Comparing Denoising Performance of DWT,WPT, SWT and DT-CWT for Partial Discharge Signals

S. H. Mortazavi S. M. Shahrtash
Center of Excellence for Automation and Operation of Power System
Iran University of Science and Technology (IUST), Tehran, Iran
Shmy58@Gmail.com Shahrtash@Iust.ac.ir

***Abstract*-Denoising is one of the major part in on-line Partial Discharge(PD) measurement system. In the recent years wavelet transform has been widely used in PD denoising area and its effectiveness discussed in many papers. In this paper, different wavelet based transforms such as DWT, WPT, SWT and DT-CWT and their performance in PD signal denoising with definition proper criteria have been investigated.**

I. Introduction

Continuous monitoring of the electrical insulations of high voltage equipments, such as electrical machines and high voltage cables, needs on-line Partial Discharge (PD) measurement systems. Such measurements can help to schedule device maintenances and prevent electrical break down.

But measured PD signals are often mixed with noises in various types that may make the PD pulses to be completely covered with, and consequently may causes mistakes in PD analysis. In this respect, denoising is one of important requirements for PD detection and features extraction systems.

Various DSP based algorithms have been proposed for PD signals denoising such as digital filters or Wavelet Transform.

Wavelet transform is a powerful mathematical tool for analysis and synthesis of digital signals. Discrete wavelet transform (DWT) decompose the signal in different levels. Each level has a unique time and frequency resolution thus it gives both time and frequency information. With proper thresholding in each level, coefficients that estimated to be noise, remove from decomposed data and with the remained coefficients, denoised signal can be constructed. In DWT, after down sampling, low frequency coefficients will be decomposed to obtain next level decomposition.

In wavelet packet transform (WPT) in addition to low frequency coefficients, high frequency coefficients also will be decomposed to obtained richer resolution.

In both DWT and WPT coefficients in each level, after filtering will be down sampled. Due to down sampling, these transforms will suffer from the lack of shift invariance. Stationary Wavelet Transform (SWT) eliminates down sampling operators at each level to obtain shift invariance property but it is a very redundant transform. Dual-Tree Complex Wavelet Transform (DT-CWT) is a transform that has a shift invariant property and a limited redundancy.

In this paper after a brief introduction of DWT, WPT, SWT and DT-CWT their features for PD denoising will be investigated and their performance for some noisy PD signal will be compared.

II. Denoising based on wavelet transform

A. DWT

Fig. 1 illustrates the Filter Bank (FB) implementation of DWT (up to two levels). As is shown, the original signal (*S*) passes through a pair of low pass (*h(n)*) and high pass (*g(n)*) filters. For perfect reconstruction, these filters must satisfy some mathematical properties [1]. Then outputs of each filter will be down sampled by a factor of two. Outputs of low pass and high pass (*g(n)*) filters are called approximation coefficients (Ca) and detail coefficients (Cd) respectively. Ca represents low frequency and Cd, the high frequency components of the signal. In DWT, For the next level of decomposition, the approximation coefficients will pass through the same low pass and high pass filters and then will be down sampled to obtain the next level detail and approximation coefficients and so on.

So the results of the DWT are a series of coefficients in one approximation and J details, where J is the number of the final decomposition level. These coefficients construct an orthogonal basis and the original signal can be reconstructed through them by applying the inverse wavelet transform (IWT).

Down-sampling plays a crucial part in the process of decomposition. While a signal is decomposed, the signal length is halved every time it passes through the filter pair, leaving the signal with a length of 1/2, 1/4, 1/8 ... of the original length, at level 1, 2, 3... and it allows to use same pair of filter in different levels and prevent redundancy.

When a noisy PD signal is decomposed, the signal and noise manifest differently in the post-decomposition results, making it possible to separate them by applying a threshold properly to each of the levels.

Wavelet based denoising procedure in general, involves three steps:

1) Decomposing: Choosing a mother wavelet and a maximum decomposition level J and then computing the decomposition coefficients at each level.

2) Thresholding: Computing threshold values for each level (for the whole set of the coefficients or for each level separately) and applying threshold (in soft or hard process) to the coefficients at each level

3) Reconstruction: Reconstructing the signal with the modified coefficients through the procedure which is reverse of decomposition.

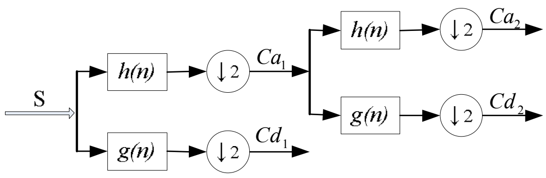

Fig. 1. FB implementation of DWT

According to these steps, three parameters should be selected; mother wavelet, maximum decomposition level and threshold values.

Proper mother wavelet can represent signal features in a few wavelet coefficients with high magnitude that can improve thresholding and consequently, denoising performance. Optimum mother wavelet selection can be done by using cross correlation function [2]. The optimum wavelet maximizes the cross correlation between the signal of interest (here PD pulses) and the mother wavelet.

Optimum decomposition level depends on signal and noise frequency characteristics and may be obtained by trial and error.

Selection of the threshold values is the most important part of denoising procedures, where small threshold values result in remained noises in the reconstructed signal and the large values may eliminate some signal features.

Donoho and Johnston developed a universal thresholding rule, which can effectively remove the Gaussian random noise [3]. It uses a fixed threshold value for all coefficients in different level. In [2] a modified version of universal thresholding rule were introduced that calculates threshold for each level separately as follow:

$$\lambda_j = \sigma_j \sqrt{2 * \log N_j} \quad (1)$$

Where N_j is the length of coefficients at jth level and σ_j is the standard deviation of noise at that level which can be estimated by [2]:

$$\sigma = \frac{MAD}{0.6745} \quad (2)$$

and MAD represents median absolute value of the coefficients.

Effectiveness of modified universal thresholding for PD denoising was investigated in [3], so in this paper we use a modified universal threshold for PD denoising.

After thresholding, signal can be recovered with modified coefficients through inverse wavelet transform.

B. WPT

WPT is a generalization of DWT. While in DWT for the next level decomposing, only the approximation coefficient(Ca) passed through the filters, in WPT both detail and approximation coefficient pass through filters[1]. Fig. 2 shows the structure of WPT decomposition by filter bank.

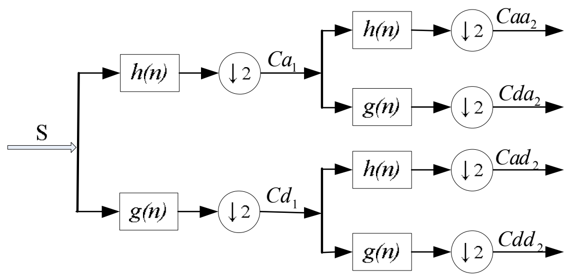

Fig. 2. FB Implementation of WPT

Decomposing the detail coefficients at each level, results a richer resolution. In a complete decomposition wavelet packet tree for a signal of length N ($N=2^j$), There are at least $2^{N/2}$ wavelet packet orthogonal bases that original signal can be reconstructed through them.

In WPT based denoising, there are the same three steps as DWT denoising, but In WPT denoising after decomposition, the best basis for signal reconstruction must be selected. Then thresholding and reconstruction of signal will be done over coefficients related to this basis. The algorithm for the best basis selection is described in [4].

In the case of adaptively chosen basis, it is proposed that calculate the threshold values as follow [5]:

$$\lambda = \sigma \sqrt{2 * \log(N * \log_2 N)} \quad (3)$$

This threshold is however not optimal and maybe eliminate some signal features (and as a result, increases MSE) but it reduces the probability of presence noise in reconstructed signal.

It must be considered that an adaptive basis selection may also find basis that better correlate the noise components and it may result presence of noise in recovered signal [1].

C. SWT

In both DWT and WPT, after filtration the coefficients will down sampled, that prevents redundancy and allow using the same pair of filter in different levels. And so, these transforms will suffer from the lack of shift invariance, which means that small shifts in the input signal can cause major variations in the distribution of energy between coefficients at deferent levels and may causes some error in reconstruction.

This problem is carried out by eliminating the down sampling steps after filtration at each level in stationary wavelet transform (SWT). By eliminating down sampling, the number of coefficients at each level is as long as original signal. Fig. 3 shows decomposition of a signal by SWT up two levels.

In decomposition a signal through a filter bank, if down sampling operators were eliminated, for the next level of decomposition the high and low pass filters must be modified. For this, the low pass and high pass filters at each level will be up sampled by putting zero between each filter's coefficients of previous level that called á trous algorithm [6]. Denoising a signal by SWT has the same three steps as DWT, which mentioned previously.

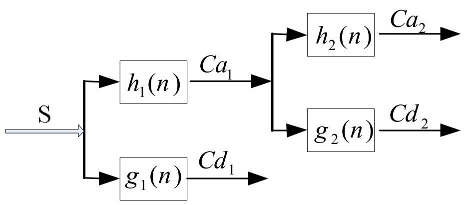

Fig. 3. FB implementation of SWT

D. DT-CWT

SWT solves the lack of shift invariance problem by eliminating down sampling operators, but it is a very high redundant transform. It represent a redundancy as much as N*J (N is the length of input signal length and J the maximum number of decomposition level) and as a result, denoising procedure consume considerable computation time.

One solution for shift invariance wavelet transform without high redundancy is complex wavelet transform which takes advantage of complex-valued filters instead of real ones. For the perfect reconstruction properties, the complex filters must be analytic. But designing of an analytic complex-valued filters for implementation in filter banks is difficult. Kingsbury introduced a more computationally efficient approach for a shift invariance transform; the Dual-Tree Complex Wavelet Transform (DT-CWT) that has the following properties [7]:

- Approximate shift invariance
- Perfect reconstruction
- Limited redundancy, independent of the number of levels it has 2:1 redundancy
- Efficient order-N computation, only twice the DWT

The DT-CWT employs two real DWT filter banks; the first DWT gives the real part of the transform while the second DWT gives the imaginary part.

The two real wavelet transforms use two different sets of filters, with each satisfying the PR conditions. The two sets of filters are jointly designed so that the overall transform is approximately analytic[8].

Filter bank used to implement the DT-CWT is shown in Fig. 4. Where, $h_0(n)/h_1(n)$ denote the low-pass/high-pass filter pair for the real part, and $g_0(n)/g_1(n)$ denote the low-pass/high-pass filter pair for the imaginary part of filter bank.

Note that the filters are themselves real; no complex arithmetic is required for the implementation of the DT-CWT. The number of coefficients in DT-CWT at each level is two times of DWT, so that the DT-CWT is a redundant transform but its redundancy, compare with SWT is a few.

In DT-CWT based denoising, the three steps of denoising (Similar to DWT based denoising) will be done for each tree (real and imaginary filter banks) separately and then the output signals of two trees are averaged to obtain the final recovered signal. The filter bank for implementing the inverse of DT-CWT is illustrated in Fig. 5.

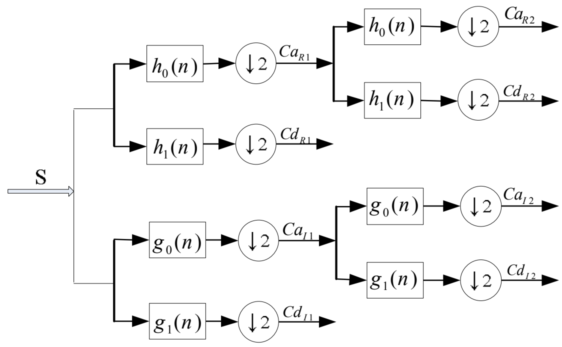

Fig. 4. FB implementation of DT CWT

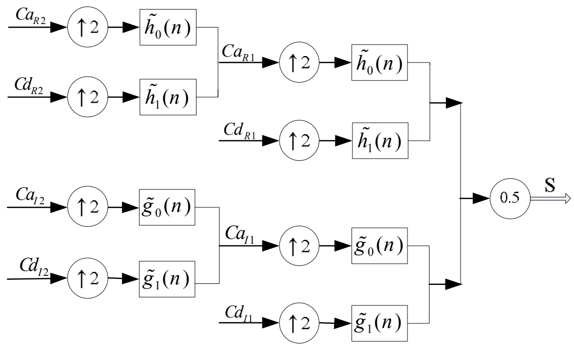

Fig. 5. Reconstruction signal through DT CWT inverse transform

III. Simulation Parameters

For investigation and comparing the performance of different wavelet based partial discharge denoising, first the mathematical models for PD pulse and the noise must be developed. Then proper criteria for measurement of denoising performance should be defined.

A. PD model

Generally there are two basic models for the simulating of PD pulses. First model represents the PD pulses through a damping exponential (DE) function as follows:

$$\begin{cases} x(t) = A e^{-\frac{t-t_0}{\tau}} & t \geq t_0 \\ x(t) = 0 & t < t_0 \end{cases} \tag{4}$$

Where A represents the amplitude of the pulse, t_0 is the time of pulse occurrence and τ is the damping factor. Fig. 6 shows an example of a PD signal with three DE pulses. Characteristics of pulses are printed in Table I.

The second model implements the damping oscillatory exponential (DOE) function for representing the PD pulses, as follows:

$$\begin{cases} x(t) = A e^{-\frac{t-t_0}{\tau}} \cos(2\pi f_0 (t-t_0)) & t \geq t_0 \\ x(t) = 0 & t < t_0 \end{cases} \quad (5)$$

Where f_0 is the frequency of oscillation and A, t_0 and τ are the same as (4). An example of a PD signal with three DOE pulses is shown in Fig. 6. A, t_0 and τ of these pulses are the same as DE pulses in Table I and f_0 of pulses are 300, 200 and 260 kHz respectively.

B. Noise simulation

The source major external interferences in PD measurements are [9]:

- Continuous sinusoidal noise from radio transmissions and communication systems.
- Stochastic noise associated with sparking, periodic pulse current of thyristor and electrical noise generated inside the detection circuit.
- Periodic and stochastic pulse shaped interferences from infrequent switching operation or lightning, arcing between adjacent metallic contacts and so on.

In addition to external interferences, there are noises due to detection system itself such as thermal noise.

Generally, there are AM (amplitude modulation) radio, FM (frequency modulation) radio and mobile communication in air. With the sampling rate of some tens of megahertz, FM radio and mobile communication signals have little effect on PD measurements and only AM radio signals are taken into account here.

For simulating sinusoidal noise a combination of a series of amplitude modulated signals as follow are used [9]:

$$e(t) = A\sum_{i=1}^{5}(c + m * \sin(2\pi f_m t)) * \sin(2\pi f_i t) \quad (6)$$

Where c represents the amplitude of the carrier wave, m is the amplitude of modulating signal, f_m *is* the frequency of modulating signal and f_i is the frequency of the carrier wave. In our simulation, the parameters used are c=1, m=0.4, f_m =1 *kHz*, f_i = 600 *kHz*, 800 *kHz*, 1000 *kHz*, 1200 *kHz*,1400 *kHz*, respectively.

Thermal and stochastic noises are formed simply by random number with zero mean and varied standard deviation.

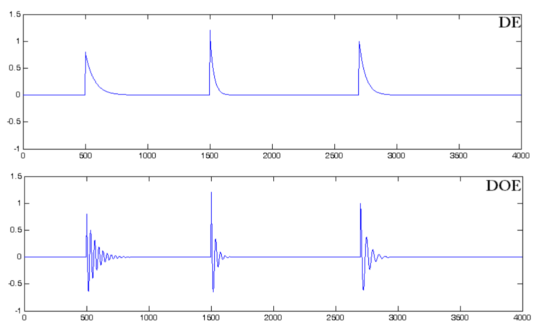

Fig. 6. An example of DE and DOE simulated signals

TABLE I
CHARACTERISTICS OF DE PULSES IN FIG 7.

	A	t_0	τ
Pulse 1	0.8	50 μs	7 μs
Pulse 2	1.2	150 μs	3 μs
Pulse 3	1	270 μs	5 μs

C. Comparison criteria

For measurement the extent of noise in signal, usually signal to noise ratio (SNR) is used. SNR in *db*, defined as [9]:

$$SNR(db) = 10 * \log_{10} \frac{E_s}{E_n} \quad (7)$$

Where E_s and E_n stand for the energy of PD signal and noise respectively.

For comparing the effectiveness of different denoising methods, we use the mean square error (*MSE*), as follow [9]:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(S(i) - R(i))^2 \quad (8)$$

Where *S(i)* and *R(i)* represent original signal without noise and recovered signal after denoising respectively and *N* is length of signal. In most cases, a smaller value of *MSE* indicates that the reconstructed data resembles the original signal better.

Cross Correlation (*CC*) function also can be used for comparing shape of the original PD pulses and reconstructed PD pulses after denoising. Cross correlation between two signals (*X* and *Y*) will be computed as follows:

$$CC(X,Y) = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}} \quad (9)$$

Where $\bar{X}$ and $\bar{Y}$, are the mean value of two signals *X* and *Y*, respectively.

It must be considered that *CC* compares only the shape between two signals and it didn't consider their magnitude.

IV. RESULTS

To compare the performance of different denoising methods, we used four simulated signals mixed with noise. Fig. 8 shows these signals and their characteristics are printed in Table II.

Assessment and performance evaluation of DWT, WPT, SWT and DT-CWT based denoising method for PD signals, were investigated through four signal mixed with the noise. Fig. 7 shows these noisy signals. Signals and noise Characteristics are printed in table II.

TABLE II
CHARACTERISTICS OF NOISY SIMULATED SIGNALS

	signal	White Noise (SNR)	AM Noise (Amplitude)
Sig. 1	DOE (Fig. 7)	-3 (db)	0
Sig. 2	DOE (Fig. 7)	-10 (db)	0
Sig. 3	DOE (Fig. 7)	-5 (db)	0.5
Sig. 4	DE (Fig. 7)	-5 (db)	0.5

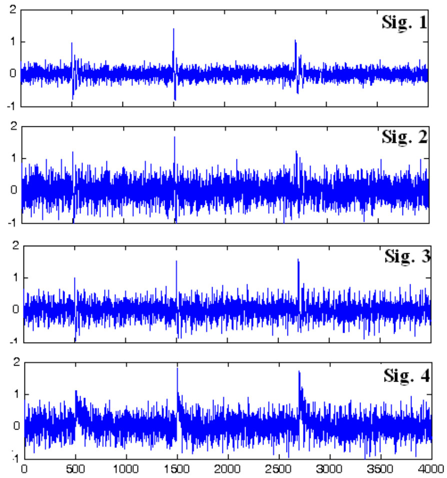

Fig. 7. Noisy simulated PD signals

Maximum decomposition level is chosen 6 level for all methods. In fact it gives best results for all methods for these signals [3]. In DWT, WPT and SWT denoising, "Coiflet-2" and "Symlet-4" were used as mother wavelet for DOE and DE types signal respectively. This selection obtained through cross correlation as mentioned before. And in DT-CWT "Antonini" filter is used. In DWT, SWT and DT-CWT, threshold values calculated by modified universal thresholding rule (1).

In WPT denoising, for best basis selection, we used SURE cost function that defined as follow:

$$E(s) = \#\{i \mid \ |s_i| \leq \lambda\} + \sum_i \min(s_i^2, \lambda^2) \qquad (10)$$

$E(s)$ represents entropy of signal (s), and $\#\{i \mid \ |s_i| \leq \lambda\}$ is the number of time instants when the signal is greater than a threshold λ. Other cost functions like Shannon entropy or threshold entropy were investigated but the results for SURE were better [3].

The times consumed for the executing of denoising procedure for different transforms are printed in Table III. These times are results of 50 iterations in the same condition for different transforms. As it shown, SWT based denoising is the most time consuming method (because of its redundancy) and the DWT based is the less.

The results of denoising by using different methods are printed in Table IV. The results are average of 50 various simulations. In each simulation PD pulses were shifted 10 samples and new noise added to signal till desired SNR was obtained. *MSE* values in Table IV are multiplied by 1000.

TABLE III
CONSUMED TIME FOR EXECUTING OF DENOISING PROCEDURE BASED ON DIFFERENT TRANSFORMS (TIME FOR 50 ITERATION)

Transform	Consumed Time (second)
DWT	2.4
WPT	10.6
SWT	41.4
DT-CWT	8.9

TABLE IV
THE RESULTS OF DENOISING SIGNALS WITH DIFFERENT TRANSFORMS

	DWT		**WPT**		**SWT**		**DT-CWT**	
	MSE	**CC**	**MSE**	**CC**	**MSE**	**CC**	**MSE**	**CC**
Sig.1	1.05	0.94	1.37	0.93	0.90	0.95	**0.7**	**0.96**
Sig.2	3.73	0.79	5.07	0.72	2.51	**0.89**	**2.38**	0.87
Sig.3	8.67	0.64	6.63	0.7	**1.83**	**0.92**	4.74	0.76
Sig.4	5.22	0.86	4.44	0.87	**1.58**	**0.95**	3.08	0.91

In the case of white noise only (Sig.1 and Sig.2), All of four methods can recover PD signals without noise but their performances are different. Implementation of DWT is the simplest among the others. It requires minimum computation and its results are acceptable. DWT compared with WPT, has better performance and require less computation. WPT needs more computation for decomposing detail coefficients and best basis selection. And as said before, in the best basis selection for noisy signal, it may select the basis that correlate with noise components better and results noise pulses in reconstructed signal. It seems that WPT is not suitable for PD denoising unless we could take advantage of some information about PD and noise frequency characteristics.

Among the all methods SWT, resulted maximum *CC* that indicates most resemblance between the shape of original and reconstructed pulses. But in the case of harsh white noise (Sig.2), DT-CWT has resulted the minimum *MSE*. According to the fewer redundancy of DT-CWT compared with SWT, and its results from Table IV, it seems that, it has the best performance among the all methods.

When signals are mixed with white noise and sine noise both (Sig. 3 and Sig. 4), SWT results the best and the other methods have a weak performance. It must considered that universal thresholding rule (or other automatic threshold selection methods) is designed for white gaussian noise and in the case of sine noises it may result the presence of noise in reconstructed signal.

When signals are mixed with different noises, we must use some information about PD and noise frequency characteristics for reliable denoising and the thresholding rules must be modified.

For example if instead of median absolute value for standard deviation estimation in (2) , we use mean absolute value, the results of denoising in some cases will improved. Table V shows the results of denoising with different methods while for thresholding we use mean absolute value in (2).

TABLE V
THE RESULTS OF DENOISING SIGNALS WITH DIFFERENT TRANSFORMS (USING MEAN FOR NOISE ESTIMATION)

	DWT		**WPT**		**SWT**		**DT-CWT**	
	MSE	**CC**	**MSE**	**CC**	**MSE**	**CC**	**MSE**	**CC**
Sig.1	1.85	0.9	2.1	0.88	1.61	**0.93**	**1.23**	**0.93**
Sig.2	4.57	0.72	4.8	0.70	4.21	**0.84**	**3.18**	0.82
Sig.3	3.71	0.78	3.32	0.806	2.83	**0.89**	**2.43**	0.87
Sig.4	3.53	0.89	4.07	0.876	3.04	**0.92**	**2.64**	**0.92**

It can be seen that except foe the SWT method, in the cases of complex noise (Sig.3 and Sig.4), the results are improved effectively. But with white noise only the results are worse than the results in Table IV. Fig. 8 shows the Sig.3, denoised with the different transforms using median absolute value in (2) for noise estimation and Fig. 9 shows the same signal (Sig.3), denoised with same conditions but using mean absolute value instead of MAD in (2) for noise estimation.

As can be seen in Fig. 8, except the signal that has denoised with SWT, there are many noisy pulses in the other signals but in the denoised signals in Fig. 9, these pulses have been eliminated.

V. CONCLUSION

We implemented various wavelet transform based denoising algorithms for some noisy PD signals. With comparing the results, it observed that in some cases DT-CWT, and in the other SWT have resulted the best. With consideration the redundancy of SWT compare with DT-CWT and the consumed time for executing the denoising algorithm (as printed in Table III), it seems that DT-CWT is the most proper transform for PD denoising. WPT results the worst in most cases. As mentioned before, adaptive basis selection may result presence of noise in recovered signal.

In the case of sinusoidal interference or low signal to noise ratio, we must use the higher threshold values than in (1). As we use the mean absolute value for noise estimation, although it may eliminate some signal features but it reduces the probability of presence of noisy pulses in the recovered signal.

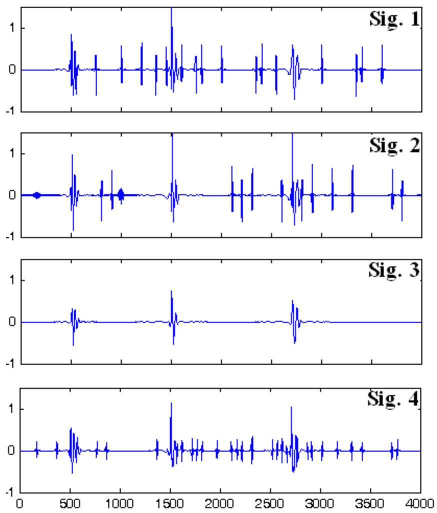

Fig. 8. Results of denoising of Sig.3 with different transform using median absolute value for noise estimation

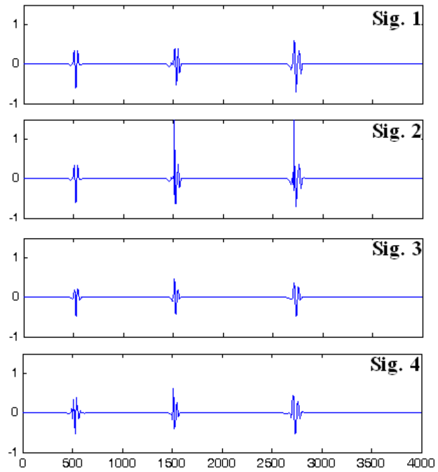

Fig. 9. Results of denoising of Sig.3 with different transform using mean absolute value for noise estimation

REFERENCES

[1] S. Mallat, A Wavelet Tour of Signal Processing. San Diego, CA: Academic, 1998.

[2] X. Ma, C. Zhou, I. J .Kemp, "Automated Wavelet Selection and Thresholding for PD Detection," IEEE Electrical Insulation Magazine, vol. 18, no.2, pp. 37-45, 2002.

[3] S. H. Mortazavi, "Wavelet Based Denoising of Partial Discharge Signals," M.Sc. Thesis, Iran University of Science and Technology(IUST), 2007,

[4] R. Coifman, M. Wickerhauser, "Entropy Based Algorithms for Best-Basis Selection," IEEE Trans. on Information Theory, vol. 38, no. 2, pp.713-718, March 1992.

[5] D. L. Donoho, I. M. Johnstone, "Ideal Spatial Adaptation via Wavelet Shrinkage," Biometrika, vol. 81, pp. 425-455, 1994.

[6] M. J. Shensa, "The Discrete Wavelet Transform Wedding À Trouse and Mallat Algorithms," IEEE Trans. on Signal Processing, vol. 40, no. 10, October 1992.

[7] N. G. Kingsbury, "The Dual-Tree Complex Wavelet Transform: a New Efficient Tool for Image Restoration and Enhancement," Proc. European Signal Processing Conference, EUSIPCO 98, Rhodes, pp. 319-322, September 1998.

[8] I. W. Selesnick, R. G. Baraniuk, N. G. Kingsbury, "The Dual-Tree Complex Wavelet Transform," IEEE Signal Processing Magazine, vol. 22, no. 6, pp. 123-151, Nov. 2005.

[9] X. Zhou, C. Zhou, I. J. Kemp, " An Improved Methodology for Application of Wavelet Transform to Partial Discharge Measurement Denoising," IEEE Trans. on Dielectrics and Electrical Insulation, vol. 12, no. 3, pp. 586-594, June 2005