A Novel Wavelet Selection Scheme for Partial Discharge Signal Detection under Low SNR Condition

Jiajia Liu, W.H. Siew, John J. Soraghan, and Euan A. Morris
University of Strathclyde
204 George Street, Glasgow, G1 1XW, UK

***Abstract*—Over the past two decades, wavelet-based techniques have been widely used to extract partial discharge (PD) signals from noisy signals. To effectively select the correct technique to minimize the effect of noise on PD detection, three aspects are considered: wavelet selection, decomposition scale, and noise or threshold estimation. For wavelet selection, popular techniques, including correlation-based wavelet selection scheme (CBWSS) and energy-based wavelet selection scheme (EBWSS), are applied to select an appropriate wavelet basis function. These two schemes, however, have their limitations. CBWSS is not as effective as expected when the signal to noise ratio (SNR) is very low. EBWSS selects the optimal wavelet that can maximize the energy ratio of the PD signal in approximation coefficients through wavelet decomposition. It is not strictly true for damped oscillating PD signals, particularly when the decomposition scale increases. As such, a novel wavelet selection scheme, wavelet entropy-based wavelet selection scheme (WEBWSS), is proposed to provide an alternative to CBWSS and EBWSS for PD denoising. PD signals are simulated and also obtained through laboratory experiments to demonstrate that this new method has better performance in the removal of noise, particularly when SNR is low.**

***Index Terms*—Wavelet-based technique, partial discharge, detection, denoising, wavelet selection, SNR, wavelet entropy**

I. Introduction

Partial discharge (PD) measurement is an effective technique for the monitoring of electrical insulation. However, PD signals are normally contaminated by noise from the environment, which increases the difficulty of their detection. To effectively extract PD signal from noisy signals, various denoising techniques, such as adaptive filter [1], [2], and the wavelet-based technique [3]–[6], have been adopted to remove noise. The wavelet-based technique has been widely used in recent years since wavelet transform can simultaneously provide signal information both in time and frequency domains. This advantage is particularly useful for the processing of non-stationary signals, e.g., PD signals.

For the application of wavelet transform, a noisy signal is decomposed into multi-scale wavelet coefficients by a selected basis function. Those wavelet coefficients associated with noise are processed by an estimated threshold, and thus, the significant features of the signal of interest are retained. Reconstruction is then performed to build the denoised signal. Based on the processes of decomposition and reconstruction, the choice of a suitable wavelet basis function is the first, and most, significant step for the application of wavelet-based denoising. As such, the investigation of an appropriate wavelet basis function for wavelet-based PD denoising has been performed in [7], [8].

A wavelet selection scheme was introduced in [7] based on the correlation coefficient between a known PD signal and wavelet waveform. This scheme is termed correlation-based wavelet selection scheme (CBWSS). The optimal wavelet is selected if it can maximize wavelet coefficients in wavelet analysis of PD signals. This approach for best wavelet selection, however, has an inherent limitation, it requires prior knowledge of PD waveforms. Also, it is not a scale-dependent wavelet selection scheme. The denoised PD signal may not be as good as expected. The most significant drawback, however, is that the PD signal is normally corrupted by the noise in the environment, which can lead to the selected wavelet being a match of the noisy PD signal rather than the pure PD signal, especially when the signal to noise ratio (SNR) is very low. In an attempt to overcome the limitation mentioned above in CBWSS, a scale-dependent energy-based wavelet selection scheme (EBWSS) was presented in [8]. The wavelet that can maximize the energy ratio of approximation coefficients at each decomposition scale is selected as the best wavelet. It has been demonstrated to outperform CBWSS [8]. In EBWSS, two typical PD waveforms, damped exponential PD pulse (DEP) and damped oscillating PD pulse (DOP), were used to demonstrate the energy criterion for the optimal wavelet selection. With further exploration in details of EBWSS, it has been found that the criterion is not strictly true for DOP signals, particularly when the decomposition scale increases. The motivation of this paper is therefore to provide an automated and data-driven selection scheme for the best wavelet selection in the context of PD denoising.

In this paper, the new wavelet entropy-based wavelet selection scheme is inspired by the concept of Shannon Entropy. Wavelet entropy can measure the randomness of the wavelet coefficients at each decomposition scale. The smaller the wavelet entropy, the lower the randomness of the wavelet coefficients. As such, the new selection scheme is proposed based on wavelet entropy, and termed wavelet entropy-based wavelet selection scheme (WEBWSS). Simulated PD signals, i.e., DEP and DOP, and PD signals obtained through laboratory experiment are used to demonstrate the performance of this novel wavelet selection scheme. Results show that it is a promising wavelet selection scheme to improve the effectiveness of PD denoising.

II. Wavelet-based technique

A. Wavelet Theory

Wavelet transform (WT) is an alternative approach to traditional methods, e.g. Fourier Transform, in signal processing. The major advantage of WT is that it can map a signal in the time-frequency plane. Due to this advantage, WT is a promising technique in the

analysis of variations in signals or images with the requirements of both time and frequency information. Generally, WT is achieved through the application of continuous wavelet transform (CWT) or discrete wavelet transform (DWT). DWT is preferable due to its representation of signals or images through its DWT coefficients without redundancy and, thus, is using less computational time. In this paper, the wavelet-based technique referred to is the DWT.

In [9], the wavelet expansion of a signal *x* can be expressed as

$$x = \sum_i \sum_j C_{j,i} \cdot \psi_{j,i}\,, \tag{1}$$

where both *i* and *j* are integer, *i* is the time-delay index and *j* is the scale index. $\{\psi_{j,i}\}$ is the expansion set of wavelet basis functions, and $\{C_{j,i}\}$ is the set of expansion coefficients, or wavelet coefficients, which is called the discrete wavelet transform of *x*. The expansion in (1) is the inverse discrete wavelet transform (IDWT). The scheme of DWT for signal decomposition is depicted in Figure 1. A signal is convolved with the low- and high-pass filters (*h* and *g*) and followed by a downsampling operation by 2. In signal processing terminology, the outputs of the low- and high-pass filters are termed approximation and detail coefficient respectively. The approximation coefficient is used as the input signal for next-scale decomposition. This decomposition is iterated until the predefined scale, J, reaches. It is important to note that the maximum decomposition scale J_{max} is defined as $log_2(N)$, where *N* is the length of the input signal. The reconstruction of the input signal, i.e., inverse DWT (IDWT), is a reverse operation as shown in Figure 1. Instead of downsampling in DWT, upsampling is involved in IDWT.

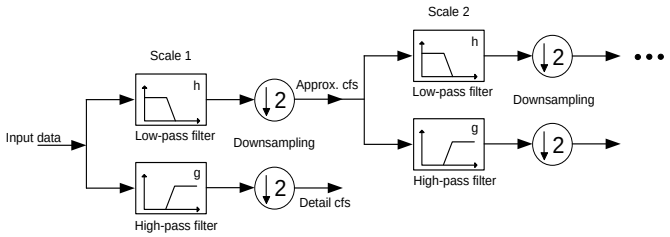

Fig. 1. The implementation of DWT in signal decomposition

B. Wavelet Denoising Technique

The wavelet denoising theory is dependent on the fundamental idea that the energy of a signal is often concentrated in only a few coefficients while the energy of noise is widely spread among all the coefficients in the wavelet domain [7]. General procedures for the wavelet-based denoising of a signal are presented as follows:

1) Apply DWT to decompose the noisy signal *s* with a selected wavelet to a predefined scale *J*, and obtain approximation coefficients a_J at the final scale *J* and detail coefficients d_j at decomposition scale *j*, where *j = 1, 2, ..., J*.
2) Estimate the threshold through a noise estimation technique and apply this threshold to detail coefficients, d_j , at decomposition scale *j* using hard or soft thresholding scheme.
3) Apply IDWT to the approximation coefficients a_J and the processed detail coefficients d_j' to reconstruct the denoised signal s'.

Based on the noise estimation technique proposed in [7], [10], the scale-dependent threshold used in this paper is estimated by

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

where $MAD|\cdot|$ is the median absolute deviation of the detail coefficients d_j at decomposition scale *j*, and n_j is the length of d_j. For the thresholding scheme, soft thresholding in [10] is used in this paper, the function is given by

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

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

C. Wavelet Entropy

The concept of wavelet entropy was derived from Shannon entropy and presented in [11]. Suppose $\{C_{j,i}\}$ are the wavelet coefficients obtained through a *J*-scale wavelet transform, in which *j* represents the decomposition scale and $j = 1, 2, \dots, J$, *i* denotes the i^{th} element in $C_{j,i}$ and $i = 1, 2, \dots, n_j$, n_j is the length of wavelet coefficients at scale *j*. The energy of wavelet coefficients at the decomposition scale *j* can be calculated by

$$E_j = \sum_i |C_{j,i}|^2\,. \tag{4}$$

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

$$p_i = \frac{|C_{j,i}|^2}{\sum_k |C_{j,i}|^2} = \frac{|C_{j,i}|^2}{E_j} \tag{5}$$

with $\sum_i p_i = 1$. Wavelet entropy $WE(j)$ is defined as follows [11]:

$$WE(j) = -\sum_i p_i \ln(p_i)\,. \tag{6}$$

Similar to Shannon entropy, wavelet entropy is applied to measure the degree of disorder of wavelet coefficients or signify the randomness of wavelet coefficients.

III. Partial Discharge Signals

Two theoretical PD pulses, i.e., damped exponential PD pulse (DEP) and damped oscillating PD pulse (DOP), are simulated using their mathematical frames derived based on two different PD detecting circuits [4], [8]. In this paper, DEP and DOP are given by the formula in [8]:

$$s_1(t) = A(e^{-\alpha_1 t} - e^{-\alpha_2 t}), \tag{7}$$

$$s_2(t) = A(e^{-\alpha_1 t}cos(w_d t - \varphi) - e^{-\alpha_2 t}cos\varphi), \tag{8}$$

where $s_1(t)$ and $s_2(t)$ are the DEP and DOP respectively. The values of A, α_1, α_2, f_d, w_d and φ used in these two equations are listed in Table 1.

TABLE 1. Values of parameters used in (7) and 8) [8].

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

The simulated sampling frequency f_s is set to $60MHz$. Figure 2 shows these two simulated PD signals both in time and frequency domains. Generally, DOP signal shown in Figure 2 (c) and (d) is closer to a real high-frequency PD signal detected from electrical power equipment in practice [8].

To develop the new scheme for practical use, real PD signals were generated through an artificial defect of a $7mm \times 7mm$ breach in the outer conductor created in a 1.5m 11 kV ethylene propylene rubber-insulated (EPR) cable sample [12]. PD signals were collected using a high frequency current transformer (HFCT). The specifications of the HFCT are listed in Table 2. Details regarding the experiment setup are depicted in Figure 3 [12].

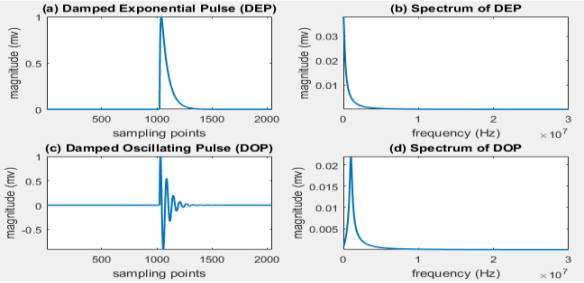

Fig. 2. (a) and (b): DEP signals simulated in time and frequency domain respectively; (c) and (d): DOP signals simulated in time and frequency domain respectively.

TABLE 2. Specifications of the HFCT.

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

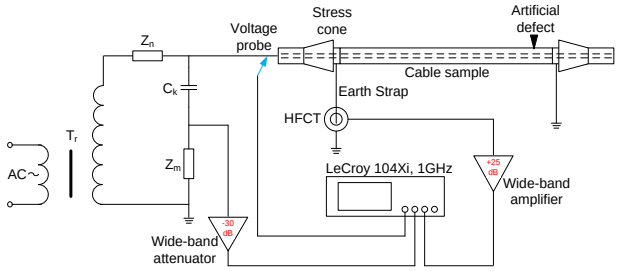

Fig. 3. PD testing of a defective 11 kV EPR cable. HFCT was used to collect PD pulses (C_k and Z_m represent the coupling capacitor and measuring impedance respectively).

Experiments were performed at various voltage levels. The PD pulses measured at 9kV are used as the real PD signals to demonstrate the new wavelet selection scheme in this paper. One PD pulse, named s_3, with 2048 sample points was selected and depicted in Figure 4.

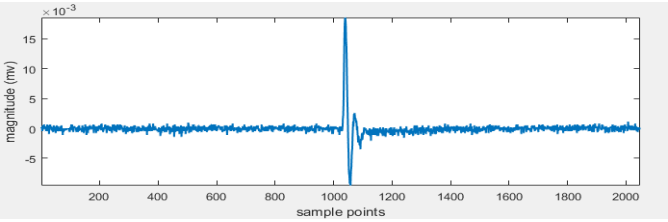

Fig. 4. Real PD pulse, s_3, detected from a defective EPR cable under 9kV AC voltage

IV. WAVELET ENTROPY-BASED WAVELET SELECTION SCHEME

In [13], it was shown that wavelet entropy value is inversely proportional to the energy concentrated in the number of wavelet coefficients. It is also known that white noise, the noise source for PD corruption in this paper, has high degree of randomness or disorder, and thus, the entropy value can describe the random characters of noise [14]. Based on this, a smaller value of wavelet entropy indicates that the wavelet used for WT decomposition can preserve more energy of the original signal in fewer coefficients and contain less white noise in the wavelet coefficients and, consequently, the wavelet used is closer to the best wavelet as expected. A new criterion for the best wavelet selection is therefore proposed, i.e., a wavelet that can have minimum wavelet entropy of the approximation coefficients at each decomposition scale through WT decomposition will be selected for denoising of PD detection. The new method has several promising advantages: it is scale-dependent, automated, and data-driven.

The general process for the proposed novel wavelet selection scheme is illustrated in the flow chart in Figure 5.

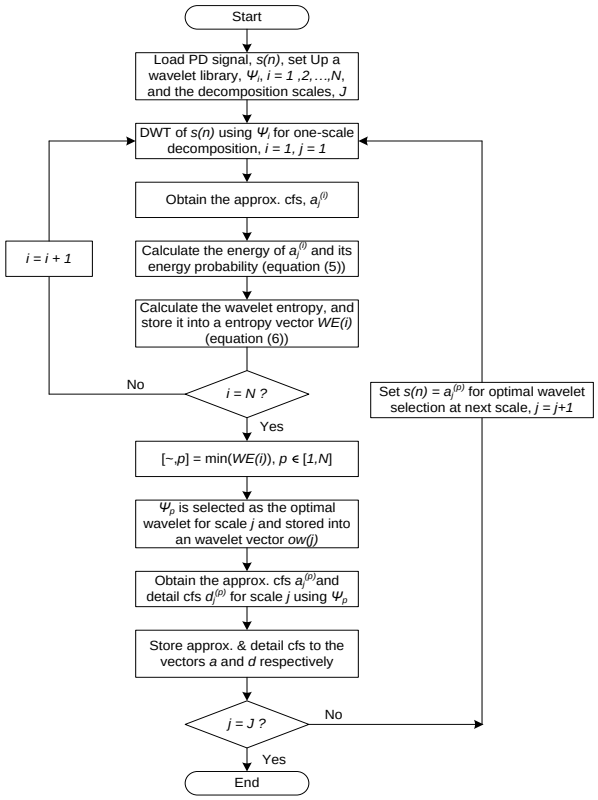

Fig. 5. Flow chart of the general process of WEBWSS.

Given a wavelet library $\{\psi_i : i = 1,2,\dots,N\}$, one wavelet of which is selected for a one-level DWT decomposition of a noisy PD signal *s(n)* each time. Next the wavelet entropy of the generated approximations is calculated based on (5) and (6). The wavelet ψ_p $(1 \le p \le N)$ that minimize the wavelet entropy of approximations will be selected as the best wavelet. The selected ψ_p is then applied for the DWT decomposition of *s(n)* for the first scale, obtaining approximation coefficients ${a_1}^{(p)}$ and detail coefficients ${d_1}^{(p)}$. Finally, ${a_1}^{(p)}$ is used as the input signal for next scale DWT decomposition, using the strategy presented above.

When the predefined decomposition scale J reaches, the best wavelet for each scale will be successfully selected.

V. RESULTS AND ANALYSIS

Parameters, e.g., magnitude error (ME), mean square error (MSE), and cross correlation (XCORR) are adopted in this paper to evaluate the performance of the denoising results of different wavelet selection schemes. ME, MSE, and XCORR are calculated by the equations as follows,

$$ME = \frac{m - m'}{m} \tag{9}$$

$$MSE = \frac{\sum_{i=1}^{N}[s(i) - s'(i)]}{N} \tag{10}$$

$$XCORR = \frac{\sum_i (s(i) - \bar{s}(i))(s'(i) - \bar{s'}(i))}{\sqrt{\sum_i (s(i) - \bar{s}(i))^2}\sqrt{\sum_i (s'(i) - \bar{s'}(i))^2}} \tag{11}$$

where m and m' are the magnitudes of $s(i)$ and $s'(i)$ respectively. $s(i)$ represents the original signal and $s'(i)$ denotes the denoised signal. N is the length of signals. $\bar{s}(i)$ and $\bar{s'}(i)$ indicate the mean of $s(i)$ and $s'(i)$ respectively. Better denoised results can be obtained with lower ME, MSE, and higher XCORR.

In this paper, PD signals are corrupted by white noise, and then, various wavelet selection schemes are used to remove the noise and evaluated by the parameters mentioned above. Two simulated PD signals, s_1 and s_2, and real PD signal, s_3, as well as their noisy signals NS_1, NS_2 and NS_3 with SNR = -10 are depicted in Figure 6. Note that the original real PD signal shown in Figure 4 is corrupted by ambient noise during experiment. To mitigate the effect of this noise on the denoising results, it has been pre-processed using the method introduced in [8]. The smoothed real PD signal is depicted in Figure 6 (e).

The related parameters used to evaluate their performance on PD detection are listed in Table 3. It can be seen from Table 3 that WEBWSS has better performance than the others of wavelet-based denoising of PD detection. Simultaneously, it also verifies the conclusion presented in [8] that EBWSS outperforms CBWSS in PD denoising.

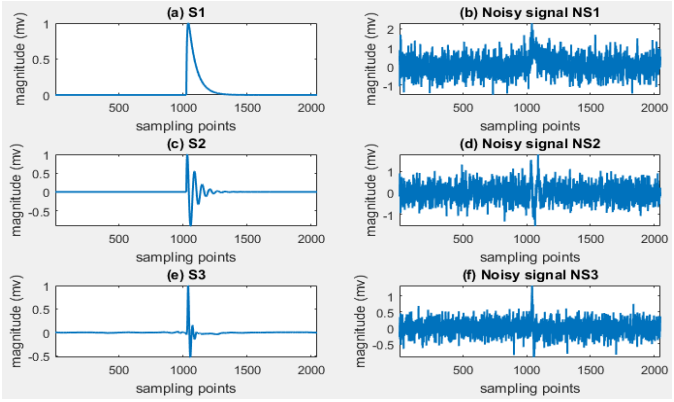

Fig. 6. (a) and (b): s_1 and its noisy signal NS_1 with SNR = -10, (c) and (d): s_2 and its noisy signal NS_2 with SNR = -10, (e) and (f): s_3 and its noisy signal NS_3 with SNR = -10.

TABLE 3. Parameters used to evaluate the performance of selection schemes.

		CBWSS	EBWSS	WEBWSS
S_1	ME	0.3701	0.3457	**0.3071**
	MSE	0.0055	0.0028	**0.0024**
	XCORR	0.8689	0.9358	**0.9495**
S_2	ME	0.5686	0.5673	**0.5238**
	MSE	0.0054	0.0044	**0.0041**
	XCORR	0.8757	0.9182	**0.9414**
S_3	ME	0.5937	0.5457	**0.4997**
	MSE	0.00035	0.00037	**0.00027**
	XCORR	0.8501	0.9369	**0.9502**

VI. CONCLUSIONS

A novel wavelet selection scheme was presented in this paper, which was shown to improve the effectiveness of wavelet-based denoising of PD detection. The wavelet entropy can measure the energy preservation and randomness of decomposed wavelet coefficients. As such, it has been developed as a criterion for the best wavelet selection in wavelet-based PD denoising. Its scale-dependent, automated, and data-driven properties enable it to be a promising technique in the context of PD detection, particularly when the SNR is low.

REFERENCES

[1] H. Borsi, "Digital location of partial discharges in HV cables," *IEEE Trans. Electr. Insul.*, vol. 27, no. 1, pp. 28–36, 1992.

[2] M. S. Mashikian, F. Palmieri, R. Bansal, and R. B. Northrop, "Location of partial discharges in shielded cables in the presence of high noise," *IEEE Trans. Electr. Insul.*, vol. 27, no. 1, pp. 37–43, 1992.

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

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

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

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

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

[8] J. Li, T. Jiang, S. Grzybowski, and C. Cheng, "Scale dependent wavelet selection for de-noising of partial discharge detection," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 17, no. 6, pp. 1705–1714, 2010.

[9] S. Alessio, *Digital Signal Processing and Spectral Analysis for Scientists: Concepts and Applications*. Springer, 2015.

[10] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[11] Y. Zhuang and J. S. Baras, "Existence and Construction of Optimal Wavelet Basis for Signal Representation," *IEEE Trans. Signal Process.*, 1994.

[12] X. Hu, W. H. Siew, M. D. Judd, and X. Peng, "Transfer function characterization for HFCTs used in partial discharge detection," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 2, pp. 1088–1096, 2017.

[13] M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, vol. 38, no. 1. IEEE Press, 1996.

[14] Z. He, X. Chen, and G. Luo, "Wavelet Entropy Measure Definition and Its Application for Transmission Line Fault Detection and Identification; (Part I: Definition and Methodology)," *2006 Int. Conf. Power Syst. Technol.*, pp. 1–6, 2006.