a complexity measure. *Chaos* **1995**, *5*, 110–117. [[CrossRef](#)]
27. Xiao, M.-X.; Wei, H.-C.; Xu, Y.-J.; Wu, H.-T.; Sun, C.-K. Combination of R-R Interval and Crest Time in Assessing Complexity Using Multiscale Cross-Approximate Entropy in Normal and Diabetic Subjects. *Entropy* **2018**, *20*, 497. [[CrossRef](#)]
28. Singh, V.; Gupta, A.; Sohal, J.S.; Singh, A. A unified non-linear approach based on recurrence quantification analysis and approximate entropy: Application to the classification of heart rate variability of age-stratified subjects. *Med. Biol. Eng.* **2018**, *57*, 741–755. [[CrossRef](#)]
29. Ryan, G.; Mosca, A.; Chang, J.-S.; Wu, E. At a Glance: Pixel Approximate Entropy as a Measure of Line Chart Complexity. *IEEE Trans. Vis. Comput. Graph.* **2018**, *25*, 872–881. [[CrossRef](#)]
30. Chen, W.; Jiang, F.; Chen, X.; Feng, Y.; Miao, J.; Chen, S.; Jiao, C.; Chen, H. Photoplethysmography-derived approximate entropy and sample entropy as measures of analgesia depth during propofol-remifentanyl anesthesia. *Int. J. Clin. Monit. Comput.* **2020**, *10*, 1007. [[CrossRef](#)]
31. Nur, B.A.; No-Weon, K.; Jae-Yong, K. Field Uniformity and Correlation Coefficient Analysis of KRIS Reverberation Chamber. *IEEE Trans. Commun.* **2018**, *101*, 2289–2296.
32. Li, S.; Qin, N.; Huang, D.; Huang, D.; Ke, L. Damage Localization of Stacker’s Track Based on EEMD-EMD and DBSCAN Cluster Algorithms. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 1981–1992. [[CrossRef](#)]
33. Bouhalais, M.L.; Djebala, A.; Ouelaa, N.; Babouri, M.K. CEEMDAN and OWMRA as a hybrid method for rolling bearing fault diagnosis under variable speed. *Int. J. Adv. Manuf. Technol.* **2017**, *94*, 2475–2489. [[CrossRef](#)]
34. Kumar, S.; Danas, K.; Kochmann, D.M. Enhanced local maximum-entropy approximation for stable meshfree simulations. *Comput. Methods Appl. Mech. Eng.* **2019**, *344*, 858–886. [[CrossRef](#)]

35. Xie, M.; Zhou, K.; Huang, Y.L.; He, M.; Wang, X.J. A White Noise Suppression Method for Partial Discharge Based on Short Time Singular Value Decomposition. *Proc. CSEE* **2019**, *3*, 915–922.
36. Fu, L.; He, Z.Y.; Mai, R.K.; Qian, Q.Q. Application of Approximate Entropy to Fault Signal Analysis in Electric Power System. *Proc. CSEE* **2018**, *28*, 68–73.
37. Jiang, F.; Zhu, Z.; Li, W.; Ren, Y.; Zhou, G.; Chang, Y. A Fusion Feature Extraction Method Using EEMD and Correlation Coefficient Analysis for Bearing Fault Diagnosis. *Appl. Sci.* **2018**, *8*, 1621. [[CrossRef](#)]
38. Shang, H.K.; Lo, K.; Li, F. Partial Discharge Feature Extraction Based on Ensemble Empirical Mode Decomposition and Sample Entropy. *Entropy* **2017**, *19*, 439. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).