Application of Shannon Wavelet Entropy and Shannon Wavelet Packet Entropy in Analysis of Power System Transient Signals

Jikai Chen *, Yanhui Dou, Yang Li and Jiang Li

School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China; fundouyh@163.com (Y.D.); liyang@nedu.edu.cn (Y.L.); lijiangdyx@163.com (J.L.)

* Correspondence: chenjikai1977@163.com; Tel.: +86-432-6480-6691

Abstract: In a power system, the analysis of transient signals is the theoretical basis of fault diagnosis and transient protection theory. Shannon wavelet entropy (SWE) and Shannon wavelet packet entropy (SWPE) are powerful mathematics tools for transient signal analysis. Combined with the recent achievements regarding SWE and SWPE, their applications are summarized in feature extraction of transient signals and transient fault recognition. For wavelet aliasing at adjacent scale of wavelet decomposition, the impact of wavelet aliasing is analyzed for feature extraction accuracy of SWE and SWPE, and their differences are compared. Meanwhile, the analyses mentioned are verified by partial discharge (PD) feature extraction of power cable. Finally, some new ideas and further researches are proposed in the wavelet entropy mechanism, operation speed and how to overcome wavelet aliasing.

Keywords: Shannon wavelet entropy; Shannon wavelet packet entropy; power system; transient power signals; wavelet aliasing; accuracy of the feature extraction

1. Introduction

The identifying and processing of transient signals are the basis for operation monitoring, fault diagnosis and power quality analysis in a power system, which also provide technical support for rely protection [1–3]. When faults occur on the high-voltage transmission lines and electrical equipment, voltage and current involve a large number of transient components of non-fundamental frequency and might change with fault location, fault resistance etc. Transient signals caused by faults are non-stationary random process, which involve transient overvoltage, voltage dips, voltage interruptions and voltage pulse [4]. During the analysis of transient signals, transient feature would be swamped by systematic noise because of its low energy and magnitude [5]. Therefore, it is key to extract transient features exactly from power signals. At the same time, it is difficult to identify transient disturbance efficiently because the quantity of transient feature is large, redundancy information exists and dimension is not unified [6]. The traditional feature extraction methods are based on the digital filter of the Fourier transform. However, the Fourier transform is invalid in analysis of non-stationary fault signals [7]. With the development of wavelet theory and entropy statistical theory, thanks to the fast development of high-speed A/D and DSP technology, SWE and SWPE are widely applied to the analysis and processing of transient signals in a power system, and their advantages have been unfolded. On the one hand, wavelet theory and wavelet packet theory are gradually improved and developed, and researches of algorithm structure and fast algorithm are deepened [8,9]. On the other hand, based on the advantages of wavelet theory and wavelet packet theory, they combine with entropy theory, which is developed toward practical direction [10–12].

From the recent achievements about SWE and SWPE in analysis of transient signals in a power system, their applications are summarized in feature extraction of transient signals and transient fault recognition. According to wavelet aliasing at adjacent scale of wavelet decomposition, the impact is analyzed on feature extraction accuracy of SWE and SWPE, and their accuracy of are compared in feature extraction. Meanwhile, the analyses mentioned are verified by PD feature extraction of power cable. Finally, some new ideas and further researches are proposed.

2. Application of SWE and SWPE in a Power System

2.1. Application of SWE in a Power System

SWE is a feature extraction method for transient signals in a power system, and professor He is an earlier specialist who applied SWE to identifying and processing of transient signals in a power system [13]. In 2001, professor He proposed the multi-resolution entropy which is Shannon wavelet time entropy, and it is applied to fault diagnosis in a power system. Then, reference [14] defines two kinds of SWE and corresponding calculation methods, and it is applied to feature extraction of transient signals. The result of experiment shows that this method is very efficient. In 2005, professor He defined Shannon wavelet energy entropy (SWEE), Shannon wavelet time entropy (SWTE), Shannon wavelet singularity entropy (SWSE) and corresponding algorithms were put forward. The mechanisms of three wavelet entropies were analyzed. Simulation results indicated that three wavelet entropies can be applied to fault detection of power system [15]. Thereafter other scholars and experts have made a lot of research work about the identifying and processing of transient signals based on the above research.

In reference [16], a wavelet-entropy-based PD de-noising method has been proposed. The features of PD are characterized by combining wavelet analysis that reveals the local features and Shannon entropy that measures the disorder. Comparing with other methods such as the energy-based method and the similarity-comparing method, the proposed wavelet-entropy-based method is more effective in PD signal de-noising. In reference [17], detection of fault type has been implemented by using Shannon wavelet entropy. Different types of faults are studied obtaining various current waveforms. These current waveforms are decomposed by using wavelet analysis. The wavelet entropies of such decompositions are analyzed reaching a successful methodology for fault classification. In reference [18], SWE is applied to system fault-detection and out-of-step protection during power swing, and stable power swings from unstable ones is distinguished by the method. Combined with the characteristic of SWE, reference [19] presents the appliance of SWE in the non-unit transient protection and the accelerated trip of transmission line protection. A new scheme of two applications of protection based on SWSE is presented. Compared with the criterion of wavelet model maximum, the criterion of accelerated trip overcomes the influence of signal magnitude.

In order to improve the accuracy of feature extraction and fault recognition, based on advantages of SWE, many experts and scholars combine SWE with neural networks, support vector machine (SVM), etc. Reference [20] analyzes the principles and the features of various voltage sags caused by power system short circuit faults, startup of induction motors and operation of power transformers, a method is proposed to identify voltage sag sources based on SWE and probability neural network. In reference [21], a grounding fault detection method based on SWE is proposed for loop net of DC system, which combines wavelet analysis with Shannon entropy to extract signal feature, and it achieves intelligent recognition. Based on Shannon entropy and SVM, reference [22] proposes a novel method for fault type recognition in distribution network. Then, the method can identify the fault types rapidly and accurately, and it can not be affected by transition resistance or faulty location and so on. Reference [23] proposes a method to identify short-time power quality disturbances based on improved SWEE and SVM. And it can classify the short-time power quality disturbances.

2.2. Application of SWPE in a Power System

Based on the analysis for SWE, it is found that when signal frequency components concentrate in the high frequency band, due to the roughness of wavelet decomposition in the high frequency band, high frequency components of similar frequency will be in the same scale and calculation accuracy of SWE is directly affected. To overcome the roughness of wavelet decomposition in the high frequency band, many experts and scholars combine wavelet packet transform (WPT) with Shannon entropy, and carry out much work on the identifying and processing of transient signals. In reference [24], the feature extraction method based on the Shannon wavelet packet time entropy (SWPTE) is proposed for the timely monitoring of distribution network and the quick identification of its operating states: normal, abnormal and faulty. With better adaptability and being immune to the network topology, line type, fault type, fault occurrence time, fault location and transition resistance, the proposed method can correctly identify the typical operating states of distribution network. Reference [25] uses SWPE to analyze the transient fault current in the protection location of the series compensated transmission line with static synchronous series compensator. In reference [26], based on WPT and Shannon entropy theory, a new method is presented to diagnosis fault for high voltage circuit breakers, and its steps and analysis are also introduced. The experimentation indicates the method can easily and accurately diagnose breaker faults, and gives an excellent resolution for fault diagnosis of HV circuit breakers.

Based on research above, many experts and scholars combine SWPE with neural networks, S-transform, SVM to extract fault feature and improve accuracy of feature extraction. Reference [27] combines Shannon entropy with WPT, and a new fault diagnosis method for power system is proposed, which achieves intelligent fault diagnosis. Aiming at the power quality and disturbance recognition, reference [28] proposes an automated recognition method based on SWPE and the modified S-transform. The experimental results show that the proposed method can effectively recognize the single and combined PQ disturbances. Based on Shannon entropy, S-transform, PSO algorithm and SVM, reference [29] proposes a novel method for mechanical fault diagnosis of high voltage circuit breakers, and it can accurately extract fault feature and classify. For a motor mechanical fault diagnosis, reference [30] combines Shannon entropy with SVM and genetic algorithm, and presents the method of motor mechanical fault diagnosis, and the experiment proves its reliability and veracity. Reference [31] build an air-gap discharge model in simulative transformer tank, collecting PD signals based on constant voltage method, utilizing wavelet packet decomposition method to partition the PD signal bands obtaining signal energy distribution in each frequency band as well as total signal energy tendency along with PD development process. The new PD parameter describing the development process and Shannon wavelet packet energy entropy (SWPEE) are proposed based on the signal energy variation in each frequency band. Due to the cyclic change of wavelet packet energy entropy, the step points of SWPE are taken as the way to effectively divide the PD development stage.

2.3. Problems Existed in SWE and SWPE in a Power System

The biggest feature of SWE and SWPE is to combine the advantages of wavelet multi-resolution analysis with the Shannon entropy theory of system complexity characterization to extract signal features hidden in the original power signal. The results deduced by SWE and SWPE involve not only the complexity and uncertainty of system model, but also the features of the fault information, which provides a new solution for transient signal feature extraction and transient fault recognition. But the recent analyses indicate that SWE and SWPE still have some problems as follows:

1. When the measured signal is more complex and contains a lot of random signals, there is the severe energy leakage and frequency aliasing in the wavelet coefficients (or reconstructed signals) with increasing of the wavelet decomposition scale [32]. So, the complexity and feature of signal can not be accurately expressed when the adjacent wavelet coefficients (or reconstructed signals) are taken as basic data to participate in the calculation of SWE and SWPE.

2.   When most signal components concentrate in the high frequency band, due to the roughness of wavelet decomposition in the high frequency band, the high frequency components of similar frequency will be in the same scale and calculation accuracy of SWE is directly reduced.

3.   For feature extraction of different transient signals, the studies, about the relationship between different entropy statistical properties and the signal feature, are still in the initial stage, which needs theoretical basis to support the transient signal extraction.

4.   For the different transient signals, sampling frequency of signals and wavelet decomposition scale exert influence on the accuracy and speed of feature extraction, but there has been no relevant research lately.

3. Comparison of Feature Extraction Accuracy and Wavelet Aliasing Effect on SWE and SWPE

3.1. Comparison in Accuracy of Feature Extraction for SWE and SWPE

SWE and SWPE are separately based on WT and WPT. Frequency decomposition of WT is shown in Figure 1, and frequency decomposition of WPT is shown in Figure 2.
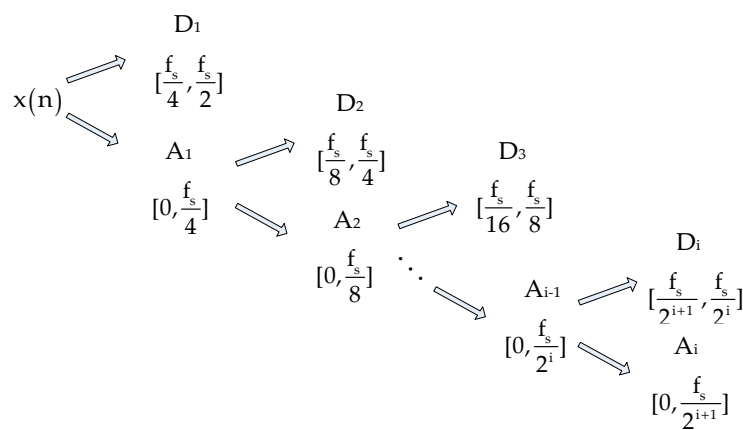

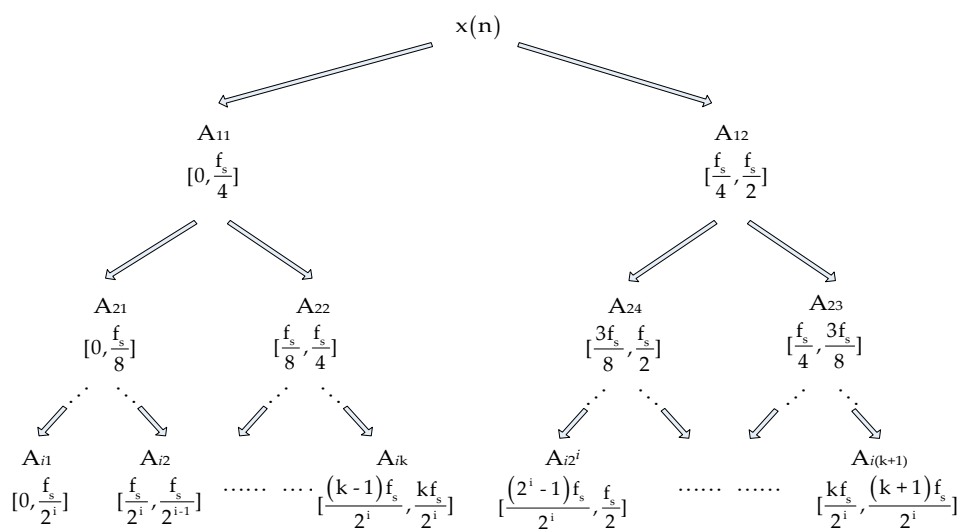**Figure 1.** Frequency decomposition of WT.

Figure 2. Frequency decomposition of wavelet packet transform (WPT).

Through the analysis, when signal frequency components concentrate in the high frequency band, due to the roughness of wavelet decomposition in the high frequency band, high-frequency components of similar frequency will be in the same scale and calculation accuracy of SWE is directly

reduced. So, feature extraction accuracy of SWPE is significantly greater than SWE. SWSE and Shannon entropy wavelet packet singular entropy (SWPSE) are cited as examples to compare their accuracy of feature extraction, and derivative process is as follows [33].

WPT is used to decompose $x(n)$ into m scales and it is reconstructed, and the frequency band of single branch reconstructed signal $D_i(n)$ and $A_i(n)$ are

$$\begin{cases} D_i(n) : \left[2^{-(i+1)}f_s, 2^{-i}f_s\right] \\ A_i(n) : \left[0, 2^{-(i+1)}f_s\right] \end{cases} \quad i = 1, 2, \cdots, m \tag{1}$$

where sampling frequency is expressed as f_s. For convenience, the $A_i(n)$ is expressed as $D_i(n)$ $(i = m+1, \cdots, 2^m)$, so $x(n) = \sum\limits_{i=1}^{2^m} D_i(n)$.

It is supposed that $D_i(n)$ involves k_i different frequency components, so

$$\begin{cases} p_m(i) = \dfrac{\lambda_i}{\sum\limits_{i=1}^{m} \lambda_i} \\ q_{ik} = \dfrac{\lambda_{ik}}{\sum\limits_{k=1}^{k_i} \lambda_{ik}} \end{cases} \tag{2}$$

where singular value of $D_i(n)$ is expressed as λ_i; the singular value of the different frequency components in $D_i(n)$ is expressed as λ_{ik}, $p_m(i)$ is simplified as p_i.

According to the correlation properties of reconstructed signals for WT, when the correlation of reconstructed signal is strong on the adjacent nodes, the frequency components are more similar. When the reconstructed signals of nodes are approximate on the adjacent nodes, the corresponding singular value will be to zero. Reversely, when the frequency components of reconstructed signals are great difference on nodes, the singular value will increase correspondingly.

According to the definition of SWSE, the $x(n)$ is calculated as follows.

$$W_{SWSE}(p_1, \cdots, p_m) = -\sum\limits_{i=1}^{m} p_i \ln p_i \tag{3}$$

Obviously, by analyzing Equations (2) and (3), the traditional SWSE does not accurately express k_i different frequency components of $D_i(n)$, therefore frequency component of each reconstructed signal is so complicated that the sum of its singular value is greater than the singular value corresponding to $D_i(n)$ without frequency sub-dividing, which is $p_i \leq \sum\limits_{j=1}^{k_i} q_i(j)$ and $q_i(j)$ is simplified as q_{ik}. The calculating process by SWPSE is listed as follows.

$$\begin{aligned} W_{SWPSE} & \left(q_{11}, \cdots, q_{1k_1}, \cdots, q_{m1}, \cdots, q_{mk_m}\right) \\ & = -\sum\limits_{i=1}^{m} \sum\limits_{k=1}^{k_i} q_{ik} \ln q_{ik} \\ & = -\left[\sum\limits_{i=1}^{m} p_i \ln p_i + \sum\limits_{i=1}^{m} \sum\limits_{k=1}^{k_i} q_{ik} \ln q_{ik} - \sum\limits_{i=1}^{m} p_i \ln p_i\right] \\ & = W_{SWSE}(p_1, \cdots, p_m) - \sum\limits_{i=1}^{m} \sum\limits_{k=1}^{k_i} q_{ik} \ln q_{ik} + \sum\limits_{i=1}^{m} p_i \ln p_i \\ & \geq W_{SWSE}(p_1, \cdots, p_m) - \sum\limits_{i=1}^{m} \sum\limits_{k=1}^{k_i} q_{ik} \ln q_{ik} + \sum\limits_{i=1}^{m} \ln p_i \sum\limits_{k=1}^{k_i} q_{ik} \\ & = W_{SWSE}(p_1, \cdots, p_m) - \sum\limits_{i=1}^{m} p_i \sum\limits_{k=1}^{k_i} \dfrac{q_{ik}}{p_i} (\ln q_{ik} - \ln p_i) \\ & = W_{SWSE}(p_1, \cdots, p_m) + \sum\limits_{i=1}^{m} p_i W\left(\dfrac{q_{i1}}{p_1}, \cdots, \dfrac{q_{ik_i}}{p_m}\right) \end{aligned} \tag{4}$$

By analyzing Equations (3) and (4), Equation (4) has an extra part $\sum\limits_{i=1}^{m} p_i W\left(\frac{q_{i1}}{p_1}, \cdots \frac{q_{ik_i}}{p_m}\right)$. According to the non-negativity of Shannon entropy, when $W\left(\frac{q_{i1}}{p_1}, \cdots \frac{q_{ik_i}}{p_m}\right) \neq 0$,

$$W_{SWPSE}\left(q_{11}, \cdots, q_{1k_1}, \cdots, q_{m1}, \cdots, q_{mk_m}\right) > W_{SWSE}\left(p_1, \cdots, p_m\right) \tag{5}$$

Based on the above, when a reconstructed signal involves the multi-frequency components, its value of SWSE is smaller than SWPSE. It means that the accuracy of description for signal complexity is affected by the extent of detail partition for the frequency band with SWSE. Therefore, when all the frequency components of the measured signal can be decomposed into the corresponding frequency band on different wavelet scales, the complexity description will be the most accurate. However, frequency band are not evenly segmented, and the roughness of the frequency band partition will increase with the scale reducing. When the multiple high frequency components of the measured signal are attributed to the same frequency band, value of SWE is smaller than SWPE. So, the accuracy of SWPE for feature extraction is higher than the SWE.

To verify effectiveness of analysis mentioned, SWSE and SWPSE is used to extract the PD feature of power cable. Due to the electromagnetic interference caused by the mass amount of cables in the tunnel, with the impedance mismatching between high frequency current transformer (HFCT) and the data acquisition equipment, the signal features (mainly between 1 and 30 MHz) of the pulse current produced by PD are often submerged by background noise, the detection results processed by software and hardware are not still good enough. When PD occurs in XLPE cable (Table 1), a high-frequency pulse current will be generated and flows from the high potential of cable core to the low potential of metal sheath, and pass in ground through the cross connection box or ground wire. Therefore, high frequency current transformer (HFCT) (band width between 0.1 and 100MHz) is connected with a cross connection box or ground wire, pulse current signal can be collected based on the principle of electromagnetic coupling, and stored in acquisition equipment through coaxial cable. The PD detection process is shown in Figure 3 [33].

Table 1. Specification of cable.

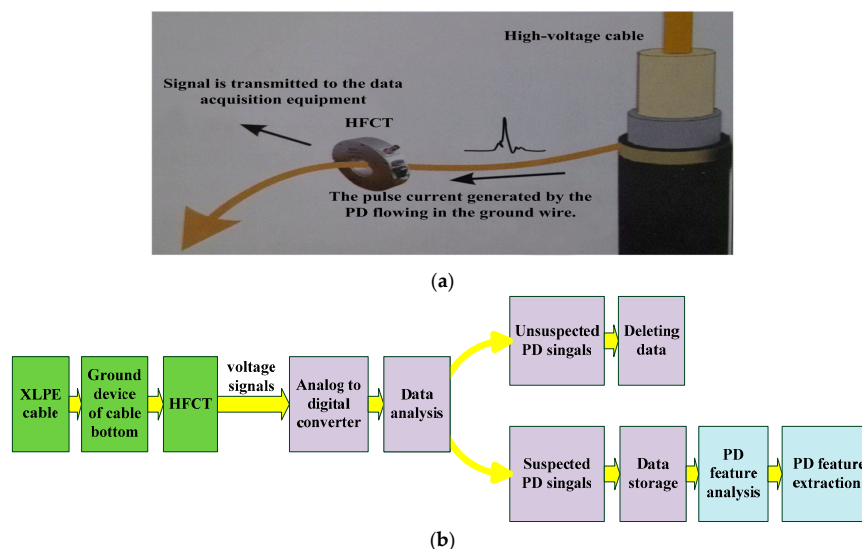| Model	Cable Core	Cross-Sectional Area	Insulation Layer	Metal Sheath	The Voltage Rating
YJLW03	Coppersplicing wire	800 mm^2	XLPE	Aluminum	127 kV/220 kV

Figure 3. Partial discharge (PD) detection process. (**a**) Working principle diagram of high frequency current transformer (HFCT); (**b**) Schematic diagram of PD detection system.

As shown in Figure 4, the HFCT is installed on the three-phase ground wire of the cross connection box, and is connected to the acquisition terminal APD-120D consisting of A/D module and data storage through the coaxial cable. The data collection on the spot is shown in Figure 5.
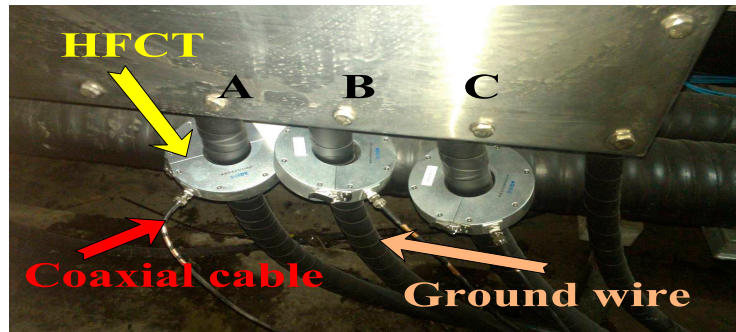
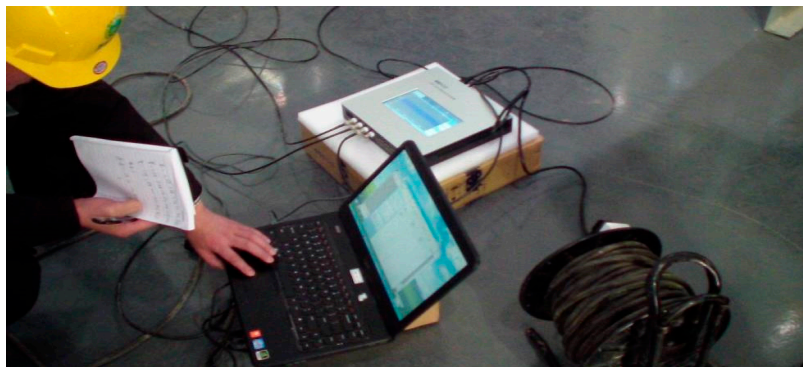

Figure 4. On-site installation of HFCT.

Figure 5. Data collection on the spot.

Using HFCT and data acquisition equipment to collect the PD signal and the sampling frequency is set at 100 MHz, and the original suspected PD signal is shown in Figure 6.
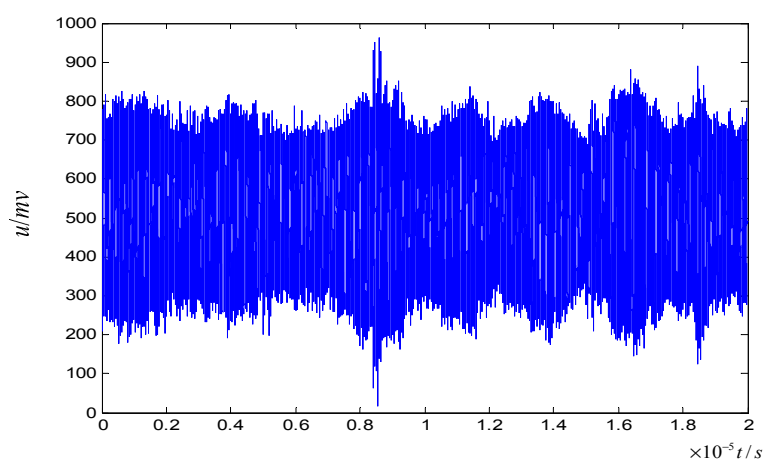
Figure 6. PD signal collected by HFCT.

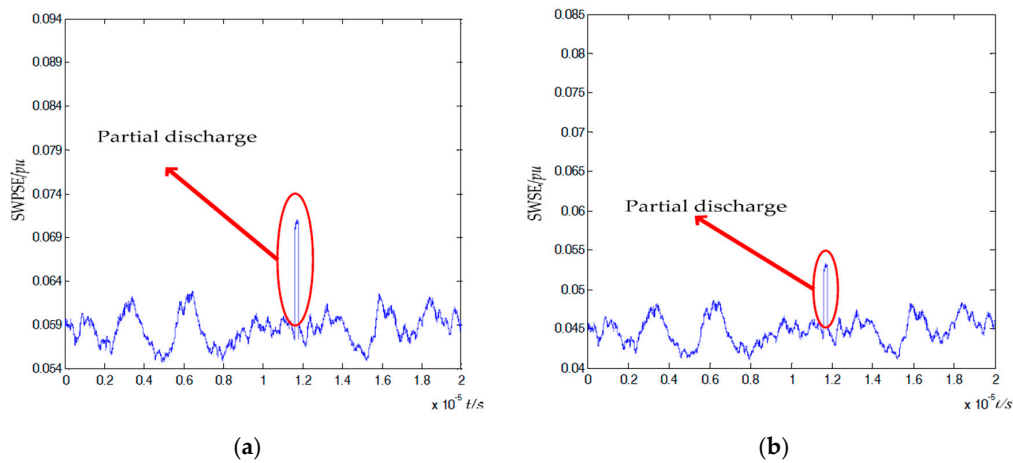SWSE and SWPSE are used to extract the PD feature of power cable and the corresponding curves are drawn in Figure 7.

Figure 7. Performance comparison of feature extraction between SWPSE and SWSE. (**a**) SWPSE; (**b**) SWSE.

By analyzing Figure 7, it is known that the accuracy of feature extraction by SWPSE is significantly higher than SWSE. So, the accuracy of SWPE for feature extraction is superior to SWE.

3.2. Wavelet Aliasing Effect on SWE and SWPE

According to the Section 2.3, there is the severe energy leakage and frequency aliasing in the wavelet coefficients (or reconstructed signals). In order to research effect of wavelet aliasing on feature extraction accuracy of the SWE and SWPE, Shannon wavelet energy entropy (SWEE) is as an example to analyze the relationship between the wavelet energy leakage and the accuracy of feature extraction [34].

Firstly, orthogonal wavelet is used to decompose the signal $x(t)$ into N scales. It supposes that wavelet aliasing does not occur at adjacent scales. So, the mathematical equation of ideal SWEE is expressed as:

$$
\begin{aligned}
W_{SWEE} &= -\sum_{i=1}^{N} p_i \ln p_i \\
&= -p_k \ln p_k - p_{k+1} \ln p_{k+1} - \sum_{i=1}^{k-1} p_i \ln p_i - \sum_{i=k+2}^{N} p_i \ln p_i
\end{aligned}
\tag{6}
$$

where

$$
\begin{aligned}
p &= \frac{E_i}{E} \\
E &= \sum_{i=1}^{N} E_i \ (i = 1, \cdots, N)
\end{aligned}
\tag{7}
$$

where $E = \sum_{i=1}^{N} E_i$ is the energy sum of N wavelet packet coefficient set or reconstructed signals. E_i is the energy sum of the *i*-th wavelet packet coefficient set or reconstructed signal. Then, it is supposed that when wavelet aliasing occurs at adjacent scales and the energy loss from k to k + 1 is expressed as α, the mathematical equation of SWEE is expressed as:

$$
\begin{aligned}
W'_{SWEE} &= -\sum_{i=1}^{N} p'_i \ln p'_i \\
&= -p'_k \ln p'_k - p'_{k+1} \ln p'_{k+1} - \sum_{i=1}^{k-1} p'_i \ln p'_i - \sum_{i=k+2}^{N} p'_i \ln p'_i
\end{aligned}
\tag{8}
$$

where $p'_i = \frac{E'_i}{E}$ $(i = 1, \cdots, N)$, $p'_k = \frac{E_k - \alpha}{E}$, $p'_{k+1} = \frac{E_{k+1} - \alpha}{E}$.

Then, the difference function $\Phi(\alpha)$ is constructed as follows:

$$
\begin{aligned}
\Phi(\alpha) \;\; &= W'_{SWEE} - W_{SWEE} \\
&= -p'_k \ln p'_k - p'_{k+1} \ln p'_{k+1} + p_k \ln p_k + p_{k+1} \ln p_{k+1} \\
&= -(p_k - p_\alpha) \ln (p_k - p_\alpha) - (p_{k+1} - p_\alpha) \ln (p_{k+1} - p_\alpha) + p_k \ln p_k + p_{k+1} \ln p_{k+1} \\
&= f(p_k - p_\alpha) - f(p_k)
\end{aligned}
\tag{9}
$$

where

$$
\begin{aligned}
f(p_k) \quad &= -p_k \ln p_k - p_{k+1} \ln p_{k+1} \\
&= -p_k \ln p_k - (C - p_k) \ln (C - p_k) \\
f(p_k - p_\alpha) \;\; &= -(p_k - p_\alpha) \ln (p_k - p_\alpha) - (p_{k+1} + p_\alpha) \ln (p_{k+1} + p_\alpha) \\
&= -(p_k - p_\alpha) \ln (p_k - p_\alpha) - (C - p_k + p_\alpha) \ln (C - p_k + p_\alpha)
\end{aligned}
\tag{10}
$$

where $C = p_k + p_{k+1}$. In order to analyze the properties of $f(p_k)$, it is transformed as follows:

$$
\begin{aligned}
f(p_k) \quad &= -p_k \ln p_k - p_{k+1} \ln p_{k+1} \\
&= C \left(-\tfrac{p_k}{C} \left(\ln \tfrac{p_k}{C} + \ln C \right) - \tfrac{p_{k+1}}{C} \left(\ln \tfrac{p_{k+1}}{C} + \ln C \right) \right) \\
&= C \left(-\tfrac{p_k}{C} \ln \tfrac{p_k}{C} - \tfrac{p_{k+1}}{C} \ln \tfrac{p_{k+1}}{C} \right) - (p_k + p_{k+1}) \ln C \\
&= CH(P) - C \ln C
\end{aligned}
\tag{11}
$$

where $H(P) = -\tfrac{p_k}{C} \ln \tfrac{p_k}{C} - \tfrac{p_{k+1}}{C} \ln \tfrac{p_{k+1}}{C}$.

Apparently, H(P) is the Shannon information entropy that has two-channel information source of the prior probability. According to the concavity of Shannon entropy function, H(P) meets the following conditions:

$$
\begin{aligned}
H(\delta P + (1 - \delta) Q) &\geq H(\delta P) + (1 - \delta) H(Q) \\
0 &\leq H(P) \leq \ln 2
\end{aligned}
\tag{12}
$$

Combined with Equations (11) and (12), $f(p_k)$ meets the following conditions:

$$
\begin{aligned}
f(\delta p_k + (1 - \delta) q_k) &\geq \delta f(p_k) + (1 - \delta) f(q_k) \\
C \ln \tfrac{1}{C} &\leq f(p_k) \leq C \ln \tfrac{2}{C}
\end{aligned}
\tag{13}
$$

So, $f(p_k)$ has the properties of the concavity, maximum and minimum values.

Based on the above, the curve of $f(p_k)$ is drawn in Figure 8. Combined with feature of wavelet aliasing, it is supposed that $0 \leq p_\alpha \leq p_k$. Then, $\Phi(\alpha)$ is discussed in two conditions as follows:

(1)　When $0 \leq p_\alpha \leq p_k$,

The curve of $f(p_k)$ corresponds to A$\hat{\mathrm{M}}$ in Figure 1. With increasing value of p_α, $\Phi(\alpha) < 0$, and $f(p_k - p_\alpha) < f(p_k)$ are found, which demonstrates $W'_{SWEE} < W_{SWEE}$ as increasing of energy leakage.

(2)　When $0 \leq p_{k+1} \leq p_k$,

If $0 \leq p_\alpha \leq p_k - p_{k+1}$, with increasing value of p_α, $\Phi(\alpha) > 0$, and $f(p_k - p_\alpha) > f(p_k)$ are found, which demonstrates $W'_{SWEE} > W_{SWEE}$ as increasing of energy leakage.

If $p_\alpha = 0$, it demonstrates that wavelet aliasing does not occur at adjacent scales and $W'_{SWEE} = W_{SWEE}$.

According to the above, there exists the close relationship among the calculation result of SWEE, wavelet aliasing and energy distribution at adjacent scale. The research of SWPEE is similar with SWEE which is not discussed in this article. To reduce the effect of wavelet aliasing on accuracy of feature extraction, Shannon entropy should be placed by entropy that has adjustment parameter. By adjusting parameter of entropy, the effect of wavelet aliasing can be reduced.
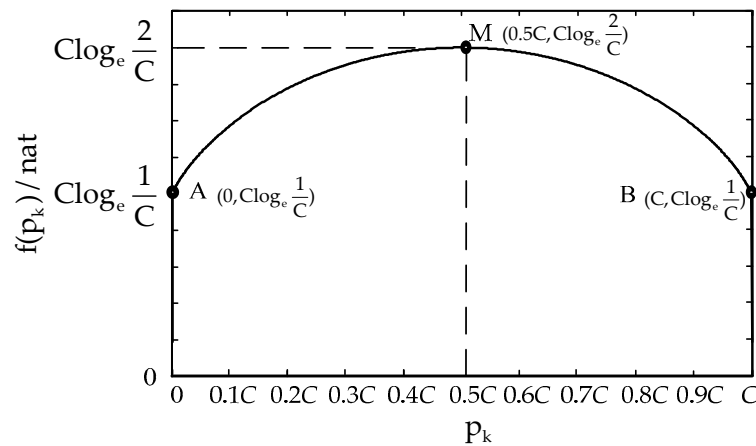
Figure 8. The curve of $f(p_k)$.

Because of adjustment parameter q of Renyi entropy, Renyi entropy is cited as an example to reduce the effect of wavelet aliasing on accuracy of PD feature extraction. Renyi entropy is an extension of Shannon entropy, Renyi entropy is equivalent to Shannon when q = 1. In many cases, the Renyi entropy has better statistical properties than the Shannon entropy when q $\neq$ 1. Renyi entropy is defined as follows [33].

$$W_{RE}(m) = \begin{cases} \frac{1}{1-q}\ln\left(\sum_n p_m^q(j)\right) & q > 0, q \neq 1 \\ -\sum_n p_m(j)\ln p_m(j) & q = 1 \end{cases} \tag{14}$$

The selection of q plays an important role in PD feature extraction. Taking a three-level system as the analysis object, according to Equation (14), the statistical results are calculated, and the corresponding relation between Renyi entropy and probability distribution are shown in Figure 9a–d. According to Figure 9a–d, when q > 0 and q $\neq$ 0, with the increase of q, the statistical range of Renyi entropy will expand for the system state of small probability event, and the statistical sensitivity of the small probability event will reduce correspondingly. On the contrary, with the decrease of q, the statistical range of small probability event is reduced, and the statistical sensitivity is increased. When q → 1, Renyi entropy is in accord with Shannon entropy. Because $W_{SE} = -\sum_{m=1}^{n} p_m \ln p_m$, when $p_i = 0$, the value of the Shannon entropy may be missing as shown in Figure 10. At this point, the Shannon entropy statistics will fails. Therefore $p_i \ln p_i = 0$ is usually defined when $p_i = 0$. Based on the analysis above, the more smaller parameter q of Renyi entropy is, the more accurate PD feature is.

Renyi wavelet packet energy entropy (RWPEE) is used to extract PD feature and q = 0.1. The result of RWPEE is shown in Figure 11. Comparing with Figures 7 and 11, PD feature is obvious by using RWPEE and it can effectively reduce the effect of wavelet aliasing on accuracy of feature extraction by selecting parameter q of Renyi entropy.
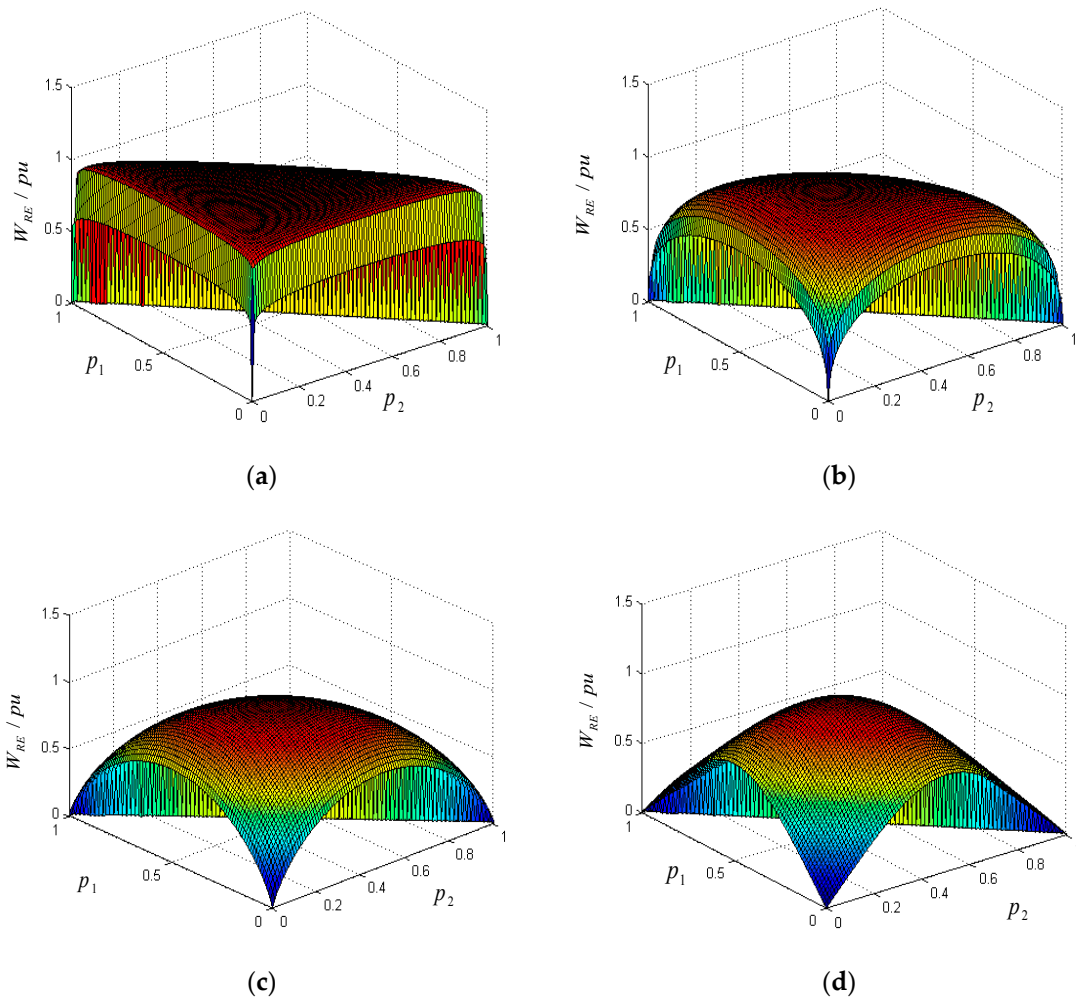
(**a**)　　　　　　　　　　　　　　　　　　　　　　　　(**b**)

(**c**)　　　　　　　　　　　　　　　　　　　　　　　　(**d**)

Figure 9. Relation between Renyi entropy with different values of q and probability distribution. (**a**) q = 0.1; (**b**) q = 0.5; (**c**) q = 0.99; (**d**) q = 2.
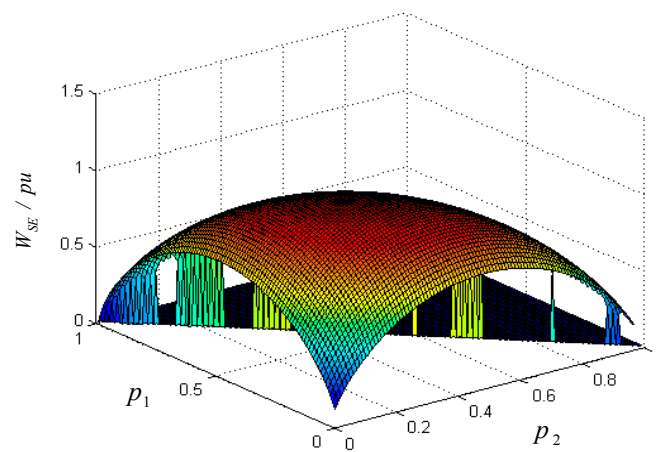

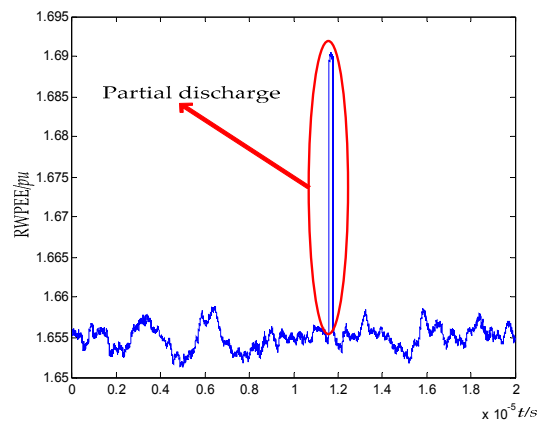**Figure 10.** Relation between Shannon entropy and probability distribution.

Figure 11. Feature extraction of Renyi wavelet packet energy entropy (RWPEE).

4. Conclusions

In this article, researches about SWE and SWPE in a power system are introduced. Problems existed in SWE and SWPE in a power system are raised. Then, it is discussed in detail for accuracy of feature extraction and effect of wavelet aliasing between SWE and SWPE. Meanwhile, the analyses mentioned are verified by PD feature extraction of power cable. Although SWE and SWPE have been widely used in feature extraction of transient signals and transient fault recognition, there are still some problems to be solved. So, the future researches include:

1. Basic theory about analysis and operation mechanism of transient signals based on wavelet and different entropy should be deeply studied and improved. For the negative influence of wavelet aliasing on SWE and SWPE, selecting different entropy of various statistical properties, optimizing the parameter of entropy, adjusting sampling frequency and selecting different orthogonal wavelet bases should be considered to reduce the effect of wavelet aliasing on accuracy of feature extraction.
2. When SWE and SWPE are applied to relay protection, there are some difficulties such as high sampling rate, complex calculation, etc. So, engineering applications put forward higher requirements for the ability of real time application. The further researches should focus on the optimizing algorithm structure and the improving operation speed of SWE and SWPE.

Author Contributions: Jikai Chen conceived the idea, and Jikai Chen, Yanhui Dou, Yang Li and Jiang Li wrote the paper. All authors have read and approved the final manuscript.

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

References

1. Chen, J.K.; Li, H.Y.; Wang, Y.J.; Xie, R.H.; Liu, X.B. A novel approach to extracting casing status features using data mining. *Entropy* **2014**, *16*, 389–404. [CrossRef]
2. Luo, G.M.; He, Z.Y.; Lin, S. Discussion on using discrepancy among wavelet relative entropy values to recognize transient signals in power transmission line. *Power Syst. Technol.* **2008**, *32*, 47–51.
3. El-Zonlcoly, A.M.; Desoulci, H. Wavelet entropy based algorithm for fault detection and classification in FACTS compensated transmission line. *Int. J. Electr. Power* **2011**, *33*, 1368–1374. [CrossRef]
4. Dehghani, M.; Khooban, M.H.; Nilcnam, T. Fast fault detection and classification based on a combination of wavelet singular entropy theory and fuzzy logic in distribution lines in the presence of distributed generations. *Int. J. Electr. Power* **2016**, *78*, 455–462. [CrossRef]

5. Fu, L.; He, Z.; Qian, Q. Feature extraction of fault transient and fault type determination for transmission ehv lines. *Proc. CSEE* **2010**, *30*, 100–106.

6. Zhou, L.; Xia, X.; Wan, Y.; Zhang, H.; Lei, P. Harmonic detection based wavelet transform. *Trans. China Electrotech. Soc.* **2006**, *21*, 67–74.

7. Zhang, B.; Sun, J. A power quality analysis method based on mallat algorithm and fast fourier transform. *Power Syst. Technol.* **2007**, *31*, 35–40.

8. He, Z.; Fu, L.; Lin, S.; Bo, Z. Fault detection and classification in EHV transmission line based on wavelet singular entropy. *IEEE Trans. Power Deliv.* **2010**, *25*, 2156–2163. [CrossRef]

9. Lin, S.; Gao, S.; He, Z.Y.; Deng, Y.J. A pilot directional protection for HVDC transmission line based on relative entropy of wavelet energy. *Entropy* **2015**, *17*, 5257–5273. [CrossRef]

10. Samui, A.; Samantaray, S.R. Wavelet Singular Entropy-Based Islanding Detection in Distributed Generation. *IEEE Trans. Power Deliv.* **2013**, *28*, 411–418. [CrossRef]

11. PiotrKowski, R.; Castro, E.; Gallego, A. Wavelet power, entropy and bispectrum applied to AE signals fordamage identification and evaluation of corroded galvanized steel. *Mech. Syst. Signal Process.* **2009**, *23*, 432–445. [CrossRef]

12. Liu, Z.; Hu, Q.; Cui, Y.; Zhang, Q. A new detection approach of transient disturbances combining wavelet packet and Tsallis entropy. *Neurocomputing* **2014**, *142*, 393–407. [CrossRef]

13. He, Z.Y.; Qian, Q.Q. Multi-resolution entropy and its application in EHV transmission line fault detection. *Electr. Power Autom. Equip.* **2001**, *21*, 9–12.

14. He, Z.Y.; Liu, Z.G.; Qian, Q.Q. Study on wavelet entropy theory and adaptability of its application in power system. *Power Syst. Technol.* **2004**, *28*, 17–21.

15. He, Z.Y.; Cai, Y.M.; Qian, Q.Q. A study of wavelet entropy theory and its application in electric power system fault detection. *Proc. CSEE* **2005**, *25*, 38–43.

16. Luo, G.; Zhang, D.; Tseng, K.J.; He, J. Impulsive noise reduction for transient Earth voltage-based partial discharge using Wavelet-entropy. *IET Sci. Meas. Technol.* **2016**, *10*, 69–76. [CrossRef]

17. El Safty, S.; El-Zonlcoly, A. Applying wavelet entropy principle in fault classification. *Electr. Power Energy Syst.* **2009**, *31*, 604–607. [CrossRef]

18. Dubey, R.; Samantara, S.R. Wavelet singular entropy-based symmetrical fault-detection and out-of-step protection during power swing. *IET Gener. Transm. Distrib.* **2013**, *7*, 1123–1134. [CrossRef]

19. Liu, Q.; Wang, Z.P.; Zheng, Z.H. Application of wavelet singular entropy theory in transient protection and accelerated trip of transmission line protection. *Autom. Electr. Power Syst.* **2009**, *33*, 79–83.

20. Jia, Y.; He, Z.Y.; Zhao, J.J. A method to identify voltage sag sources in distribution network based on wavelet entropy and probability neural network. *Power Syst. Technol.* **2009**, *33*, 63–69.

21. Li, D.H.; Wang, B.; Ma, Y.X. Grounding fault detection based on wavelet entropy and neural network for loop net of DC system. *Electr. Power Autom. Equip.* **2008**, *28*, 51–54.

22. Wang, Y.S.; Tan, Z.Y.; Liu, X.M. Fault type recognition for distribution network based on wavelet singular entropy and support vector machine. *Power Syst. Prot. Control* **2011**, *39*, 16–20.

23. Li, G.Y.; Wang, H.L.; Zhao, M. Short-time power quality disturbances identification based on improved wavelet energy entropy and SVM. *Trans. China Electrotech. Soc.* **2009**, *24*, 161–167.

24. Yu, N.H.; Li, C.J.; Yang, J.; Cai, M.; Dong, B.; Gong, L.Y.; Ma, Y.Y. Operating state feature extraction based on wavelet-packet time entropy for distribution network. *Electr. Power Autom. Equip.* **2014**, *34*, 64–71.

25. Liu, Q.; Chang, Y.; Xu, Y.; Hao, W. Fault position identification for series compensated transmission lines with sssc based on improved wavelet packet entropy. *Autom. Electr. Power Syst.* **2011**, *35*, 65–70.

26. Sun, L.J.; Hu, X.G.; Ji, Y.C. Fault diagnosis for high voltage circuit breakers with improved characteristic entropy of wavelet packet. *Proc. CSEE* **2007**, *27*, 103–108.

27. Zhang, J.; Wang, X.G.; Li, Z.L. Application of neural network based on wavelet packet-energy entropy in power system fault diagnosis. *Power Syst. Technol.* **2006**, *30*, 72–80.

28. Liu, Z.G.; Cui, Y.; Li, W.H. Combined power quality disturbances recognition using wavelet packet entropies and S-transform. *Entropy* **2015**, *17*, 5811–5828. [CrossRef]

29. Huang, N.T.; Chen, H.J.; Zhang, S.X.; Cai, G.W.; Li, W.G.; Xu, D.G.; Fang, L.H. Mechanical fault diagnosis of high voltage circuit breakers based on wavelet time-frequency entropy and one-class support vector machine. *Entropy* **2016**, *18*, 1–17. [CrossRef]

30. Zhang, Y.N.; Wei, W.; Wu, L.L. Motor mechanical fault diagnosis based on wavelet packet, Shannon entropy, SVM and GA. *Electr. Power Autom. Equip.* **2010**, *30*, 87–91.

31. Chen, W.; Xie, B.; Long, Z.; Cui, L.; Li, Y.; Zhou, Q.; Chen, X. Stage identification in air-gap discharge of oil-impregnated paper insulation based on wavelet packet energy entropy. *Proc. CSEE* **2016**, *36*, 563–569.

32. Chen, J.K.; Li, H.Y.; Yang, S.Y.; Kou, B.Q. Application of wavelet packet singularity entropy and PSD in power harmonics detection. *Trans. China Electrotech. Soc.* **2010**, *25*, 193–199.

33. Chen, J.K.; Dou, Y.H.; Wang, Z.H.; Li, G.Q. A novel method for PD feature extraction of power cable with renyi entropy. *Entropy* **2015**, *17*, 7698–7712. [CrossRef]

34. Chen, J.K.; Li, G.Q. Tsallis wavelet entropy and its application in power signal analysis. *Entropy* **2014**, *16*, 3009–3025. [CrossRef]