Signal denoising techniques for partial discharge measurements

Please check the document version of this publication:

Signal Denoising Techniques for Partial Discharge Measurements

S. Sriram
Midas Communication Technologies Pvt. Ltd.
Chennai 600036, India

S. Nitin
Department of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta 30332, USA

K. M. M. Prabhu
Department of Electrical Engineering
Indian Institute of Technology Madras
Chennai 600036, India

and ## M. J. Bastiaans
Department of Electrical Engineering, P.O. Box 513, Bldg. Eh 3. 24,
Eindhoven University of Technology
5600 MB Eindhoven, The Netherlands

ABSTRACT

One of the major challenges of on-site partial discharge (PD) measurements is the recovery of PD signals from a noisy environment. The different sources of noise include thermal or resistor noise added by the measuring circuit, and high-frequency sinusoidal signals that electromagnetically couple from radio broadcasts and/or carrier wave communications. Sophisticated methods are required to detect PD signals correctly. Fortunately, advances in analog-to-digital conversion (ADC) technology, and recent developments in digital signal processing (DSP) enable easy extraction of PD signals. This paper deals with the denoising of PD signals caused by corona discharges. Several techniques are investigated and employed on simulated as well as real PD data.

Index Terms — **FFT, LMS, matched filtering, notch filtering, partial discharge (PD), PD models, real PD data, recursive least squares (RLS), short-time Fourier transform (STFT), wavelets, Wigner-Ville distribution (WVD).**

1 INTRODUCTION

WHEN high voltages are applied across insulators with defects, the insulators may either totally breakdown, or partial discharges (PDs) may occur. Partial discharges are localized discharges confined to the insulation system. They are caused due to the presence of weak spots, such as voids in the insulation where degradation takes place. It is well known that PD measurements are widely employed in testing power apparatus after manufacture [1]. However, there is a recent trend to extend them to on-site measurements, where the major problem encountered is the strong coupling of external noise, particularly discrete spectral interference (DSI). For a sensitive PD measure-

ment, these disturbing signals have to be rejected. In the case of development and routine tests, the PD measurements are carried out in the manufacturer's shielded laboratories, with filtered mains, to reach the demanded measurement sensitivity. However, the problems faced in PD measurements performed in unshielded laboratories as well as on-site conditions, is the strong coupling of external noise, particularly from broadcasting stations. This noise can be suppressed by several analog and digital techniques.

The giant strides made in digital signal processing (DSP) techniques provide a vast scope for significant reduction of these types of interference. Several of these techniques have been discussed in the literature [2-10]. In this paper, different denoising techniques are applied to simulated PD

data as well as real PD data obtained from the Department of High Voltage Engineering, Indian Institute of Science, Bangalore, India. The PD data is obtained from corona discharges. The simulated PD data is acquired from a standard PD model, an ac stochastic model proposed in [11]. The PD pulse shape employed is also obtained from the literature [8, 12]. We have considered several DSP techniques [13-23], such as short-time Fourier transform (STFT) and Wigner-Ville distribution (WVD)-based denoising methods to recover the PD pulses.

The paper is organized as follows: Section 2 describes the PD model and the PD pulse shape used in this paper. Section 3 discusses the noise model used for the simulated PD data and Section 4 describes the set-up used to obtain real PD data. Section 5 explains the various denoising techniques employed in this paper. Section 6 discusses the results obtained and Section 7 concludes the paper.

2   PD MODEL AND PULSE SHAPE

Partial discharge (PD) models can be divided into two types: (a) deterministic and (b) stochastic. In the case of deterministic models, a discharge is assumed to occur as soon as the local electric field exceeds the PD inception field, as opposed to a stochastic model where a time lag is caused while waiting for an initial electron to trigger a discharge. In this paper, we use a stochastic model to predict the time of occurrence of partial discharges. This model has been shown to produce results that are very close to experimental data [11].

2.1   ac STOCHASTIC MODEL

In this model [11], a PD process is treated as a stochastic process consisting of short duration discharges and charge carrier drift/recombination intervals between these discharges. Two assumptions are made:

1) The physical parameters of the PD system are time-independent, except for space or surface charges around the PD defect, which may drift and recombine in the electric field.

2) It is sufficient to consider the resulting electric field, E_i, being produced by the charge distribution at the PD defect. Thus, the time-dependent part of the physical state of the PD system is completely described by the internal field E_i.

During a discharge at time t, the total electric field $E_{tot}(t)$ is given by

$$E_{tot}(t) = E_o(t) + E_i(t), \quad (1)$$

This field drops to a residual field $\pm E_{res}$ ($+ E_{res}$ if $E_{tot}(t) > 0$ and vice versa), where $E_o(t)$ is the external driving field. The drift/recombination process changes E_i according to the following differential equation:

$$\dot{E}_i(t) = f(E_i(t), E_o(t)). \quad (2)$$
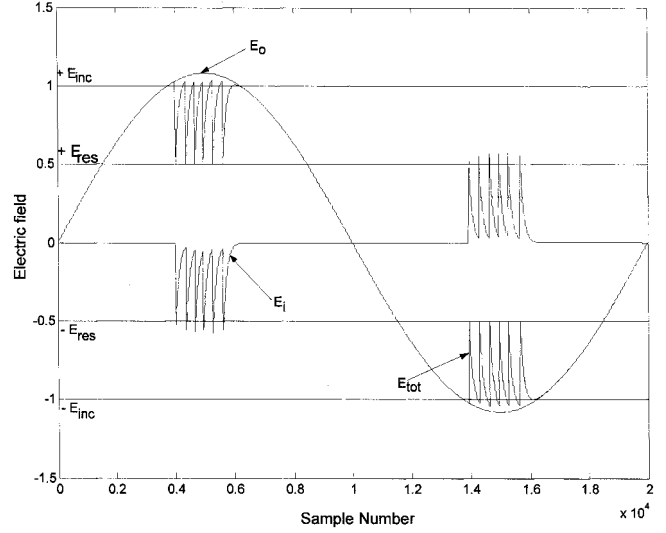
Figure 1. Typical electric field plots for the AC stochastic model.

If the exact way in which $E_i(t)$ decays is not very important, we can take

$$\dot{E}_i(t) = \frac{-E_i}{\tau_{dec}}. \quad (3)$$

Both processes, i.e., discharge (jump of E_i) and drift/recombination (continuous change of E_i) are coupled by the discharge probability in the following way. Let $c \cdot dt$ be the probability that a fast discharge occurs in the time interval $\{t, t + dt\}$. We have,

$$c = c(E_i(t), E_o(t)). \quad (4)$$

The value of c is zero when E_{tot} is less than E_{inc} (the inception field) and for $E_{tot} > E_{inc}$, it depends on the mechanism of first electron supply. Hence, the PD process is completely determined by the parameters c, E_{res}, f ($E_i(t)$, $E_o(t)$) and $E_o(t)$. Typical (random) trajectories of $E_i(t)$ and $E_{tot}(t)$ for a corona discharge are shown in Figure 1. The parameter values used in the simulation are shown in Table 1.

2.2   PD PULSE SHAPE

Having modelled the location of partial discharge occurrences, we need to simulate the shapes of PD pulses as well. It has been shown in the literature [8, 12] that real PD pulses can be represented by a double-exponential function of the form

$$f(t) = \exp(-t/\tau) - \exp(-2t/\tau). \quad (5)$$

The time constant τ is chosen to be 5 μs.

Table 1. Parameter values used in simulation (Corona discharge).

c (s^{-1})	τ(s)	$\hat{E}_o$	E_{res}
$((E_{tot}(t)-1)/0.025)^3$	120×10^{-6}	1.08	0.5

Note: $\hat{E}_o$ and E_{res} are given in terms of E_{inc}.

3   NOISE MODEL

On-site PD measurements are seriously affected by interference signals arising from different sources. Thermal or resistor noise is added by the measuring circuit, whereas the high-frequency sinusoidal signals from radio broadcasts and/or carrier wave communications are electromagnetically coupled and their amplitude is usually higher than that of the PD signal itself.

A typical on-site PD measurement system is modelled in [7]. According to this model, the PD pulses are affected by white Gaussian noise and high-frequency sinusoidal noise. The sinusoidal noise component, also referred to as discrete spectral interference (DSI), is made up of 10-11 high-frequency sinusoids at different power levels.

4   ACQUISITION OF REAL PD DATA

The real data for corona discharge was acquired with the setup shown in Figure 2. The test object, in this particular instance happens to be a point-plane gap to produce corona pulses. This is a very simple type of PD with periodic and regular pulses, also termed as 'trichel pulses'. The PD current pulse is generated and starts flowing through C_b and Z_m, and is sensed by the measuring impedance Z_m. Some part of the current flows through the transformer also. Here, a PD-free transformer is used so that the PD pulses due to the test object alone can be gathered. Since the voltages used are low (in the range of 3 to 10 kV) such a PD-free condition is easily met. The input to the HV transformer is controlled using a variac. In reality, the bad contacts of the variac, harmonics and extraneous pulsed interferences generated elsewhere in a power system (e.g., from thyristor drive circuits, arcing contacts, SMPS, etc.) are conductively coupled into the measuring circuit. So, sometimes a series high frequency (HF) filter is used at the input, and also a parallel rejection filter is used in series between the HV transformer output and the test object. However, since our goal was to study interferences, such a filter was not used in this case.
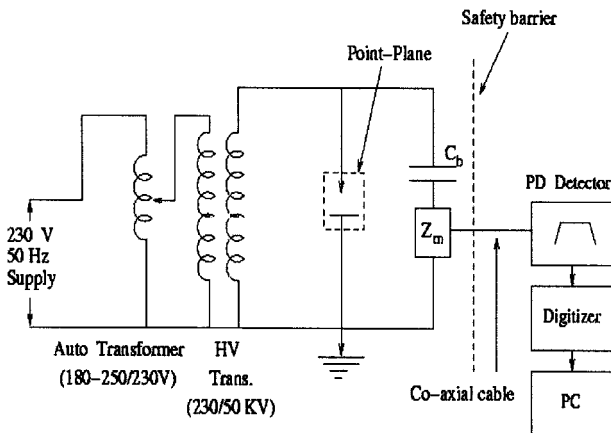

Figure 2. Test setup for acquiring real PD data.

The measuring impedance Z_m is a passive parallel RLC combination, which has a bandpass characteristic. This, together with an active bandpass filter (the heart of the PD detector)-cum amplifier, performs the task of 'quasi-integration' and assists in separating the PD pulses (of very low amplitude) buried in the capacitive displacement currents of relatively large magnitudes. The f1 and f2 values of the bandpass filter (which are switchable) were set to 10 and 300 kHz, respectively. The peak of the quasi-integrated pulse is proportional to the charge of the PD pulses (due to the calibration procedure), and hence is used as a measure of insulation degradation. These repetitive quasi-integrated PD pulses occur in both half cycles of the applied sine wave. The digitizer samples these and acquires them, which in turn are transferred into a personal computer (PC) via the general purpose instrumentation bus (GPIB). We have used the HP VEE graphical programming environment to control the transfer of data.

5   DENOISING TECHNIQUES

Various techniques have been mentioned in the literature to extract the PD pulses from noisy data. In addition to this, we have also implemented time-frequency analysis methods, such as STFT and WVD-based methods. The different DSP techniques investigated in this paper are listed in Table 2.

A brief description of these denoising techniques is given below:

5.1   FAST FOURIER TRANSFORM (FFT)-BASED DENOISING

Given a sequence of N samples $f(n)$, indexed by $n = 0,1,...,N$-1, the discrete Fourier transform (DFT) is defined as

$$F(k) = \sum_{n=0}^{N-1} f(n)e^{-j2\pi kn/N}, \, k = 0,1,...,N-1. \quad (6)$$

$F(k)$ are often called the 'Fourier Coefficients' or 'Harmonics'. The sequence $f(n)$ can be calculated from $F(k)$ using the inverse discrete Fourier transform (IDFT):

$$f(n) = \frac{1}{N} \sum_{k=0}^{N-1} F(k)e^{j2\pi kn/N}, \, n = 0,1,...,N-1. \quad (7)$$

The fast Fourier transform (FFT) is a DFT algorithm that reduces the number of computations for N points from $2N^2$ to $2N \log_2(N)$. In FFT-based signal denoising [9], the FFT of the signal is taken, and the FFT coefficients above a certain threshold are dropped. Then, the signal is reconstructed using an inverse FFT (IFFT). This method is mainly used to remove the sinusoidal noise, which has pronounced peaks in the frequency domain. We have used two kinds of FFT-based techniques, one with a constant

Table 2. Denoising Techniques.

Method No.	Denoising Technique
1a	Fast Fourier Transform (FFT), constant threshold
1b	Fast Fourier Transform (FFT), frequency-dependent threshold
2a	Low-pass filtering (Butterworth filter)
2b	Low-pass filtering (Chebyshev filter)
2c	Low-pass filtering (Inverse Chebyshev filter)
2d	Low-pass filtering (Elliptic filter)
3a	Wigner-Ville Distribution (WVD)
3b	Wigner-Ville Distribution (WVD)
4a	Short-Time Fourier Transform (STFT)
4b	Short-Time Fourier Transform (STFT)
5a	Least Mean Squares (LMS)
5b	Leaky LMS
5c	Sign-error LMS
5d	Sign-data LMS
5e	Sign-sign LMS
5f	Normalised LMS
5g	Kurtosis-driven LMS
5h	Kurtosis-driven LMS
5i	Adaptive Recursive LMS
5j	Cascade adaptive filtering
6a	Frequency-Domain Adaptive Filtering (FDAF) using DFT
6b	Frequency-Domain Adaptive Filtering (FDAF) using DCT
6c	Frequency-Domain Adaptive Filtering (FDAF) using DWT
7a	Recursive Least Squares (RLS)
7b	Exponentially-Weighted Recursive Least Squares (EWRLS)
8	Matched Filtering
9a	Notch filtering, algorithm 1, direct implementation
9b	Notch filtering, algorithm 1, lattice filter implementation
9c	Notch filtering, algorithm 2, direct implementation
9d	Notch filtering, algorithm 2, lattice filter implementation
10a	Wavelet-based (Thresholding)
10b	Wavelet-based (Mallat's algorithm)

threshold for all frequencies (Method 1a) and the other with the threshold varying according to the formula,

$$threshold = (j-1)^2 * 60 + 10, \qquad j = 1,...,l, \qquad (8)$$

where an l − point FFT is taken (Method 1b).

5.2 LOW PASS FILTERING

The PD pulses are concentrated in the low frequency region, whereas the sinusoidal components are usually in the high frequency region. Hence, for low levels of white noise, a low pass filter can extract the PD pulses efficiently. The different kinds of digital low pass filters [13]

considered in this paper are shown in Table 3. The cut-off frequency specified is the normalised value (a value of 1 corresponds to half the sampling rate). Increasing the filter order improves the performance of the filter.

5.3 WIGNER-VILLE DISTRIBUTION (WVD)-BASED DENOISING

The WVD of a signal is a time-frequency representation of the signal [14]. It is given by

$$W_x(n,m) = \frac{1}{2N} \sum_{k=0}^{N-1} x_k x_{n-k}^* e^{j\pi m(2k-n)/N}. \qquad (9)$$

In [15], the signal is segmented into blocks and the WVD of each block ($W_x(t, f)$) is taken and processed as follows. At every time t, $W_x(t, f)$ is integrated in a small range about a centre frequency f i.e., $(f - \Delta f, f + \Delta f)$. If this value exceeds a pre-set threshold, it indicates the presence of a signal at time t. The basis of this method is the fact that the WVD is a quadratic time-frequency representation, as evidenced by the following equation:

$$\int_{-\infty}^{\infty} W_x(t,f)\,df = |x(t)|^2 \qquad (10)$$

In this method (Method 3a), the signal is divided into blocks of length 100, Δf is taken to be 20 Hz and the threshold chosen is 0.3.

An alternative WVD-based method (Method 3b) involves masking of the time-frequency response [16]. As before, the signal is divided into blocks of length 100 and $W_x(n,\omega)$ is calculated. This is then multiplied by $H(n,\omega)$, which is a 2-D matrix with ones in the frequency range of the signal and zeroes elsewhere. This has the effect of masking the noise components.

$$Y(n,\omega) = W_x(n,\omega) H(n,\omega) \qquad (11)$$

If the desired signal, $s(n)$, satisfies $s(n) = 0$, $n < n_o$ and $n \geq n_o + L$, then it can be reconstructed from $Y(n,\omega)$ by the synthesis algorithm given in [16].

5.4 SHORT-TIME FOURIER TRANSFORM (STFT)-BASED DENOISING

The STFT of a signal is another time-frequency representation of the signal [14]. The expression for the STFT is given by:

$$S(n,m) = \sum_{k=0}^{N-1} x(k) w(n-k) e^{-j2\pi km/N}, \qquad (12)$$

Table 3. Low-pass filter characteristics.

Method No.	Filter type	Order	Cut-off frequency	Passband	Stopband
2a	Butterworth	3	0.5/π	−	−
2b	Chebyshev	3	0.5/π	0.5 dB ripple	−
2c	Inverse Chebyshev	3	0.5/π	−	ripple 20 dB down
2d	Elliptic	3	0.5/π	0.5 dB ripple	ripple 20 dB down

which is nothing but the DFT of $x(k)w(n-k)$. Here $w(n)$ represents the finite-length window. Usually a Hanning window is used. The length of the window determines the time and frequency resolutions of the STFT. Reducing this length produces a reduction in the frequency resolution and an increase in the time resolution. Increasing the length has the opposite effect.

The first denoising scheme (Method 4a) involves calculating the STFT of the input signal, multiplying it by a 2D array of ones and zeroes (masking) and then reconstructing the signal from the modified STFT [17]. One such reconstruction method (which minimizes the distance measure) is given by

$$x(n) = \frac{\sum_{m=-\infty}^{\infty} w(m-n)f_m(n)}{\sum_{m=-\infty}^{\infty} w^2(m-n)}, \quad (13)$$

where $f_m(n) = \text{IDFT}(S(n,k))$, for $n = m$.

The second method (Method 4b) involves estimating the signal from the modified STFT magnitude [17]. First the STFT magnitude of the input signal is calculated. Then, a first estimate of the actual signal x_i (usually white noise, i.e., mean 0 and variance 1) is made, its STFT is calculated and that STFT magnitude is replaced by the available STFT magnitude. Then, the signal is estimated using the previous method (13) and this is taken as the next estimate of the signal, x_{i+1}. The above process is then repeated for a total of 10 iterations.

5.5 LEAST MEAN SQUARES (LMS) APPROACH

The LMS method is an adaptive, iterative gradient search method, discussed in [18] and [19]. Its update equation is given by:

$$W(n+1) = W(n) + \mu \cdot e(n) \cdot X(n), \quad (14)$$

where W is filter coefficient vector, e an error vector, i.e., the difference between the desired filter output and the actual filter output and X an input vector.

The desired signal is taken to be either the input signal itself, or the input delayed by one sample. The first LMS-based denoising method (Method 5a) implements (14). Here, the length of the filter is chosen to be 10 and $\mu = 7 \times 10^{-7}$. Several variants of the LMS algorithm are also implemented in this paper.

1) *The leaky LMS algorithm [19]:* Here (Method 5b), the update equation is

$$W(n+1) = \gamma \cdot W(n) + \mu \cdot e(n) \cdot X(n). \quad (15)$$

γ is a constant leakage factor lying between 0 and 1. Leakage allows the impact on the filter coefficient vector of any single input sample to decay with time. The filter length is chosen to be 10, $\gamma = 0.9$ and $\mu = 7 \times 10^{-7}$.

The following three methods [19] reduce the computational cost but are not as efficient.

2) *Sign-error LMS or Pilot LMS:* Here (Method 5c), the update equation is

$$W(n+1) = W(n) + \mu \cdot \text{sgn}[e(n)] \cdot X(n). \quad (16)$$

The filter length is chosen to be 10 and $\mu = 3 \times 10^{-6}$.

3) *Sign-data LMS or Clipped LMS:* Here (Method 5d), the update equation is

$$W(n+1) = W(n) + \mu \cdot e(n) \cdot sgn[X(n)]. \quad (17)$$

The filter length is chosen to be 10 and $\mu = 3 \times 10^{-20}$.

4) *Sign-sign LMS or Zero-forcing LMS:* Here (Method 5e), the update equation is

$$W(n+1) = W(n) + \mu \cdot sgn[e(n)]sgn[X(n)]. \quad (18)$$

The filter length is chosen to be 10 and $\mu = 3 \times 10^{-6}$.

5) *Normalized LMS [18]:* This method (Method 5f) is used to prevent instability in the LMS algorithm:

$$W(n+1) = W(n) + \mu \cdot e(n) \cdot X(n) / (\gamma + X^T(n) \cdot X(n)) \quad (19)$$

Here the filter length is 10, $\mu = 7 \times 10^{-10}$ and $\gamma = 0$ [18].

6) *Kurtosis driven LMS [20]:* Two methods are employed here. Both use the 'kurtosis' of the error which is defined as $C = E(e^4(n)) - 3E(e^2(n))$. Method 5g uses the following update equation:

$$W(n+1) = W(n) + \mu \cdot |C| \cdot e(n) \cdot X(n). \quad (20)$$

Here, the length of the filter is 10 and $\mu = 7 \times 10^{-8}$ [20]. Method 5h uses the following update equation:

$$W(n+1) = W(n) + \mu_{max} \cdot (1 - e^{-\alpha|C|}) \cdot e(n) \cdot X(n). \quad (21)$$

We have used a filter length of 10, $\alpha = 7 \times 10^{-10}$ and $\mu_{max} = 100$.

7) *Adaptive recursive LMS [19] (Adaptive IIR filtering):* This method (Method 5i) is governed by the following set of equations:

$$y(n) = \sum_{k=0}^{N_F} a_k x(n-k) + \sum_{k=1}^{N_B} b_k y(n-k) \quad (22)$$

$$A^T = [a_o, a_1, \cdots, a_{N_F}] \quad (23)$$

$$B^T = [b_1, b_2, \cdots, b_{N_B}] \quad (24)$$

$$A(n+1) = A(n) - 2k_1 e(n) X(n) \quad (25)$$

$$B(n+1) = B(n) - 2k_2 e(n) Y(n) \quad (26)$$

Here $N_B = N_F = 30$, $k_1 = 10^{-5}$ and $k_2 = 10^{-10}$.

8) *Cascade adaptive filtering [21]:* The convergence can be improved by cascading a number of adaptive filters

(Method 5j). If l_k denotes the length of each filter, M denotes the number of stages, and $x_k(n)$ the input to stage k then, the following equations define the system:

$$x_1(n) = X(n) \qquad (27)$$

$$x_{k+1}(n) = e_k(n) \qquad (28)$$

$$e_k(n) = x_k(n) - \sum_{k=1}^{l_k} w_k^m(n)x_k(n-m), \qquad m = 1, \cdots, M \qquad (29)$$

We have used $M = 10$, $l_k = 2$ and $\mu = 7 \times 10^{-7}$.

5.6 FREQUENCY-DOMAIN ADAPTIVE FILTERING (FDAF) [18, 19]

Adaptive filtering can be done in the frequency domain (FDAF) as well, in which case it is computationally efficient compared to the LMS algorithm. The update equation in this case is given by

$$W_i(n+1) = W_i(n) + \mu \cdot e(n) \cdot X_i^*(n), \text{ for } i = 1, ..., N, \qquad (30)$$

where W, e, X are the DFTs or discrete cosine transforms (DCTs) or discrete wavelet transforms (DWTs) of the corresponding quantities in the time-domain and N is the filter length. The DCT of an N-point sequence, $x[n]$, is given by [18]:

$$C[k] = \sum_{n=0}^{N-1} x[n] \cos\left(\frac{(2n+1)k\pi}{2N}\right) \qquad k = 0, ..., N-1. \qquad (31)$$

The DCT is a real transform unlike the DFT. The DWT is discussed in Section 5.10. It can be shown that the LMS algorithm requires $2N^2 + N$ real multiplications, while the FDAF (using the FFT) requires $3N \log_2(N/2) + 4N$ multiplications.

We have implemented three kinds of frequency domain adaptive filters. All of them use $N = 8$, $\mu = 7 \times 10^{-7}$. Method 6a uses the DFT, method 6b uses the DCT and method 6c uses the DWT (with a Daubechies wavelet - db9).

5.7 RECURSIVE LEAST SQUARES (RLS) AND EXPONENTIALLY WEIGHTED RECURSIVE LEAST SQUARES (EWRLS) METHODS [19]

The weaknesses of the LMS method are:

1) It is slow.

2) It exhibits generally non-uniform convergence.

To overcome these pitfalls, the recursive least squares (RLS) method (Method 7a) has been proposed. This method can be described by the following recursion procedure:

$$R^{-1}(n) = R^{-1}(n-1) - \frac{R^{-1}(n-1)X(n)X^T(n)R^{-1}(n-1)}{1 + X^T(n)R^{-1}(n-1)X(n)} \qquad (32)$$

Initialize $W(0)$, $R^{-1}(-1)$. Then for $n = 0, 1, ...$

$$e(n) = d(n) - W^T(n)X(n) \qquad (33)$$

$$W(n+1) = W(n) + \alpha \cdot R^{-1}(n) \cdot e(n) \cdot X(n) \qquad (34)$$

We have used a filter length of 10 and α (the weighting factor) $= 7 \times 10^{-7}$.

Another related method (Method 7b), which has been proposed, is the exponentially weighted RLS method (EWRLS). Its recursion equations are:

$$e(n) = d(n) - W^T(n-1)X(n) \qquad (35)$$

$$\alpha(n) = \frac{1}{\lambda + X^T(n)R^{-1}(n-1)X(n)} \qquad (36)$$

$$W(n+1) = W(n) + \alpha(n) \cdot R^{-1}(n-1) \cdot e(n) \cdot X(n) \qquad (37)$$

$$R^{-1}(n) = \frac{R^{-1}(n-1) - \alpha(n)R^{-1}(n-1)X(n)X^T(n)R^{-1}(n-1)}{\lambda} \qquad (38)$$

Usually $0 < \lambda \leq 1$. $\lambda < 1$ implies that more weight is given to the most recent errors. This is useful in the case of non-stationary data. We have used a filter length of 10 and $\lambda = 0.95$.

Even though the RLS and EWRLS methods converge faster than the LMS method, it should be noted that they are computationally more expensive.

5.8 MATCHED FILTERING [8]

A matched filter arrangement, where the impulse response of the filter $h(t)$ is matched to the signal input $\varphi(t)$, is described by

$$h(t) = \varphi(T_o - t), \qquad (39)$$

where T_o corresponds to the sampling period. The output signal-to-noise ratio (SNR) is maximum for a matched filter. Method 8 uses a matched filter.

5.9 NOTCH FILTERING

A notch filter can be used to remove the sinusoidal noise present in the signal. The transfer function for a notch filter [22] is given by:

$$H(z) = \frac{(1 - re^{j\theta}z^{-1}) \cdot (1 - re^{-j\theta}z^{-1})}{(1 - \alpha re^{j\theta}z^{-1}) \cdot (1 - \alpha re^{-j\theta}z^{-1})}$$

$$= \frac{1 + w_1 z^{-1} + w_2 z^{-2}}{1 + \alpha w_1 z^{-1} + \alpha w_2 z^{-2}}, \qquad (40)$$

where $w_1 = -2r\cos\theta$, $w_2 = r^2$ and α is a parameter which determines the bandwidth of the notch. In this method, the notch is implemented using an infinite im-

pulse response (IIR) lattice notch filter, whose transfer function is given by

$$H(z) = \frac{1 + k_o(1+k_1)z^{-1} + k_1 z^{-2}}{1 + a_o(1+a_1)z^{-1} + a_1 z^{-2}}. \qquad (41)$$

The constraints to be satisfied for this to be a notch filter are given in [22]. Two adaptation algorithms (1 and 2) proposed in [22] have been implemented in this paper. We have used $\lambda = 0.9023$ and $\alpha = 0.9025$. Methods 9a and 9b use adaptation algorithm 1, whereas methods 9c and 9d use adaptation algorithm 2. Methods 9a and 9c implement the filter directly, whereas the other two, implement lattice filters.

5.10   WAVELET DENOISING

An excellent introduction to wavelets is given in [23]. The wavelet expansion for a function $f(t)$ is given by

$$f(t) = \sum_k \sum_j a_{j,k} \psi_{j,k}(t), \qquad (42)$$

where both j and k are integer indices and the $\psi_{j,k}(t)$ are the wavelet expansion functions that usually form an orthonormal basis. The set of coefficients $a_{j,k}$ are called the *discrete Wavelet transform* (DWT) of $f(t)$.

Two denoising methods based on wavelets have been implemented in this paper. The first one uses thresholding of wavelet coefficients [23] and the second one uses Mallat's algorithm [7].

5.10.1   DENOISING BY THRESHOLDING [23]

Donoho has proposed the following scheme for denoising:

1)   Compute the discrete wavelet transform (DWT) of the noisy signal.

2)   Perform thresholding.

Hard thresholding:

$$\text{thresholded value } \delta_\lambda(t) = \begin{cases} x(t), & \text{if } |x(t)| > \lambda \\ 0, & \text{otherwise} \end{cases} \quad (43)$$

Soft thresholding:

thresholded value $\delta_\lambda(t)$

$$= \begin{cases} \text{sgn}(x(t)) \cdot (|x(t)| - \lambda), & \text{if } |x(t)| > \lambda \\ 0, & \text{otherwise} \end{cases} \quad (44)$$

3)   Reconstruct the signal by performing an inverse DWT (IDWT) with the thresholded coefficients.

This method (Method 10a) is based on the property of the wavelet coefficients, forming an unconditional basis. This means that the wavelet coefficients drop off rapidly and this property is useful in removing noise. The MATLAB® routine 'wden' was used for this method. The

threshold selection rule was based on the principle of Stein's unbiased risk. Soft thresholding was performed with no multiplicative threshold rescaling. Wavelet decomposition was performed at level 3 using the biorthogonal ('bior 4.4') wavelet.

5.10.2   DENOISING USING MALLAT'S ALGORITHM [7]

Here (Method 10b), using Mallat's algorithm, the signal is decomposed into several scales, the detail coefficients are neglected and the signal is reconstructed using only the approximation coefficients. This is done because significant information about the noise components is contained in the detail coefficients and hence by neglecting them, we eliminate the noise. A 3 level decomposition was performed using Daubechies wavelet ('db10').

6   RESULTS

The denoising methods have been evaluated by applying them to both simulated and real PD data. To obtain the simulated PD data, the time of occurrence of partial discharges is simulated using the stochastic PD model given in Section 2 and noise is added according to the model given in Section 3. 20,000 samples of the input signal are taken, with the separation between any two samples being 1 μs. Therefore, the sampling rate is 1 MHz. The methods are characterized on the basis of their mean square errors (MSEs) and the time taken to perform denoising. For a particular method, if X denotes the noisy PD sequence and Y denotes the denoised output sequence, the normalized MSE between the input and output sequences is defined as

$$\text{MSE} = \frac{\sum_{i=0}^{n} [Y(i) - X(i)]^2}{n \sigma_n^2}, \qquad (45)$$

where n is the number of samples for which X is defined (i.e., the PD pulse region), and σ_n^2 is the total noise power. As can be seen from equation (45), if the method rejects all the noise components, the MSE is unity. In other words, the closer the MSE of a method is to unity, the better the method is at extracting the PD pulses. We have calculated the MSE of a particular denoising method as the mean of the MSEs at four different noise power levels. The results are shown in Table 4.
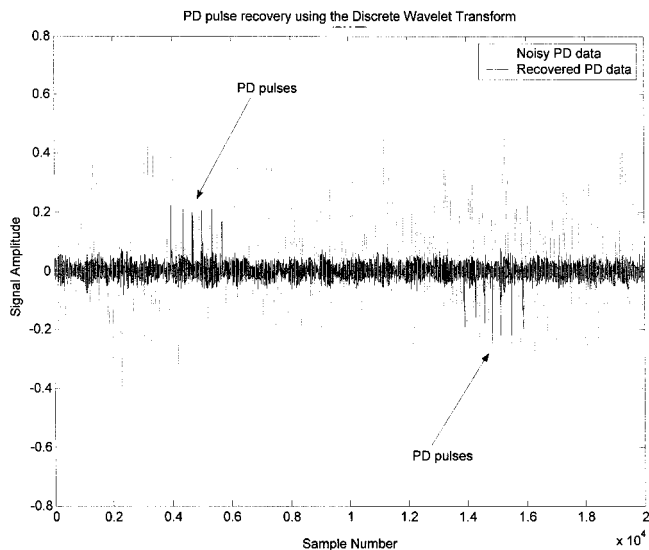
It can be seen from Table 4 that the MSEs of the low-pass filtering methods (Methods 2a-2d) and the wavelet denoising methods (Methods 10a, 10b) are closest to unity. The low-pass filtering methods are also the fastest ones (on the basis of computation time). These methods, though, will not perform effectively when low-frequency noise components are present. We conclude, therefore, that the wavelet denoising methods are the best in extracting the PD pulses. The performance of the wavelet de-
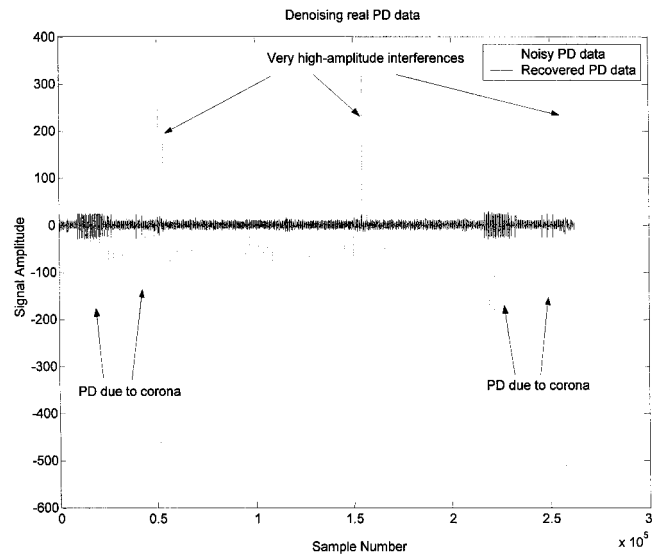
Table 4. Performance of the denoising methods.

Method	MSE	Time taken (s)
1a	0.9502	20.5645
1b	0.9536	20.4195
2a	0.9789	0.3978
2b	1.0045	0.1055
2c	0.9835	0.2428
2d	0.9809	0.7513
3a	2.8618	830.275
3b	1.0421	250.8
4a	1.0202	132.1875
4b	1.1629	788.01
5a	0.9659	5.558
5b	3.5526	5.8008
5c	0.9714	22.9203
5d	0.9669	24.118
5e	0.9652	24.769
5f	0.9484	13.2265
5g	0.9649	12.245
5h	0.9626	12.7538
5i	0.9539	23.6843
5j	2.9626	15.0965
6a	0.8751	8.4895
6b	0.8758	8.2545
6c	0.8735	106.618
7a	1.0600	7.2633
7b	0.9662	16.524
8	1.1772	19.3005
10a	0.9822	1.7298
10b	0.9870	53.885

Table 5. Performance of the notch filters.

Method	MSE	Time taken (s)
9a	0.9648	2.4083
9b	0.6602	2.2080
9c	0.9728	1.3470
9d	0.9693	1.3470

Figure 4. Notch filtering + wavelet-based denoising of real PD data.

noising method using Mallat's algorithm is illustrated in Figure 3. Methods 3a, 5b and 5j have the worst MSEs and are hence the worst methods. It is also observed that Methods 3b and 4b do not recover the sign of the PD pulses.

The notch filtering methods (Methods 9a-9d) are characterized by adding only sinusoidal noise. The performance of these methods is shown in Table 5. It is observed that the notch filtering methods using adaptation

algorithm 2 (refer Section 5.9), i.e., methods 9c and 9d perform much better. They also denoise the input signal faster.

The real PD data is considerably noisy, and it is difficult to distinguish between partial discharge pulses and noise. The notch filters (Methods 9c and 9d) best extract the PD pulses. This indicates a high sinusoidal noise content in the real PD data. As seen in Figure 4, the notch filter removes the low frequency, high amplitude interference. Additional wavelet denoising (Methods 10a, 10b) smoothens the signal, and makes it easier to locate the PD pulses. With this combined usage of notch filtering and wavelet denoising, partial discharge pulses are identified correctly. We conclude, therefore, that this combination can be used to extract PD pulses due to corona from a mixture of sinusoidal and white Gaussian noise, effectively.

Figure 3. PD pulse recovery using wavelet-based denoising (Mallat's algorithm).

7 CONCLUSION

In this paper, we looked at the problem of denoising PD signals caused by corona discharges. Several denoising methods were evaluated by applying them to both simulated and real PD data. The methods were characterized on the basis of their mean square errors and the execution time. It was observed that a combination of notch filtering and wavelet-based denoising performs well on both simulated and real PD data.

ACKNOWLEDGEMENT

REFERENCES

[1] *High Voltage Test Techniques - Partial Discharge Measurements*, IEC Standard 60270, 2000.

[2] H. Borsi, "A PD Measuring and Evaluation System Based on Digital Signal Processing", IEEE Trans. Dielect. Electr. Insul., Vol. 7, pp. 21–29, 2000.

[3] U. Schichler, H. Borsi and E. Gockenbach, "Application of Digital Filters for On-site Partial Discharge Detection on High Voltage Cables", IEEE Int. Conf. Properties and Applications of Dielectric Materials, Brisbane, Australia, pp. 618–621, 1994.

[4] A. Zargari and T. R. Blackburn, "Application of Adaptive Filters for the Estimation of Partial Discharge Signals in Noisy Environments", IEEE Int. Conf. Properties and Applications of Dielectric Materials, Seoul, Korea, pp. 212–215, 1997.

[5] U. Köpf and K. Feser, "Rejection of Narrow-Band Noise and Repetitive Pulses in On-site PD Measurements", IEEE Trans. Dielectr. Electr. Insul., Vol. 2, pp. 433–447, 1995.

[6] C. H. Knapp, R. Bansal, M. S. Mashikian and R. B. Northrop, "Signal Processing Techniques for Partial Discharge Site Location in Shielded Cables", IEEE Trans. Power Del., Vol. 5, pp. 859–865, 1990.

[7] W. Hang, T. Kexiong and Z. Deheng, "Extraction of Partial Discharge Signals using Wavelet Transform", 5th IEEE Int. Conf. Properties and Applications of Dielectric Materials, Seoul, Korea, pp. 322–325, 1997.

[8] M. W. J. van Hest, *Multi Resolution Analysis of Partial Discharge Signals*, M.S. thesis, Technische Universiteit Eindhoven, 1999.

[9] V. Nagesh, *Application of Digital Signal Processing Techniques to Partial Discharge Measurements*, Ph.D. dissertation, Indian Institute of Science, Bangalore, India, 1993.

[10] V. Nagesh and B. I. Gururaj, "Evaluation of Digital Filters for Rejecting Discrete Spectral Interference in On-Site PD Measurement", IEEE Trans. Elect. Insul., Vol. 28, pp. 73–85, 1993.

[11] C. Heitz, "A Generalized Model for Partial Discharge Processes based on a Stochastic Process Approach", J. Phys. D: Appl. Phys., Vol. 32, pp. 1012–1023, 1999.

[12] X. Dong, D. Zhu, C. Wang, K. Tan and Y. Liu, "Simulation of Transformer PD Pulse Propagation and Monitoring for a 500 kV Substation", IEEE Trans. Dielectr. Electr. Insul., Vol. 6, pp. 803–813, 1999.

[13] T. W. Parks and C. S. Burrus, *Digital Filter Design*, New York: John Wiley and Sons, 1987.

[14] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and Quadratic Time-Frequency Signal Representations", IEEE Signal Processing Mag., Vol. 9, no. 2, pp. 21–67, 1992.

[15] P. Rao and F. J. Taylor, "Detection and Localization of Narrow-Band Transient Signals using the Wigner Distribution", J. Acoustical Soc. America, Vol. 90, pp. 1423–1434, 1991.

[16] G. Faye, G. F. Boudreaux-Bartels and T. W. Parks, "Time-varying Filtering and Signal Estimation using Wigner Distribution Synthesis Techniques", IEEE Trans. Acoust., Speech, Signal Processing, Vol. 34, pp. 442–451, 1986.

[17] D. W. Griffin and J. S. Lim, "Signal Estimation from Modified Short-Time Fourier Transform", IEEE Trans. Acoustics, Speech, Signal Processing, Vol. 32, pp. 236–243, 1984.

[18] J. R. Treichler, C. R. Johnson Jr. and M. G. Larimore, *Theory and Design of Adaptive Filters*, New York: John Wiley and Sons, chs. 2–4, 1987.

[19] P. M. Clarkson, *Optimal and Adaptive Signal Processing*, CRC Press Inc., ch. 6, 1993.

[20] D. I. Pazaitis and A. G. Constantinides, "A Novel Kurtosis-Driven Variable Step-size Algorithm", IEEE Trans. Signal Processing, Vol. 47, pp. 864–872, 1999.

[21] P. Prandoni and M. Vetterli, "An FIR Cascade Structure for Adaptive Linear Prediction", IEEE Trans. Signal Processing, Vol. 46, pp. 2566–2571, 1998.

[22] N. I. Cho, C. H. Choi and S. U. Lee, "Adaptive Line Enhancement by using an IIR Lattice Notch Filter", IEEE Trans. Acoust., Speech, Signal Processing, Vol. 37, pp. 585–589, 1989.

[23] C. S. Burrus, R. A. Gopinath and H. Guo, *Introduction to Wavelets and Wavelet Transforms*, Prentice-Hall International, chs. 1, 2, 3 and 10, 1998.

S. Sriram was born in Bangalore, India in 1979. He received the B.Tech degree in electrical engineering from the Indian Institute of Technology, Madras, India, in 2001 and is expected to get his M.S. degree in electrical engineering from the same institute in 2005. He is currently working as a Hardware Design Engineer with Midas Communication Technologies Pvt. Ltd., Madras, India. His areas of interest include communication hardware design, FPGA design and signal processing.

S. Nitin received the B.Tech degree in electrical engineering from the Indian Institute of Technology, Madras, India, in 2001 and the M.S. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, in 2003, where he is now working towards the Ph.D. degree. His current research interests are in signal processing with emphasis on image and video processing algorithms and quality Evaluation techniques.

K. M. M. Prabhu (SM'93) obtained the B.Sc. (Engg.) degree in electronics and communication engineering, the M.Sc. (Engg.) degree in applied electronics engineering and the Ph.D. degree in digital signal processing in 1971, 1973 and 1981, respectively. He joined the Department of Electrical Engineering, Indian Institute of Technology (IIT) Madras, in 1981, where he is currently a Professor. While at IIT Madras, he has been a Principal Investigator for many projects funded by the Defense Research and Development Organization. He has published more than 110 papers in international journals and international conference proceedings. He spent a sabbatical, from August 1999 to August 2000, with the Signal Processing Systems Group, Eindhoven University of Technology, Eindhoven, The Netherlands. He has been a reviewer for a number of IEEE, IEE and other international journals. His current research interests include DSP algorithms and their implementation, digital filter structures, time-frequency signal analysis and wavelet-based signal processing.

M. J. Bastiaans (S'69-M'69-SM'85) received the M.Sc. degree in electrical engineering (with honours) and the Ph.D. degree in technical sciences from the Technische Universiteit Eindhoven (Eindhoven University of Technology), Eindhoven, The Netherlands, in 1969 and 1983, respectively. In 1969 he was an Assistant Professor and since 1985 an Associate Professor with the Department of Electrical Engineering, Technische Universiteit Eindhoven, in the Signal Processing Systems Group, where he teaches electrical circuit analysis, signal theory, digital signal processing, and Fourier optics and holography. His main current research interest is in describing signals by means of a local frequency spectrum (for instance, the Wigner distribution function, the windowed Fourier transform, Gabor's signal expansion, etc.). He is the author and co-author of about 150 papers in international scientific journals and proceedings of scientific conferences. Dr. Bastiaans is a Fellow of the Optical Society of America.