An Automatic Tool for Partial Discharge De-noising via Short Time Fourier Transform and Matrix Factorization

An Automatic Tool for Partial Discharge De-noising via Short Time Fourier Transform and Matrix Factorization

Yuan Yan（颜源）, Riccardo Trinchero, Igor Simone Stievano, Hongjie Li（李洪杰）, and Yanzhao Xie（谢彦召）

***Abstract*—This paper develops a fully automatic tool for the de-noising of partial discharge (PD) signals occurring in electrical power networks and recorded in on-site measurements. The proposed method is based on the spectral decomposition of the PD measured signal via the joint application of the short-time Fourier transform and the singular value decomposition. The estimated noiseless signal is reconstructed via a clever selection of the dominant contributions, which allows us to filter out the different spurious components, including the white noise and the discrete spectrum noise. The method offers a viable solution which can be easily integrated within the measurement apparatus, with unavoidable beneficial effects in the detection of important parameters of the signal for PD localization. The performance of the proposed tool is first demonstrated on a synthetic test signal and then it is applied to real measured data. A cross comparison of the proposed method and other state-of-the-art alternatives is included in the study.**

***Index Terms* — electrical networks, partial discharge, de-noising, short time Fourier transform, singular value decomposition.**

I. INTRODUCTION

CONCEPT of smart grid is now a robust and well defined design format allowing the bi-directional interactions between electric utilities and power components in grids. This enables the real-time monitoring of the healthy condition of the different interconnected blocks, thus allowing the development of tools for optimizing their maintenance and operation actions [1]. In the above scenario, partial discharge (PD) can be considered one of the most harmful insulation aging factors playing a crucial role in the healthy condition assessment of power components [2].

Nowadays, many commercial digital integrated circuits with advanced high-speed data acquisition features are available, and they also integrate with ultra-wide band sensors. These sensors allow achieving advanced PD diagnosis, including PD localization, recognition, and classification [3-5]. However, on-site PD measurements turn out to be susceptible to field noise interferences, with unavoidable detrimental effects of both white noise and discrete spectrum disturbance [6]. Therefore, de-noising is an essential step in on-site PD measurement. However, high-level noise, variable waveforms of PD signals, and limited computational resource in most commercial devices make on-site PD de-noising yet challenging.

In last two decades, many papers have been published on alternative pulse signal (e.g., PD) de-noising approaches, being wavelet transform (WT), empirical mode decomposition (EMD), local mean decomposition (LMD), and variational mode decomposition (VMD). In [7-12], WT has proven to have an outstanding PD de-noising performance for extreme noisy signals if the mother wavelet and decomposition level are selected properly. Prior knowledge of the characteristics of the PD signal is however required to determine the mother wavelet and decomposition level, which is always difficult to have in field due to variable waveforms of PD signals. In [13,14] iteration-based algorithms are proposed to automatically determine the optimal mother wavelet and decomposition level. Despite the mentioned nice automatic feature of these methods, a common unfavorable aspect of these methods is their low efficiency, which means large computational time. In [15,16] the adaptative EMD and LMD algorithms are proposed. Their main advantage is that no prior knowledge is required, but they suffer from mode mixing problems, possibly leading to signal distortion. In [17] and [18], ensemble EMD and LMD are developed to overcome the problem, but they require many iterative cycles and calculations. In [19], VMD is proposed to specifically address this issue, and it has successfully been applied for de-noising PD or fault signals [20, 21]. However, the success of VMD depends highly on the pre-set parameters, including the mode number and bandwidth control parameter.

To remedy these deficiencies in above de-noising approaches, the non-parametric and self-adaptive methods based on singular value decomposition (SVD) have been explored in recent years [22-28]. In [22], Abdel-Galil et al. carried out a pioneer study of de-noising via SVD. In [23], an adaptive principal components selection algorithm is developed to automate the SVD-based PD de-noising process. In [24, 25], some fast SVD algorithms are developed by reducing the rank of the input matrix to be decomposed. In [26, 27], some hybrid methods

* This work was supported in part by China Scholarship Council. (corresponding author: igor.stievano@polito.it)

Yuan Yan is with State Key Laboratory of Insulation and Power Equipment, School of the Electrical Engineering, Xi'an Jiaotong University, Xi'an, 710049, China, and also with Department of Electronics and Telecommunications, Politecnico di Torino, Turin, 10129, Italy. (y220928y@stu.xjtu.edu.cn)

Riccardo Trinchero, and Igor Simone Stievano are with the Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy. (e-mail: riccardo.trinchero@polito.it; igor.stievano@polito.it).

Hongjie Li, and Yanzhao Xie are with State Key Laboratory of Insulation and Power Equipment, School of the Electrical Engineering, Xi'an Jiaotong University, Xi'an, 710049, China. (e-mail: hjli@mail.xjtu.edu.cn, yzxie@xjtu.edu.cn)

combining SVD with WT or EMD are presented, where SVD is used to remove severe discrete spectrum noise components with significant larger singular values. In all the above papers, the SVD-based de-noising scheme is done as follows: first, a noisy PD signal is embedded into a Hankel matrix; then, applying SVD to the Hankel matrix results into a singular value vector and multiple components; finally, the components with larger singular values are used to reconstruct the noiseless PD or the unwanted discrete spectrum noise that needs to be subtracted. However, the investigation in [28] shows that the de-noising scheme does not perform well for signals with very low signal-to-noise ratio. In this case, several floors appear in the singular value plot, which makes it very difficult to select the optimal value of the number of dominant contributions. In addition, when the singular values of the PD signal and of the discrete spectrum noise have the same magnitude, the de-noising scheme can hardly differentiate them. In [29], a PD de-noising based on the so-called generalized S-transform and module time-frequency matrix is proposed. In this approach, the de-noising performance strongly depends on the accuracy of the estimated frequency of the discrete spectrum noise, which requires high frequency resolution of the S-transform and thus results in a high-rank matrix, which in turn may lead to large computational time.

To overcome the above limitations and offer a de-noising solution with improved accuracy and efficiency at the same time, this paper proposes an alternative approach that jointly combines the SVD and the time-frequency complex matrix (or spectrogram) of a noisy PD signal obtained by the short-time Fourier transform (STFT). It can be verified that the latter transform and the time-frequency domain allow a clearer decomposition of the PD signal into its different constitutive components. The functional, i.e., noiseless, PD signal can be therefore estimated by filtering out the spurious and well-separated noisy terms. In addition, the procedure can be wrapped-up in terms of an automatic tool in which only two parameters are needed, one related to the frequency pass-band behavior of the used PD sensor and the other related to the field noise level. The performance of the developed de-noising algorithm is compared with some state-of-the-art alternatives, including the Hankel matrix-based adaptive SVD (H-ASVD), WT, and EMD algorithms. In addition, the proposed approach is first applied to a synthetic PD signal and then to real measured data to demonstrate its strengths and features. The obtained results also prove the improved features of the method in extreme conditions with very small signal-to-noise ratio, as it produces a better noiseless PD reconstructed signal.

In summary, the main contributions of this paper are:

- A well-defined and robust step-by-step procedure for de-noising PD signals is defined;
- An automatic tool with the smallest number of tuning parameters (two in the proposed scheme) is generated. The parameters are chosen with a simple rule, and the low sensitivity of the de-noising accuracy to their value has been confirmed.
- The feasibility and strengths of the proposed de-noising scheme on a synthetic signal are verified, as we have proved its generality for different possible shapes of PD signals and its robustness to large noise levels;
- An unbiased comparison with alternative state-of-the-art solutions is carried out;
- The proposed tool has been applied to real on-site measurements carried out by a high frequency current transformer and an ultra-high frequency antenna.

The rest of this paper is structured as follows. Section II discusses the modeling of typical PD signals and noise. In Section III, the proposed adaptive and automatic PD de-nosing tool is detailed. In Section IV, three de-noising performance metrics are introduced and adopted to quantify the effectiveness of the proposed algorithm on a synthetic PD test case. In the same section, a cross-comparison involving the mentioned state-of-the-art approaches is carried out. In Section V, the feasibility of the developed tool is confirmed by its application to real measured signals. Conclusions and final remark are given in Section VI.

II. PD AND NOISE MODELING

A. PD signal modeling

PD is a localized electrical discharge that only partially bridges the insulation between conductors, and it can or cannot occur adjacent to a conductor [30]. According to many laboratory studies, the PD process can be approximately equivalent to the *Townsend* discharge process, in which a current pulse is generated due to the movement of ionized electrons and ions under the stress of an external electric field [31]. The electrons move faster due to their lighter weight, leading to a fast-rising edge of the current pulse, whereas the ions move slowly due to their heavier weight, leading to a slow-falling edge.

Therefore, at the PD source, a PD signal start at time zero can be approximately modeled by a double exponential pulse (DEP), which can be formulated as [6]

$$DEP(t) = A_1 \cdot (e^{-t/\tau_1} - e^{-t/\tau_2}) \tag{1}$$

where A_1 is the amplitude, and τ_1 and τ_2 ($\tau_1 > \tau_2$) are the time constants. DEP signal is often detected in the line-type power equipment, such as cables and overhead lines. Most of detected PD signals are however oscillating pulses due to the effects of both the propagation path and the transfer function of the used sensor. Therefore, single exponential and double exponential attenuation oscillation pulses (SEOP and DEOP) are used [26]. They are formulated as:

$$SEOP(t) = A_2 \cdot e^{-t/\tau_3} \cdot \sin(2\pi f_{c1} t) \tag{2}$$

$$DEOP(t) = A_3 \cdot (e^{-t/\tau_4} - e^{-t/\tau_5}) \cdot \sin(2\pi f_{c2} t) \tag{3}$$

where A_2 and A_3 are the amplitude parameters, τ_3, τ_4, and τ_5 are time constants, and f_{c1} and f_{c2} are the oscillation frequencies.

In this work, the above three types of pulses (DEP, SEOP, and DEOP) are concatenated sequentially to generate the signal shown in Fig. 1(a). The above signal represents an initial

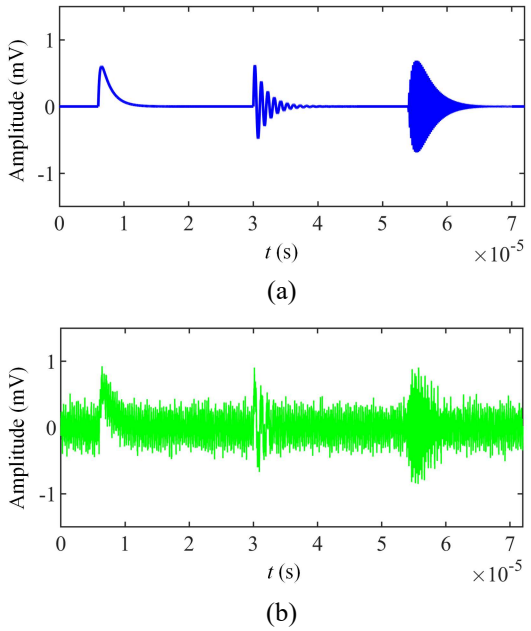


Fig. 1. Synthetic noiseless (a) and noisy (b) PD signal with signal-to-noise ratio of -1.55dB.

TABLE I
PARAMETERS OF THE SYNTHETIC PD SIGNAL OF FIG. 1

Parameter	A_1 /mV	A_2 /mV	A_3 /mV	τ_1 /ns	τ_2 /ns
Value	1	0.7	1.8	150	23
Parameter	τ_3 /ns	τ_4 /ns	τ_5 /ns	f_{c1} /MHz	f_{c2} /MHz
Value	200	77	230	1	5

synthetic test case which simulates the simultaneous occurrence of multiple PDs, which is a common phenomenon in field measurements. Also, it offers a fully controllable test case, which highlights and verifies the features and strengths of the proposed de-noising scheme for increasing levels of superimposed noise. The parameters of DEP, SEOP, and DEOP signals used in this paper are listed in Tab. I. The PD signal is uniformly sampled at a sampling frequency $f_s = 125\text{MHz}$ and stored in a vector. For the signals of Fig. 1, the number of sampling points is 9000.

B. Noise characteristics

In on-site PD measurements, the sensitivity and accuracy of PD detection are always reduced by two types of noise: (i) the white noise and (ii) the discrete spectrum noise. Discrete spectrum noise mainly arises from carrier communication, ratio communication, high-order harmonic, etc., whilst white noise is mainly caused by equipment thermal noise, ground noise, rand noise, etc. [32]. In this work, the frequencies of two harmonics in the discrete spectrum noise are set to 3 and 7 MHz and their corresponding amplitudes are set to 0.15 and 0.1 mV, respectively. White noise is a zero mean gaussian sequence with 0.1 mV standard deviation.

The above disturbances are added to the noise-free PD signal to generate the noisy signal shown in Fig. 1(b). It is important to point out that the synthetic noisy signal in the Figure is characterized by a critical low signal-to-noise ratio, and thus it is a challenging test case which can be effectively used to verify the benefits of the proposed de-noising solution and to compare with state-of-the-art alternatives.

III. PRINCIPLE OF THE PROPOSED DE-NOISING TOOL

This section presents the proposed automatic tool for PD de-noising according to the scheme in Fig. 2 and data processing following all the different involved blocks from left to right. It consists of six steps which are detailed in the following subsections.

A. Time-frequency transform via STFT (step 1)

According to the procedure of Fig. 2, the sampled noisy PD signal like the one shown in Fig. 1(b) is stored in a vector

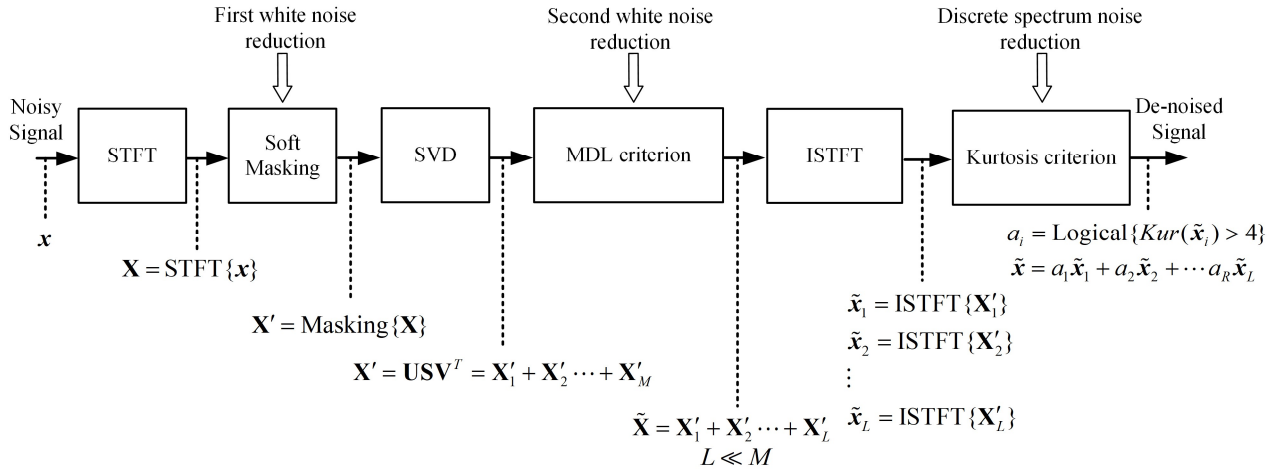


Fig. 2. Proposed step-by-step automatic procedure for PD de-noising.

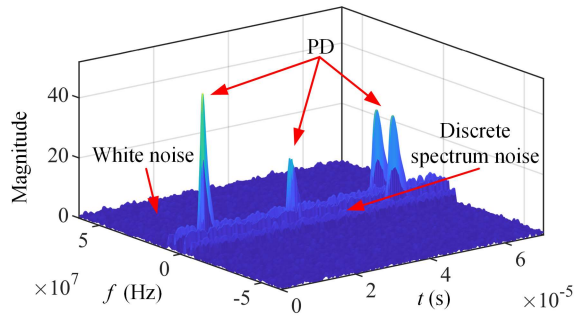


Fig. 3. Magnitude of the spectrogram of the synthetic noisy signal shown in Fig. 1(b) and location of the different signal components.

$\boldsymbol{x} = [x(t_1), x(t_2), \ldots, x(t_K)]^T$ defined as:

$$x(t_k) = s(t_k) + w(t_k),\ t_k = (k-1)\Delta t,\ k = 1, 2, \ldots, K \tag{4}$$

where $s(t_k)$ is the discretized noise-free PD signal, $w(t_k)$ represents the superimposed noise (e.g., white and discrete spectrum noise), $\Delta t = 1/f_s = 1/(125\text{MHz})$ is the sampling interval, and K is the number of time samples (e.g., 9000 in the example signal). Then, $\boldsymbol{x}$ is transformed into a spectrogram $\mathbf{X}$ via the STFT.

This step represents the most important initial data processing since the local features of a nonstationary PD signal cannot be well expressed in the time-domain or in the frequency-domain only. Hence, a time-frequency analysis is more suitable for representing a PD signal with its inherent pulsed characteristics.

STFT is a widely used time-frequency tool for studying nonstationary signals and it has been proven to be effectively used in this field [33]. The discrete STFT of the discrete signal $\boldsymbol{x}$ can be written as:

$$X(f_m, t_n) = \sum_{k=1}^{K} x(t_k) g(t_k - t_n) e^{-j2\pi t_k f_m} \Delta t \tag{5}$$

where $g(\cdot)$ is a Gaussian window function with length of M (e g., 200 in the following illustrative example in this section), $t_n = (n-1)\Delta t$ ($n = 1, 2, \cdots, N$, N=K-M+1) and $f_m = (2m - M)/(2M\Delta t)$ ($m = 1, 2, \cdots, M$) are the discrete time and frequency, respectively, and $X(f_m, t_n)$ (also labeled as $X_{m,n}$) is the element in the m^{th} row and n^{th} column of the spectrogram $\mathbf{X} \in \mathbb{C}^{M \times N}$. The absolute value of the spectrogram $\mathbf{X}$ obtained by applying STFT to the synthetic noisy signal in Fig. 1(b) is shown in Fig. 3, in which the PD pulses emerge clearly while they are drowned out by noise in time-domain.

B. Soft masking (step 2)

White noise is preliminarily reduced by applying the following soft masking to the spectrogram $\mathbf{X}$, resulting into an updated spectrogram $\mathbf{X}'$.

Figure 3 shows the magnitude of $\mathbf{X}$, highlighting that both the PD signal and the discrete spectrum noise turn out to be localized in the specific zones of the spectrogram characterized by a remarkably high amplitude (in magnitude). On the other hand, white noise produces a uniformly distributed background. This magnitude difference between the dominant signal (e.g., PD signal) and the white noise allows us to first preliminarily filter out the noise by applying the following soft masking on the spectrogram [34]:

$$X'_{m,n} = \begin{cases} T_{m,n} \cdot X_{m,n} & \text{if } |X_{m,n}| < 3\varepsilon \\ X_{m,n} & \text{if } |X_{m,n}| \geq 3\varepsilon \end{cases} \tag{6}$$

where ε is the standard deviation of $\mathbf{X}$. The threshold value of 3ε is suitably adjusted on the basis of the signal-to-noise ratio and is chosen according to [35], and $T_{m,n}$ are the attenuation coefficients defined as:

$$T_{m,n} = \left| \frac{X_{m,n}}{3\varepsilon} \right|^q \tag{7}$$

where q is the attenuation control factor. A large value of q increases the white noise suppression, but at the same time it leads to large distortion of the PD signal. Therefore, the value of q should be set carefully, and it will be thoroughly discussed in Section IV. In the following illustrative example in this section, q is set to 1. After the soft masking, the spectrogram $\mathbf{X}$ is updated to a new spectrogram $\mathbf{X}'$, in which parts of white noise is removed while discrete spectrum noise and some residual white noise still exist. Finally, it is important to point out that the soft masking improves the signal-to-noise ratio of the PD signal, thus facilitating the subsequent matrix factorization and component selection.

C. Matrix factorization via SVD (step 3)

In this step, the spectrogram $\mathbf{X}'$ is decomposed into M components via the singular value decomposition, allowing to split the original signal into multiple distinguishable components, in which the functional PD signal, the discrete spectrum noise and the remaining white noise can be separated. This can be also justified by the different localization of signal components in Fig. 3, which are characterized by different shapes and strengths, in magnitude.

Applying SVD to the spectrogram $\mathbf{X}'$ (with $M < N$) can generate three decomposed matrices, yielding:

$$\begin{aligned} \mathbf{X}' &= \mathbf{U}\mathbf{S}\mathbf{V}^H \\ &= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^H + \sigma_2 \boldsymbol{u}_2 \boldsymbol{v}_2^H + \cdots \sigma_M \boldsymbol{u}_M \boldsymbol{v}_M^H \\ &= \mathbf{X}'_1 + \mathbf{X}'_2 + \cdots + \mathbf{X}'_M \end{aligned} \tag{8}$$

where $\mathbf{U}$ is a $M \times M$ orthonormal matrix ($\mathbf{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_N]$, $\boldsymbol{u}_i \in \mathbb{C}^{M \times 1}$), $\mathbf{S}$ is a real $M \times N$ rectangular matrix with the singular values $\sigma_1 > \sigma_2 > \cdots > \sigma_M$ in the diagonal entries, $\mathbf{V}$ is $M \times M$ a square orthonormal matrix ($\mathbf{V} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_M]$, $\boldsymbol{v}_i \in \mathbb{C}^{M \times 1}$), and $\mathbf{X}'_i$ is the i^{th} decomposed component of $\mathbf{X}'$. PD de-noising can be achieved by selecting the components which best represent the contribution of the noiseless PD signal.

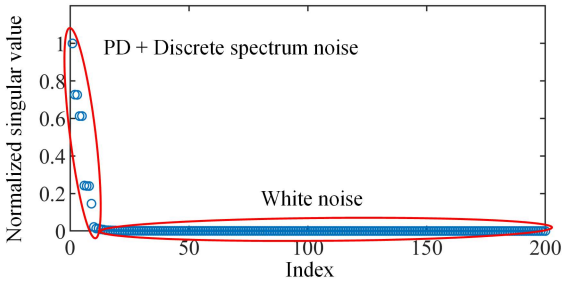


Fig. 4. Singular values of the spectrogram $\mathbf{X}'$.

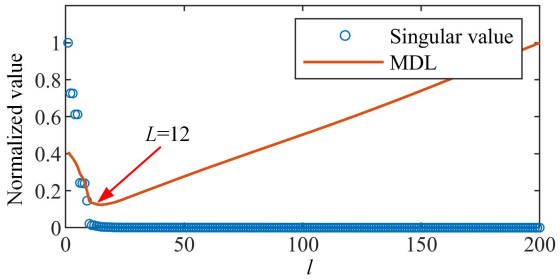


Fig. 5. MDL curve of the singular values.

Roughly speaking, it can be verified that the more concentrated and large-amplitude objects (e.g., the PD signal and the discrete spectrum noise) in the spectrogram $\mathbf{X}'$ correspond to the components with larger singular values, whereas the scattered and low-amplitude objects (e.g., the white noise) correspond to the components with small singular values as highlighted in Fig. 4. Therefore, the mix consisting of the PD signal and of the discrete spectrum noise can be possibly extracted by selecting the principal components associated to larger singular values, thus reducing the residual white noise.

D. Principal components selection via minimum description length (MDL) criterion (step 4)

This step implements an automatic rule for the selection of the optimal L ($L \ll M$) components arising from the dominant singular values as shown in Fig. 4.

Traditionally, the principal components associated to larger singular values in the SVD are selected via the threshold criterion $\sigma_{l+1}/\sigma_1 \geq \delta$ ($l=0,1,\ldots,M-1$), where δ is an artificial threshold value referring to the signal-to-noise ratio of the spectrogram $\mathbf{X}'$. Since we do not know the signal-to-noise ratio, which is different depending on the PD signal and noise level, we use the MDL criterion to automatically select the number of the principal components. The function of the MDL with repect to the singular values of the spectrogram $\mathbf{X}'$ is defined as [36,37]:

$$\mathrm{MDL}(l) = -N \log \left(\frac{\prod_{i=l+1}^{M} \sigma_i}{\left(\frac{1}{M-l} \sum_{i=l+1}^{M} \sigma_i \right)^{M-l}} \right) + \frac{1}{2} l(2M-l) \log N \quad (9)$$

where σ_i is the i^{th} singular value of $\mathbf{X}'$, M is the number of singular values (or rows) of $\mathbf{X}'$, and N ($N > M$) is the number of the columns of $\mathbf{X}'$. The number of effective singular values can be determined by minimizing MDL(l) as [36,37]:

$$L = \arg \min_{0 \leq l \leq M-1} \mathrm{MDL}(l) \quad (10)$$

To better explain the effect of the above criterion, the MDL curve of the singular value vector of Fig. 4 is shown in Fig. 5. L=12 corresponds to the minimum of the MDL curve, defining that only the first 12 components, instead of 200 in total, can successfully be used to represent the dominant contributions of the spectrogram $\mathbf{X}'$. Therefore, $\mathbf{X}'$ can be updated to a compressed spectrogram $\tilde{\mathbf{X}} = \mathbf{X}'_1 + \mathbf{X}'_2 \cdots + \mathbf{X}'_L$ ($\mathbf{X}'_i \in \mathbb{C}^{M \times N}$, $i = 1, 2, \ldots, L$) by selecting the first L components, thereby removing most of the remaining white noise.

E. Time-domain signal reconstruction via inverse STFT (step 5)

Each component in the spectrogram $\tilde{\mathbf{X}}$ is transformed back to time domain via Inverse STFT (ISTFT), resulting into L signal modes, in which the PD signal and the discrete spectrum noise are clearly distinguishable. The conversing back to time-domain is important not only for retrieving the estimated noiseless PD signal but also because, as observed before, the singular values of the PD signal and of the discrete spectrum noise may have the same magnitude, making their possible separation hard.

Hence, each component $\mathbf{X}'_i$ of the spectrogram $\tilde{\mathbf{X}}$ is converted back to time domain via inverse STFT:

$$\tilde{x}_i(t_k) = \frac{1}{M} \sum_{n=1}^{N} \sum_{m=1}^{M} X'_i(f_m, t_n) g(t_k - t_n) e^{j2\pi t_k f_m} \quad (11)$$

where $\tilde{x}_i(t_k)$ is the i^{th} reconstructed signal mode, and $X'(f_m, t_n)$ (also labeled as $X'_{m,n}$) is the element in the m^{th} row and n^{th} column of the sub-spectrogram $\mathbf{X}'_i$.

Again, for example signal of Fig. 1(b), the L reconstructed signal modes $\tilde{\boldsymbol{x}}_1$, $\tilde{\boldsymbol{x}}_2$, …, and $\tilde{\boldsymbol{x}}_L$ are shown in Fig. 6. From visual inspection, it can be observed that the PD signal and the discrete spectrum noise are spread in different contributions, or modes (e.g., #1, #4, #5, #8, #9, #10, #11, #12). This observation suggests that it is also possible filtering out the discrete spectrum noise and estimate the noiseless PD signal by selecting the most modes with an impulse-like shape. This is done in the next, i.e., last, step.

F. Mode selection via Kurtosis criterion (step 6)

The Kurtosis parameter is used to separate the contributions associated with the PD signal from those arising from the discrete spectrum noise in the L reconstructed modes. Kurtosis is a statistical measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution [38].

Qualitatively, sequences with high kurtosis parameter tend to have heavy tails, or outliers. Sequences with low kurtosis value tend to have light tails or lack outliers. In the reconstructed

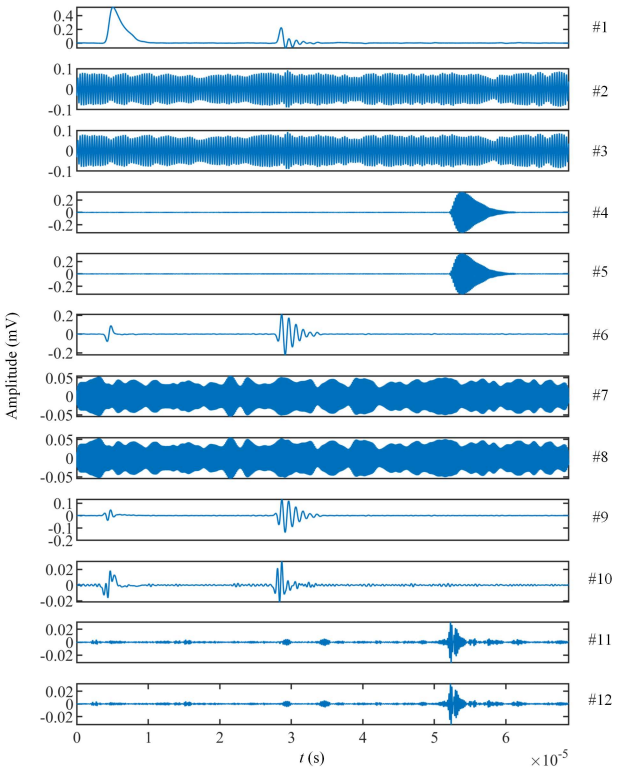


Fig. 6. Signal modes reconstructed from the first twelve components with larger singular values.

modes, the signal components associated with the pulses (e.g., PD signal) performs like some outliers, thus produce to a large kurtosis value, whereas periodic signals (e.g., the discrete spectrum noise) produce a low kurtosis value.

The Kurtosis value of a discretized signal is defined as:

$$Kur(\tilde{x}_i)=\frac{\frac{1}{K}\sum_{k=1}^{K}\left[\tilde{x}_i(t_k)-\bar{x}_i\right]^4}{\left(\frac{1}{K}\sum_{k=1}^{K}\left[\tilde{x}_i(t_k)-\bar{x}_i\right]^2\right)^2} \tag{12}$$

where $\tilde{x}_i$ is the i^{th} signal mode reconstructed by the i^{th} component of $\mathbf{X}'$, and $\bar{x}_i$ is the average of $\tilde{x}_i$. The calculated kurtosis values of the modes collected in Fig. 6 are shown in Fig. 7. From this figure, we can see the modes which are more closely related to the PD signal are these with larger kurtosis values (i.e., #1, #4, #5, #8, #9, #10, #11, #12).

The modes with larger kurtosis value are selected to reconstruct the final estimated noiseless signal $\tilde{\boldsymbol{x}}$, which can be rewritten as:

$$\tilde{\boldsymbol{x}}=a_1\tilde{\boldsymbol{x}}_1+a_2\tilde{\boldsymbol{x}}_2+\cdots a_L\tilde{\boldsymbol{x}}_L \tag{13}$$

where $a_i=\text{Logical}\{kur\{\tilde{\boldsymbol{x}}_i\}>\beta\}$,where Logical{.} is the logical judgment function and returns 1 or 0. β is a threshold set to 4, an empirical value suggested in [27, 38]. The final estimated noiseless signal $\tilde{\boldsymbol{x}}$ via (13) is shown in Fig 8, where most of white noise and discrete spectrum noise are successfully removed.

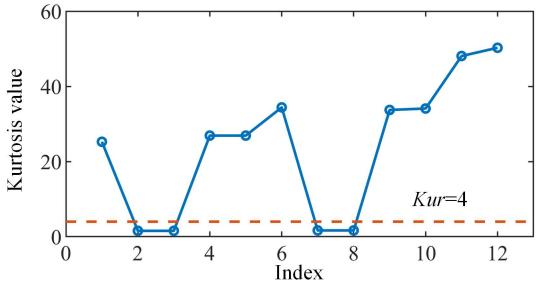


Fig. 7. Kurtosis value of the modes in Fig. 6.

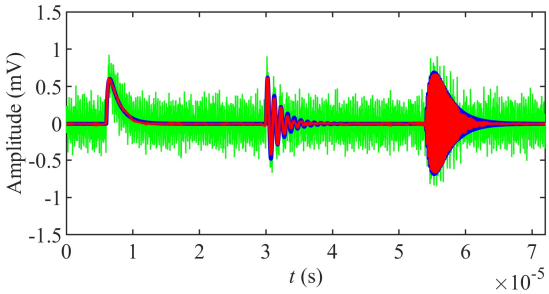


Fig. 8. De-noised signal obtained by summing modes #1, #4, #5, #8, #9, #10, #11, and #12 in Fig. 6. The green, blue, and red lines denote the noisy, the original, and the de-noised signals, respectively.

IV. Results: Synthetic PD Signals

This Section collects the results of the application of the proposed tool to the synthetic PD test case. The simulations are conducted using MATLAB software running on an Intel(R) Core i7-10750H processor with 2.60-GHz clock frequency and 16-GB RAM. In the simulations, three evaluation metrics are defined to quantitatively assess the de-noising performance of the proposed algorithm, which is compared with the alternative WT [7], EMD [17], and H-ASVD [23] tools. Also, as briefly mentioned before, in the proposed algorithm, only two parameters (the window length in the STFT and the attenuation control factor of the soft masking) need to be determined in advance; a rule for their choice is also presented.

A. Evaluation metrics

Three evaluation metrics are defined below to quantify the quality of the estimated noiseless PD signal [27] ($s(t_k)$ and $\tilde{x}(t_k)$ denote the discrete noise-free and de-noised signals, respectively).

1) Signal-to-noise ratio (*SNR*) is used to measure the background noise reduction:

$$SNR=10\cdot\log_{10}\frac{\sum_{k=1}^{K}s^2(t_k)}{\sum_{k=1}^{K}\left[\tilde{x}(t_k)-s(t_k)\right]^2} \tag{14}$$

2) Normalized correlation coefficient (*NCC*) is used to evaluate the similarity of waveform between the original and de-noised signals. It is defined as:

$$NCC = \frac{\sum_{k=1}^{K} s(t_k)\cdot \tilde{x}(t_k)}{\sqrt{\left[\sum_{k=1}^{K} s^2(t_k)\right]\cdot\left[\sum_{k=0}^{K-1} \tilde{x}^2(t_k)\right]}} \quad (15)$$

3) Root-mean-square error (*RMSE*) is used to evaluate the waveform distortion of the de-noised signal compared with the original signal:

$$RMSE = \sqrt{\frac{1}{K}\sum_{k=1}^{K}\left[\tilde{x}(t_k) - s(t_k)\right]^2} \quad (16)$$

Higher *SNR* and *NCC*, and lower *RMSE* represent a better de-noising performance.

B. Effect of the window length in STFT

The window length M in STFT, which is equal to the number of rows of the time-frequency spectrogram $\mathbf{X}$, is inversely proportional to the frequency resolution Δf of the spectrogram, which can be formulated as

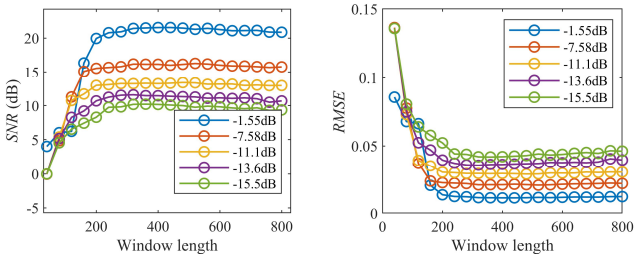
$$\Delta f = \frac{f_s}{M} \quad (17)$$

For a fixed value of the sampling frequency f_s, a small window length can lead to a large frequency resolution of the spectrogram, which may be insufficient to distinguish a PD signal from the noise, especially the discrete spectrum noise, whereas a large window length can lead to a small frequency resolution, resulting in redundant computations.

We de-noise the noisy PD signal synthesized in Section II via the proposed algorithm with varying window lengths and noise levels. The window length M is changed from 40 to 800 sampling points, the *SNR* of the noisy signal is changed from -1.55 to -15.5-dB by increasing the amplitudes of the two types of noises in equal proportions, and the attenuation control factor q of the soft masking in equation (7) is set as 1.

Figure 9 provides a compact picture about the effect of the window length M in STFT on the de-noising performance of the proposed algorithm through the defined two evaluation metrics, *SNR* and *RMSE*. It can be observed that the de-noising performance improves as the window length increases in the initial part pf the curves, and then it becomes nearly flat once the number of sampling points exceed the critical value of 200. This critical value is equivalent to the frequency resolution of the spectrogram of 0.625MHz, which is just sufficient to distinguish the last two PD pulses with dominant frequencies of 1MHz and 5MHz, respectively, from the discrete spectrum noise with dominant frequencies of 3MHz and 7MHz. Once the window length is smaller than the critical value, the frequency resolution of the spectrogram is insufficient to distinguish the PD pulses from the discrete spectrum noise, resulting in a significant drop in the de-noising performance as shown in the front parts of the curves in Fig. 9. Therefore, the window length should not be too small.

Figure 10 shows the effect of the window length M on the computational time of the proposed algorithm. It can be observed that the efficiency decreases as the window length increases, suggesting a selection of the smallest value of M may lead to sufficiently good performance indexes (such as those considered in Fig. 9).



(a) (b)

Fig. 9. Effect of the window length M on (a) *SNR* and (b) *RMSE* in varying noise levels.

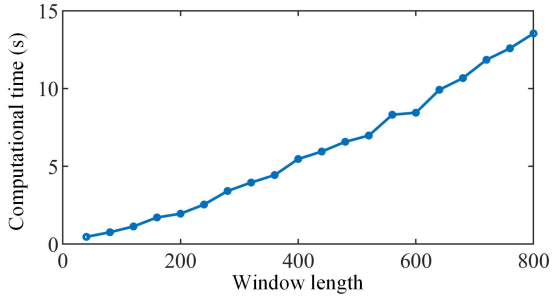


Fig. 10. Effect of the window length M on the overall computational time required by the proposed algorithm.

Referring to the frequency pass-band of the used PD sensor, an empirical formula allowing to determine the window length M is:

$$M \geq \frac{f_s}{\Delta f} = \frac{\alpha \cdot f_s}{f_{\text{high}} - f_{\text{low}}} \quad (18)$$

where f_{high} and f_{low} are the upper and lower cutoff frequencies of the used PD sensor, and α is the ratio between the passband frequency of the sensor and the acceptable frequency resolution. In practical applications, it is recommended to set α a bit larger than some dozens (e.g., 20 to 60) to ensure sufficient frequency resolution of the spectrogram for good de-noising performance and simultaneously less computation time.

C. Effect of the attenuation control factor in the soft masking

The effect of the attenuation control factor of the soft masking on the two evaluation metrics *SNR* and *RMSE* is shown in Fig. 11, where the attenuation control factor, q, is changed from 0 to 6, the SNR of the noisy signal is changed from -1.55 to -15.5-dB, and α in (18) is set as 48. The above value corresponds to the window length $M = 48 \cdot 125\ \text{MHz} / (30\ \text{MHz} - 0.5\ \text{MHz}) \approx 200$ sampling points if a PD sensor with the pass-band from 0.5 MHz to 30 MHz is used, which is sufficient to detect the synthetic PD signal. It can be observed that with the increase of the attenuation control factor, the de-noising performance is obviously improved at first, indicating that soft masking works

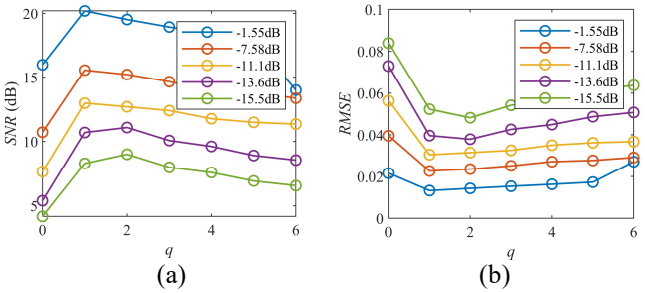


Fig. 11. Effect of the attenuation control factor q on (a) *SNR* and (b) *RMSE* in varying noise levels. (q equal to 0 denotes the soft masking is disenabled)

first effectively, and then gradually deteriorates. The initial increase in the de-noising performance is because more white noise is removed; the subsequent drop in de-noising performance is because part of the useful PD signal is also removed, resulting in some signal distortions. In addition, comparing the optimal q values for different noise levels, it can be found that a relatively small q value produces the best noise reduction effect at low noise levels, while a relatively large q value produces the best noise reduction effect at high noise levels. Referring to the simulation results, the attenuation control factor, q, is recommended to be set between 0.5 to 2.5 and it should be relatively large for high noise levels and relatively small for low noise levels.

D. De-noising

Figure 12 collects the results of the de-noising of the simulated PD signal with high-level noise (*SNR*=-11.10dB) via the proposed STFT-SVD, H-ASVD, WT, and EMD algorithms. It can be seen from Fig. 12(a) that the H-ASVD algorithm can hardly remove the discrete spectrum noise. In Fig. 12(b), the WT algorithm has good behavior for the first and second pulses, but it basically fails to reconstruct the third pulse. The difference in the noise reduction effect on the PD pulses is attributed to the selected mother wavelet (dB8); the failure of the last PD pulse de-noising is due to the insufficient frequency resolution requiring to separate the PD pulses (with a dominant frequency of 5MHz) with the discrete spectral noises with sinusoidal components at 3MHz and 7MHz. In Fig. 12(c), the EMD technique can only discriminate the PD signal vaguely, and a large amount of white noise and waveform distortions still remain. In Fig.12(d), compared with the results of H-ASVD, WT, and EMD algorithms, the proposed STFT- SVD algorithm can be effectively used to reduce both the white noise and the discrete spectrum noise in all three types of PD pulses, although a small amount of noise remains. In Fig. 12(e), where the soft masking is enabled, the de-noising performance is improved, thus proving the benefit of this additional step in the de-noising procedure.

The evaluation metrics of all the algorithms are listed in Tab. II, where it can be observed that the proposed STFT-SVD algorithm with the soft masking has the largest signal-to-noise ratio (the highest *SNR*), and the least waveform distortion (the lowest *RMSE* and the highest *NCC*).

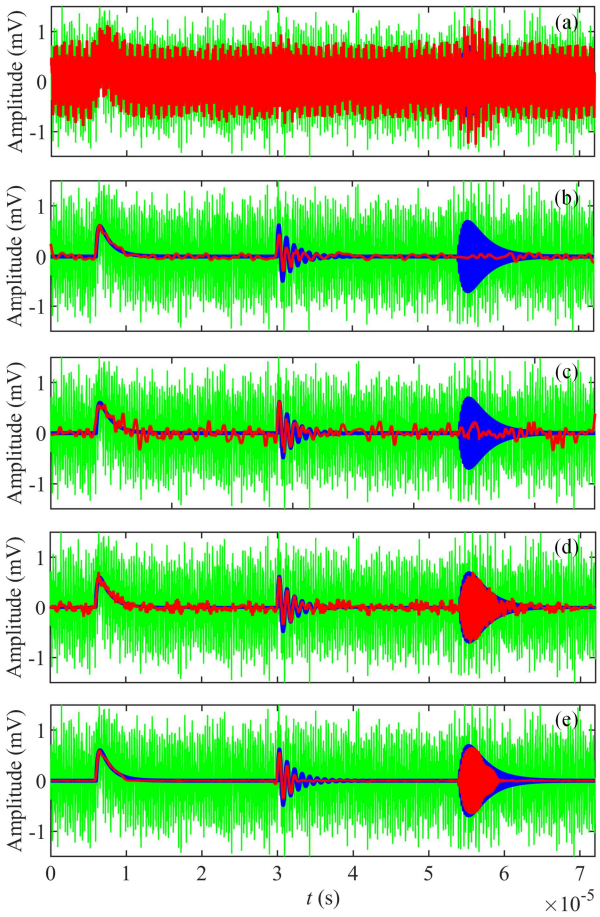


Fig. 12. De-noising results of (a) H-ASVD (the size of Hankel matrix: 450×9000), (b) WT (the mother wavelet: dB8), (c) EMD (the ensemble number: 300), (d) the proposed STFT-SVD without the soft masking (M=200, q=0), and (e) the proposed STFT-SVD with the soft masking (M=200, q=2). The green, blue, and red lines denote the noisy signal (*SNR*=-11.10 dB), the original signal, and the de-noised signal, respectively.

TABLE II
EVALUATION METRICS COMPARISON OF THE ALGORITHMS

Method	*SNR* (dB)	*NCC*	*RMSE*
STFT-SVD (L=200, q=0)	7.53	0.912	0.057
STFT-SVD (L=200, q=2)	12.97	0.975	0.030
WT	2.10	0.627	0.107
EMD	0.236	0.518	0.132
H-ASVD	-9.17	0.273	0.390

V. RESULTS: MEASURED DATA

This Section collects the results of the application of the proposed de-noising tool to the measurements carried out by two different ultra-wide band PD sensors widely used for on-

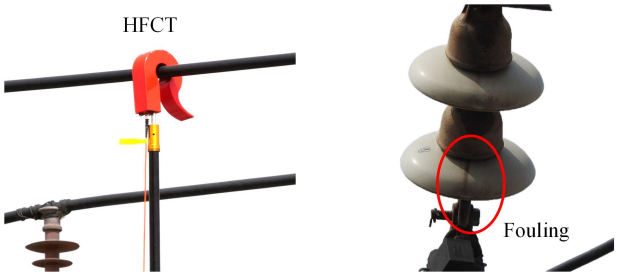


Fig. 13. Case I: layout of the high frequency current transformer on a 10-kV CC line (a) and picture of the PD defect on an insulator caused by fouling (b).

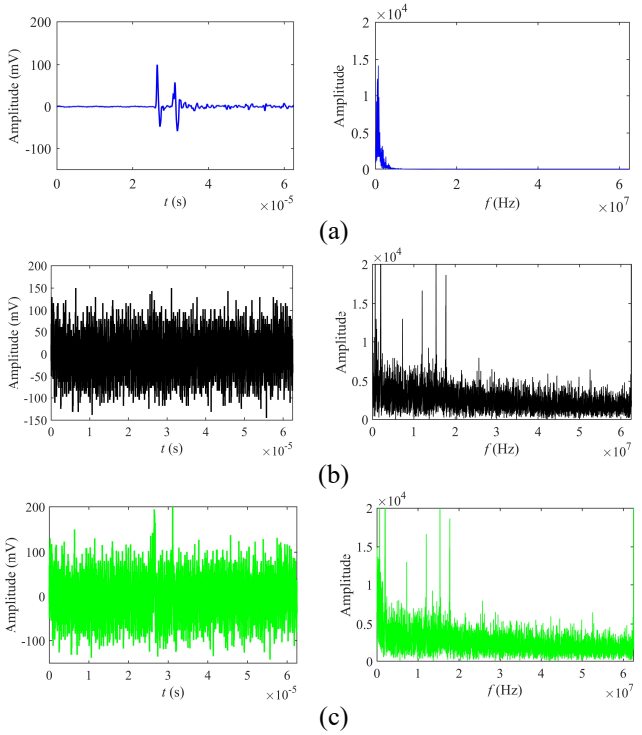


Fig. 14. Case I: time-domain waveforms and frequency-domain spectrums of the HFCT-measured PD signal (a), the HFCT-measured noise (b), and their combination (c).

site PD measurement in cables, transformers, motors, switchgears, overhead lines, etc. Two test cases are considered: the first one involves a high frequency current transformer (HFCT) and the second one an ultra-high frequency (UHF) antenna. The achieved de-noising performance is compared with the ones obtained through the alternative WT [7], EMD [17], and H-ASVD algorithms [23].

A. Case I: De-noising of a PD signal measured by HFCT

In this first real application test case, the proposed STFT-SVD algorithm is applied to a PD signal measured by means of a HFCT to demonstrate its de-noising performance. To have a controlled environment and test, two measurements are carried out: one is related to a PD signal with the smallest possible noise corruption and the other involves the contribution of the noise only, being this latter associated to another conductor and time

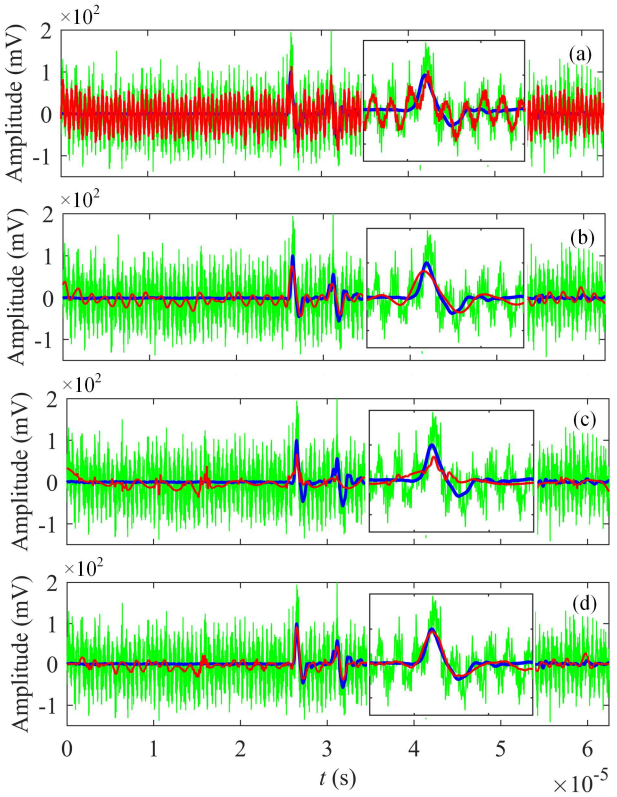


Fig. 15. Case I: de-noising results of real noisy PD signal by (a) H-ASVD (the size of Hankel matrix: 390×7810), (b) WT (the mother wavelet: dB8), (c) EMD (the ensemble number: 300), and (d) the proposed STFT-SVD (M=200, q=2). The green, blue, and red lines denote the noisy, original, and de-noised signal, respectively. The insets zoom in the PD pulses.

window where only the spurious disturbance is recorded. The PD signal originates from a soiled insulator in a 10-kV covered conductor (CC) line, as shown in Fig. 13. The passband frequency of the HFCT is in the range of 0.5-30 MHz and the sampling frequency for data collection is 125 MHz.

Figure 14 collects both the time waveforms and the corresponding frequency spectra of the PD measured signal and noise. Specifically, panel (a) corresponds to the PD signal measured as closer as possible to the insulator, to reduce the detrimental effects of noise and to produce a reference response which is eventually compared with the reconstructed PD. Panel (b) and panel (c) corresponds to the noisy contrition and to the PD noisy response, respectively. The PD noisy response is generated by summing the noiseless and the noisy waveforms. It is important to point out that the noisy waveform of Fig. 14(b) includes both the effects of the white noise and of the discrete spectrum noise, as discussed in the synthetic test example considered in Section II. Also for this case, the PD pulse, for both the time- and the frequency-domain waveforms, is massively cluttered by the effect of the large superimposed noise, thus making the de-noising a challenge.

De-noising is carried out starting from the time-domain measured response of Fig. 14(c) by using all the considered algorithms. In the proposed STFT-SVD algorithm, the tuning parameters are defined as follows. According to (18) and the

suggested range of α between 20 to 60, M should be chosen in the range between 83 and 254. It is important to point out that the discussion of the role of M in Sec. IV.B and the additional simulation analyses have proven that the overall accuracy of the method has a very low sensitivity to this parameter. For this test case, the window width value is set to 200 sampling points (corresponding to the ratio $\alpha \approx 48$), which enables the frequency resolution to be 0.625 MHz. A similar reasoning and behavior holds for the other parameter (i.e., the attenuation control factor in the soft masking) which is set to 2, due to the high-level noise.

Fig. 16 offers the results of the cross comparison. We can find that it illustrates the superior performance of the proposed tool which yields a very good reconstructed PD signal (see Fig. 15(d)). From the responses in the figure, we can observe that the H-ASVD algorithm fails to remove the discrete spectrum noise (see Fig. 15(a)), the WT algorithm effectively reduces most of noise but leads to a visible distortion (see Fig. 15(b)), and the EMD algorithm can only discriminate the PD signal vaguely and leads to a significant distortion (see Fig. 15(c)). To sum up, this test has proven the nice features of the proposed de-noising tool for in field measurements via HFCT.

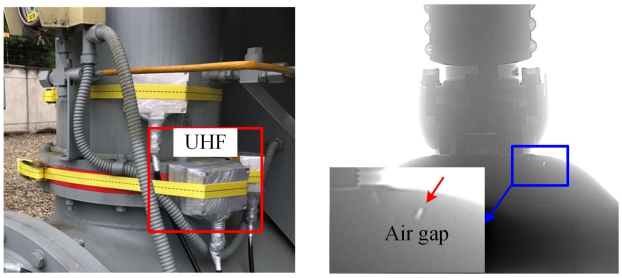


Fig. 16. Case II: layout of a UHF sensor on a 110-kV gas-insulated switchgear (a) and picture of X-ray imaging inspection of the detected internal air gap PD defect of a basin insulator (b).

B. Case II: De-noising of a PD signal measured by UHF

In this second comparison, a real noisy PD signal is measured by means of a UHF in 110 kV gas-insulated switchgear as shown in Fig. 16 (a). The PD was caused by an internal air gap in a basin insulator, as shown in Fig. 16(b). The passband frequency of the UHF antenna is in the range of 300-1500 MHz. The sampling frequency of the data collection device is 10 GHz. As done in Case I, the time-domain waveform and the corresponding frequency spectrum of the noisy signal are shown in Fig. 17. It can be observed also in this case that the PD pulse is almost drowned out by the noise. The edge of the PD pulse is zoomed in Fig. 17(c), where a visible oscillation with a specific frequency appears from about time t_0, which is credibly regarded as the starting point of the PD pulse.

Based on the observations of the tuning parameters used in test case I, the same values of M and q are selected for this second example (hence M=200 and q=2). In this case, the window width value in the proposed STFT-SVD algorithm

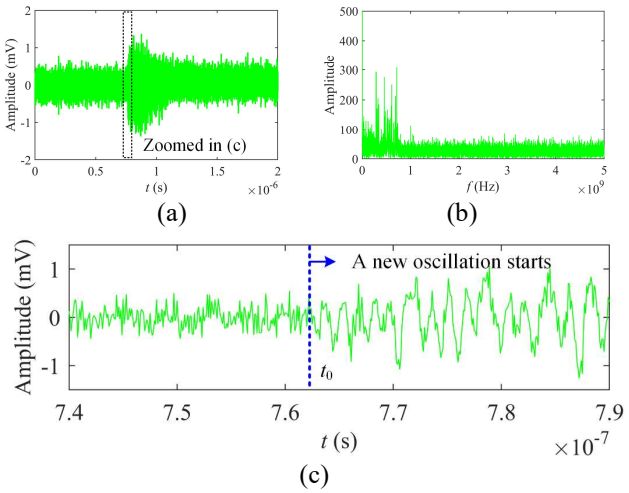


Fig. 17. Case II: time-domain waveform (a) and frequency-domain spectrum (b) of a real noisy PD signal in a 110-kV gas-insulated switchgear, measured by a UHF sensor. The starting of the PD pulse is zoomed in (c).

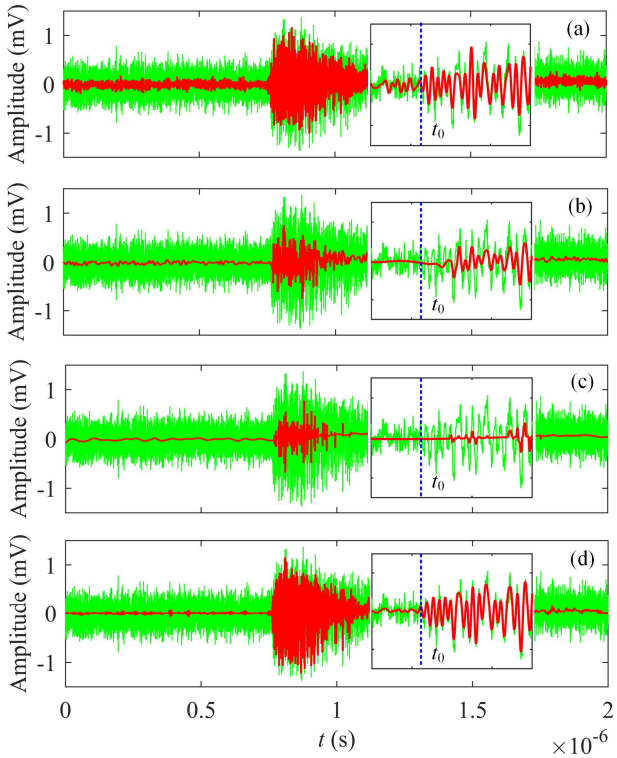


Fig. 18. Case II: de-noising results of a real noisy UHF-measured PD signal by (a) H-ASVD (the size of Hankel matrix: 1000×20000), (b) WT (the mother wavelet: dB8), (c) EMD (the ensemble number: 300), and (d) the proposed STFT-SVD (M=200, q=2). The green line denotes the noisy signal, and red line denotes the de-noised signal. The insets zoom in the starting of the PD pulse.

corresponds to the ratio $\alpha \approx 24$ in (18), which enables a frequency resolution of 50MHz.

The results of this cross comparison are collected in Fig. 18.

It demonstrates the superior performance of the proposed tool, which gives a good performance in detecting the starting of the PD pulse (see Fig. 18(d)). This is essential for PD source localization, which is nowadays one of the common functions required for PD diagnosis on gas-insulated switchgears in order to improve their maintenance efficiency. From the responses in the Figure we can observe that the H-ASVD algorithm is effective in reducing most of the white noise, but some discrete spectrum noise remains, making it quite difficult to identify the start of the PD pulse (see Fig. 18(a)). The WT and EMD algorithms effectively reduce most of the noise, but lead to remarkable distortions, causing the loss of the pulse starting information (see Fig. 18(b) and (c)). In summary, this test has verified the nice features and benefits of the proposed de-noising algorithm in detecting the pulse starting for field PD localizations via UHF.

VI. Conclusion

In this paper, an automatic PD de-noising algorithm is proposed. It is based on a well-defined automatic procedure which involves the joint application of the STFT and the SVD tools. The former time-frequency transform enables the separation of the functional behavior of the PD signal and of the possibly superimposed discrete spectrum noise and the white noise background disturbance. The latter matrix decomposition allows to separate all the different signal contributions in terms of the dominant terms (or modes) and filter out all the noisy terms. The algorithm also embeds some additional features, including a soft-masking mechanism and the optimal selection of the terms leading to the accurate estimation of the noiseless PD signal. The proposed method is demonstrated on a first test case involving a synthetic PD signal which is fully controllable and can be used to validate and stress the tool in extreme conditions (e.g., with a very large impact of noise). A second test case is then considered, where the real measured data are processed and de-noised. A cross comparison with some state-of-the-art alternatives is carried out for all the test cases. To sum up:

1) The proposed algorithm is fully automatic. Minimum description length and Kurtosis criterions are used to automatically select the dominated components to reconstruct the estimated noiseless PD signal;
2) The synthetic test case and the simulation results prove that the proposed de-noising scheme can adaptively reduce, selectively, both of the white noise and the discrete spectrum noise in all the considered (three representative types) PD pulses;
3) The white noise reduction is further improved by applying a soft masking to the time-frequency spectrogram and by filtering out the remaining white noise (together with the discrete spectrum noise) by the selection of PD-dominant contributions.
4) The application of the tool to real measured data demonstrates its effectiveness, as it causes a significantly less waveform distortion than other state-of-the-art alternatives. So this tool has a qualified capacity to detect the important parameters of the signal for advanced PD waveform-based applications, such as PD localization.

Future works will investigate a possible application of the proposed algorithm to an embedded system, which may offer a cheap and effective alternative solution for on-line PD monitoring.

Acknowledgment

This work was supported by the China Scholarship Council. Thanks to Jiayi Wang from the Sichuan Electric Power Research Institute of State Grid for providing real on-site UHF-measured partial discharge data used in this paper.

References

[1] G. C. Montanari, R. Hebner, P. Seri and R. Ghosh, "Self-Assessment of Health Conditions of Electrical Assets and Grid Components: A Contribution to Smart Grids," in *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1206-1214, Mar. 2021.

[2] G. C. Montanari and A. Cavallini, "Partial discharge diagnostics: from apparatus monitoring to smart grid assessment," in *IEEE Electrical Insulation Magazine*, vol. 29, no. 3, pp. 8-17, May-June 2013.

[3] P. Wagenaars, P. A. A. F. Wouters, P. C. J. M. van der Wielen and E. F. Steennis, "Influence of Ring Main Units and Substations on Online Partial-Discharge Detection and Location in Medium-Voltage Cable Networks, " in *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 1064-1071, Apr. 2011.

[4] F. Alvarez, J. Ortego, F. Garnacho and M. A. Sanchez-Uran, "A clustering technique for partial discharge and noise sources identification in power cables by means of waveform parameters, " in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 1, pp. 469-481, Feb. 2016.

[5] G. V. R. Xavier, H. S. Silva, E. G. da Costa, A. J. R. Serres, N. B. Carvalho and A. S. R. Oliveira, "Detection, Classification and Location of Sources of Partial Discharges Using the Radiometric Method: Trends, Challenges and Open Issues, " in *IEEE Access*, vol. 9, pp. 110787-110810, 2021.

[6] S. Sriram, S. Nitin, K. M. M. Prabhu and M. J. Bastiaans, "Signal denoising techniques for partial discharge measurements," in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 12, no. 6, pp. 1182-1191, Dec. 2005.

[7] X. Ma, C. Zhou and I. J. Kemp, "Automated wavelet selection and thresholding for PD detection," in *IEEE Electrical Insulation Magazine*, vol. 18, no. 2, pp. 37-45, March-April 2002.

[8] R. Hussein, K. B. Shaban, and A. H. El-Hag, "Wavelet transform with histogram-based threshold estimation for online partial discharge signal denoising, " I*EEE Trans. Instrum. Meas*., vol. 64, no. 12, pp. 3601–3614, Dec. 2015.

[9] J. Seo, H. Ma and T. Saha, "Probabilistic wavelet transform for partial discharge measurement of transformer," in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 22, no. 2, pp. 1105-1117, Apr. 2015.

[10] M. Ghorat, G. B. Gharehpetian, H. Latifi and M. A. Hejazi, "A New Partial Discharge Signal Denoising Algorithm Based on Adaptive Dual-Tree Complex Wavelet Transform," in *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 10, pp. 2262-2272, Oct. 2018.

[11] J. Tang, S. Zhou, and C. Pan, "A denoising algorithm for partial discharge measurement based on the combination of wavelet threshold and total variation theory," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 3428–3441, Jun. 2020.

[12] S. Zhou, J. Tang, C. Pan, Y. Luo and K. Yan, "Partial Discharge Signal Denoising Based on Wavelet Pair and Block Thresholding," in *IEEE Access*, vol. 8, pp. 119688-119696, 2020.

[13] C. F. F. C. Cunha, A. T. Carvalho, M. R. Petraglia, and A. C. S. Lima, "A new wavelet selection method for partial discharge denoising," Electr. Power Syst. Res., vol. 125, pp. 184–195, Aug. 2015.

[14] Ö. Altay, Ö. Kalenderli. "Wavelet base selection for de-noising and extraction of partial discharge pulses in noisy environment, " in *IET Science, Measurement & Technology*, vol. 23, no.9(3), pp. 276-284, Apr. 2015.

[15] M.-Y. Lin, C.-C. Tai, Y.-W. Tang, and C.-C. Su, "Partial discharge signal extracting using the empirical mode decomposition with wavelet transform," in *Proc. 7th Asia–Pacific Int. Conf. Lightning*, pp. 420–424, Nov. 2011.

[16] J. Xie, F. Lv, M. Li, and Y. Wang, "Suppressing the discrete spectral interference of the partial discharge signal based on bivariate empirical mode decomposition," in *Int. Trans. Electr. Energy Syst.*, vol. 27, no. 10, p. e2407, Oct. 2017.

[17] S. Zhang *et al*., "An Adaptive CEEMDAN Thresholding Denoising Method Optimized by Nonlocal Means Algorithm," in *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6891-6903, Sept. 2020.

[18] Yang, Yu, Junsheng Cheng, and Kang Zhang, "An ensemble local means decomposition method and its application to local rub-impact fault diagnosis of the rotor systems," Measurement, 45, no. 3 (2012): 561-570.

[19] Dragomiretskiy, Konstantin, and Dominique Zosso, "Variational mode decomposition," *IEEE transactions on signal processing*, 62, no. 3 (2013): 531-544.

[20] Ni, Qing, J. C. Ji, Ke Feng, and Benjamin Halkon, "A fault information-guided variational mode decomposition (FIVMD) method for rolling element bearings diagnosis," *Mechanical Systems and Signal Processing*, 164 (2022): 108216.

[21] Zhang, Jun, Junjia He, Jiachuan Long, Min Yao, and Wei Zhou, "A new denoising method for UHF PD signals using adaptive VMD and SSA-based shrinkage method," *Sensors*, 19, no. 7 (2019): 1594.

[22] T. K. Abdel-Galil, A. H. El-Hag, A. M. Gaouda, M. M. A. Salama and R. Bartnikas, "De-noising of partial discharge signal using eigen-decomposition technique," in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 15, no. 6, pp. 1657-1662, Dec. 2008.

[23] M. B. Ashtiani and S. M. Shahrtash, "Partial discharge de-noising employing adaptive singular value decomposition," in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 21, no. 2, pp. 775-782, Apr. 2014.

[24] K. Zhou, M. Li, Y. Li, M. Xie, Y. Huang, "An improved denoising method for partial discharge signals contaminated by white noise based on adaptive short-time singular value decomposition," in *Energies*. Vol. 12, no. 18, pp. 3465-3481, Jan. 2019.

[25] S. Govindarajan, J. Subbaiah, A. Cavallini, K. Krithivasan and J. Jayakumar, "Partial Discharge Random Noise Removal Using Hankel Matrix-Based Fast Singular Value Decomposition," in *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 7, pp. 4093-4102, Jul. 2020.

[26] L. Li, X. Wei, "Suppression Method of Partial Discharge Interferences Based on Singular Value Decomposition and Improved Empirical Mode Decomposition,"in *Energies*. vol. 14, no. 24, pp. 8579-8601, Jan. 2021.

[27] J. Zhong, X. Bi, Q. Shu, M. Chen, D. Zhou and D. Zhang, "Partial Discharge Signal Denoising Based on Singular Value Decomposition and Empirical Wavelet Transform," in *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 11, pp. 8866-8873, Nov. 2020.

[28] Karami, Hossein, and Gevork B. Gharehpetian, "Limitations of Partial Discharge De-noising of Power Transformer Using Adaptive Singular Value Decomposition," in *2019 International Power System Conference* (PSC), pp. 777-781. IEEE, 2019.

[29] Y. Liu, W. Zhou, P. Li, S. Yang, Y. Tian, "An ultrahigh frequency partial discharge signal de-noising method based on a generalized S-transform and module time-frequency matrix," in *Sensors*. Vol. 16, no. 6, pp. 941-960, Jun. 2016.

[30] High-Voltage Test Techniques—Partial Discharge Measurements, IEC standard 60270, 2015.

[31] H. Okubo, N. Hayakawa and A. Matsushita, "The relationship between partial discharge current pulse waveforms and physical mechanisms," in *IEEE Electrical Insulation Magazine*, vol. 18, no. 3, pp. 38-45, May-June 2002.

[32] J. Long, X. Wang, W. Zhou, J. Zhang, D. Dai and G. Zhu, "A Comprehensive Review of Signal Processing and Machine Learning Technologies for UHF PD Detection and Diagnosis (I): Preprocessing and Localization Approaches," in *IEEE Access*, vol. 9, pp. 69876-69904, 2021.

[33] H. E. Rojas, M. C. Forero and C. A. Cortes, "Application of the local polynomial Fourier transform in the evaluation of electrical signals generated by partial discharges in distribution transformers," in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 1, pp. 227-236, Feb. 2017.

[34] Sawada, Hiroshi, Shoko Araki, Ryo Mukai, and Shoji Makino, "Blind extraction of dominant target sources using ICA and time-frequency masking," IEEE Transactions on Audio, Speech, and Language Processing, 14, no. 6, pp. 2165-2173, NOV, 2006.

[35] Sharif, Waqas, Michael Muma, and Abdelhak M. Zoubir, "Robustness analysis of spatial time-frequency distributions based on the influence function," *IEEE transactions on signal processing*, 2013, 61, no. 8, pp. 1958-1971.

[36] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387-392, April 1985.

[37] C. J. Zarowski, "The MDL criterion for rank determination via effective singular values," in *IEEE Transactions on Signal Processing*, vol. 46, no. 6, pp. 1741-1744, Jun. 1998.

[38] DeCarlo, Lawrence T. "On the meaning and use of kurtosis." *Psychological methods* 2, no. 3, pp. 292-307, 1997.