applied sciences

MDPI

Article

Partial Discharge Signal Denoising Algorithm Based on Aquila Optimizer–Variational Mode Decomposition and K-Singular Value Decomposition

Jun Zhong, Zhenyu Liu and Xiaowen Bi *

College of Electrical Engineering, Sichuan University, Chengdu 610225, China; zhongjun55@163.com (J.Z.); zhenyul748@gmail.com (Z.L.)
* Correspondence: xiaowen_b913@163.com

Abstract: Partial discharge (PD) is a primary factor leading to the deterioration of insulation in electrical equipment. However, it is hard for traditional methods to precisely extract PD signals in increasingly complex engineering environments. This paper proposes a new PD signal denoising method combining Aquila Optimizer–Variational Mode Decomposition (AO-VMD) and K-Singular Value Decomposition (K-SVD) algorithms. Firstly, the AO algorithm optimizes critical parameters of the VMD algorithm. For the PD signal overwhelmed by noise, the AO-VMD algorithm can decompose it and reconstruct it by using kurtosis. In this process, the majority of the noise is removed, and the characteristics of the original signal are shown. Subsequently, the K-SVD algorithm performs sparse decomposition on the signal after OA-VMD, constructs a learned dictionary, and captures the characteristics of the signal for continuous learning and updating. After the dictionary learning is completed, the best matching atoms from the dictionary are selected to precisely reconstruct the original noiseless signal. Finally, the proposed method is compared with three traditional algorithms, Adaptive Ensemble Empirical Mode Decomposition (AEEMD), SVD-VMD, and the Adaptive Wavelet Multilevel Soft Threshold algorithm, on the simulated signal and the actual engineering signal. The results both demonstrate that the algorithm proposed by this paper has superior noise reduction and signal extraction performance.

Keywords: partial discharge; AO; VMD; sparse decomposition; K-SVD; kurtosis

Citation: Zhong, J.; Liu, Z.; Bi, X. Partial Discharge Signal Denoising Algorithm Based on Aquila Optimizer–Variational Mode Decomposition and K-Singular Value Decomposition. *Appl. Sci.* **2024**, *14*, 2755. https://doi.org/10.3390/app14072755

Academic Editor: Filipe Soares

Received: 12 February 2024
Revised: 15 March 2024
Accepted: 21 March 2024
Published: 25 March 2024

1. Introduction

PD signal is a crucial component of monitoring electrical equipment and can accurately reflect the severity of insulation deterioration [1]. Engineers can evaluate the safety and stability of high-voltage equipment using the PD signal. However, due to the complexity of the engineering work environment, PD signals will inevitably be contaminated with various types of noise. White Gaussian noise and narrowband noise have the most significant impact on PD signals [2]. How to eliminate noise, determine the edge, and precisely extract the waveform of PD signals are currently essential research issues.

In PD signal extraction, PD signals are often polluted by various forms of noise interferences, including continuous periodic noise like system harmonics, impulse interferences from power electronics, and random disturbances such as corona discharges. Additionally, white noise, encompassing thermal and ground network disturbances, further complicates signal clarity. Effective noise mitigation is essential for precise PD signal analysis [3].

Currently, some traditional signal denoising techniques are employed for noise reduction in PD signals, including Wavelet Transform (WT) [4], Empirical Modal Decomposition (EMD) [5], and SVD [6]. WT is recognized for its proficient time-frequency analysis capabilities, but selecting an optimal wavelet basis for signal decomposition remains challenging, and the choice of wavelet basis and decomposition level significantly affects the denoising outcome [7]. In contrast to WT, EMD generates basic functions directly from the signal,

offering adaptive decomposition into multiple modal components with notable robustness. However, EMD can encounter modal aliasing issues, which can hinder effective noise reduction [8]. SVD is also always applied for noise suppression, but it requires manual determination of effective singular value orders, making the process susceptible to human error [9]. In recent research on denoising PD signals, researchers have increasingly turned to advanced and hybrid versions of traditional algorithms such as the AEEMD algorithm [10], the integration of SVD with the VMD algorithm [11], and the Adaptive Wavelet Multilevel Soft Threshold algorithm [12].

VMD improves upon the traditional EMD method, offering enhanced noise filtering, especially for white noise in vibration signals. The choice of VMD parameters critically affects the quality of the decomposition outcome [13], and a mix of other methods is needed to suppress periodic interference and white noise effectively [14]. Sparse Dictionary Learning, a machine learning algorithm that combines the principles of sparse representation and dictionary learning, has recently gained traction in signal processing. This sparse dictionary learning algorithm is divided into two distinct types: the Analysis Dictionary algorithm and Learning Dictionary algorithm. Although the structure of analysis dictionary is easy to access, the construction of prior knowledge depends on the target signal, which has limitations on the unknown fault features [15]. Denoising algorithms based on analysis dictionaries have been applied in PD signal denoising [16,17]. However, the learning dictionary method exemplified by the K-SVD algorithm has been unexplored in signal processing.

2. Basic Theory

2.1. VMD Decomposition Principle

2.1.1. Basic Principle

The VMD algorithm is improved by Dragomiretskiy and Zosso in 2014 [18]. The VMD algorithm evolves from the EMD algorithm, with both fundamentally decomposing signals into various modes [19]. Specifically, VMD seeks to decompose a real-valued input signal into a predetermined number of modes, denoted as K, with each mode identified as μ_k. These modes are designed with finite bandwidths, ensuring that each extracted mode occupies a specific frequency range without overlapping with others. The primary goal of VMD is to ensure that the sum of these K modes, μ_k, precisely reconstructs the original input signal. The process involves a constrained optimization problem that aims to minimize the bandwidth of each mode while adhering to this reconstruction constraint. This optimization challenge is addressed through a variational approach, seeking the optimal set of modes that meet these criteria. This constrained optimization variational problem is expressed as follows:

$$\min_{\{\mu_k\}\{\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot \mu_k \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad \text{s.t.} \sum_{k=1}^{K} \mu_k(t) = x(t) \tag{1}$$

where $x(t)$ is the original input signal, $\mu_k(t)$ and ω_k are the Band-Limited Intrinsic Mode Functions (BLIMFs) and their central frequency, $\delta(t)$ is the Fermi–Dirac distribution function, and $(\delta(t) + j/\pi t)\mu_k$ is the analytic signal obtained by convolution of each mode $\mu_k(t)$ through Hilbert transform. The exponential term $e^{(-j\omega_k t)}$ indicates that the analytic signal is modulated to the baseband, and the BLIMF bandwidth can be estimated by calculating the L^2 norm.

Then, by introducing the augmented Lagrangian L and quadratic penalty factor α, this constraint problem can be changed into an unconstraint problem as follows:

$$L(\{\mu_k\},\{\omega_k\},\lambda) = \alpha \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi \mathrm{t}} \right) \cdot \mu_k \right] e^{-j\omega_k t} \right\|_2^2 \left\| x(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + < \lambda(t), x(t) - \sum_{k=1}^{K} u_k(t) > \tag{2}$$

Lastly, it is necessary to find the saddle point of the augmented Lagrangian L in a sequence of iterative sub-optimizations, denoted as the Alternate Direction Method of Multipliers (ADMM). In this process, it must initialize μ_k, ω_k, and λ continuously as follows:

$$\hat{\mu}_k^{n+1}(\omega) = \frac{\hat{x}(\omega) - \sum_{i<k} \hat{\mu}_i^{n+1}(\omega) - \sum_{i>k} \hat{\mu}_i^{n}(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{3}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \left| \hat{\mu}_i^{n+1}(\omega) \right|^2 \mathrm{d}\omega}{\int_0^\infty \left| \hat{\mu}_i^{n+1}(\omega) \right|^2 \mathrm{d}\omega} \tag{4}$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^{n}(\omega) + \tau \left(\hat{x}(\omega) - \sum_{k=1}^{K} \hat{\mu}^{n+1}(\omega) \right) \tag{5}$$

In Formulas (3)–(5), $\wedge$ is a Fourier transformation, and τ is the time step of dual rise method.

It can set a discrimination accuracy ε when it satisfies the following formula:

$$\frac{\sum_{k=1}^{K} \left\| \hat{\mu}_k^{n+1} - \hat{\mu}_k^{n} \right\|_2^2}{\left\| \hat{\mu}_k^{n} \right\|_2^2} < \varepsilon \tag{6}$$

The iteration stops, obtaining K modal functions satisfied by Formula (1) by inverse Fourier transform.

2.1.2. Optimizing Parameters—AO-VMD

In 2021, Abualigah et al. introduced the Aquila Optimizer algorithm, a new intelligent meta-heuristic optimization algorithm [20]. The AO algorithm is a novel meta-heuristic optimization algorithm inspired by the hunting behaviors of the Aquila (eagle) in nature. This optimization method is designed to effectively navigate the search space for optimal solutions, mimicking Aquila's strategic phases in catching prey [21].

In the optimization of VMD parameters, the AO is employed to navigate the parameter space, with the fitness function defined as the ratio of kurtosis to envelope entropy. The AO algorithm is initialized with a predefined population size, dimensionality for the optimization, and the maximum number of iterations. This approach ensures that the VMD is configured to yield the most accurate signal representation by leveraging the exploratory and exploitative capabilities of AO. The flowchart of the AO-VMD algorithm is shown in Figure 1.

In this study, the AO primarily optimizes two critical parameters of VMD: the number of decomposition levels K and the penalty factor α.

2.1.3. Selecting IMFs and Reconstructing the Signal-Kurtosis

After signal decomposition by the VMD, choosing the appropriate Intrinsic Mode Functions (IMFs) for signal reconstruction is a crucial research focus. Several criteria, such as correlation [22] and energy content [23], can guide the selection of IMFs. Kurtosis is particularly apt for the selection of IMFs due to its sensitivity to the atypical, transient

burst characteristics of PD signals. Researchers already proved the utility of kurtosis in the extraction and identification of significant PD signal denoising [24].

The kurtosis is calculated by dividing the random variable's fourth central moment by the standard deviation's fourth power [25]. A dimensionless factor assesses how far the signal deviates from the normal distribution. The formula for kurtosis is as follows:

$$K = \frac{\int_{-\infty}^{+\infty}[x(t) - \bar{x}]^4 p(x)dx}{\sigma^4} \tag{7}$$

where $x(t)$ is the instantaneous amplitude, $\bar{x}$ is the average of the amplitude, $p(x)$is the probability density, and σ is the standard deviation.

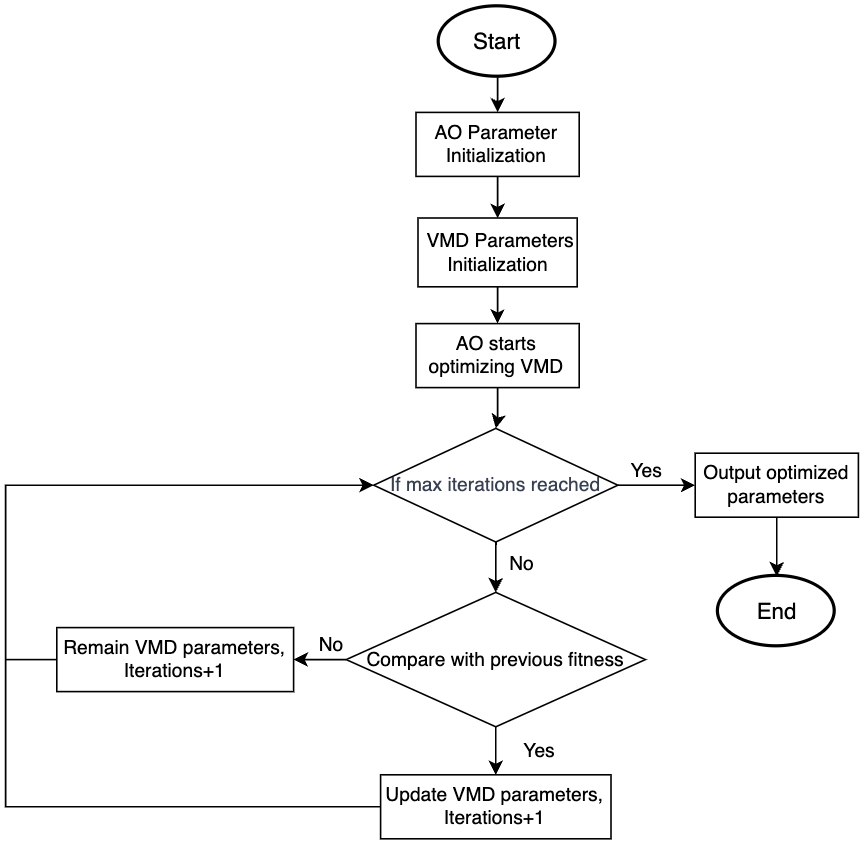


Figure 1. Flowchart of the AO-VMD.

The discretization formula is as follows:

$$K = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i - \bar{x}_i}{\sigma_t}\right)^4 \tag{8}$$

where x_i is the value of the signal, $\bar{x}_i$ is the average of the siganl value, N is the length of the signal, and σ_t is the standard deviation.

2.2. Sparse Dictionary Learning

The principle of Sparse Dictionary Learning rests on the assumption that signals can be compactly represented by a sparse linear combination of basis functions, known as atoms, which are collected in a dictionary. Sparse Dictionary Learning involves two main components: the construction of the dictionary and the sparse decomposition of the signal over this dictionary.

2.2.1. Sparse Decomposition

Sparse representation of a signal involves expressing it as a sum of a minimal number of basic elements, or atoms, from a dictionary [26]. Mathematically, this is formulated as:

$$x = D\alpha + r \tag{9}$$

where $x \in R^n$ is the signal to be represented, $D \in R^{n*m}$ is the dictionary composed of m atoms, $\alpha \in R^m$ is the sparse coefficient vector with a few non-zero entries, and r is the residual of the representation.

The objective is to identify the most sparse coefficient vector α that minimizes the residual r, thereby attaining a succinct representation of the signal x. Orthogonal Matching Pursuit (OMP) is commonly utilized for this purpose, as it incrementally selects atoms from the dictionary D that exhibit the highest correlation with the signal, thereby progressively building the sparsest possible representation. OMP operates on a greedy algorithmic principle, whereby it iteratively chooses the dictionary atoms that have the greatest correlation with the current residual in the approximation of signal x.

2.2.2. Dictionary Learning

Dictionary Learning is a critical process within the framework of Sparse Representation, where the goal is to learn an optimal set of atoms that constitute the dictionary D. This learning process is designed to tailor the dictionary to the specific characteristics of the signal or data set at hand, thus enabling a more efficient and representative sparse representation [27].

In mathematical terms, Dictionary Learning seeks to solve the optimization problem:

$$\underset{D,\,\alpha}{\text{minimize}} \|\|X - D\alpha\|\|_F^2 \quad \text{subject to} \quad \|\alpha_i\|_0 \leq T \quad \forall i \tag{10}$$

where X represents the data matrix, D is the dictionary to be learned, α is the sparse coefficient matrix, $||\cdot||_F$ denotes the Frobenius norm, $||\alpha_i||_0$ represents the l_0 norm of the i-th column of α, indicating the number of non-zero entries, and T is the sparsity threshold.

Sparse dictionaries can be categorized into Analysis Dictionary and Learning Dictionaries. The Analysis Dictionary is constructed using predefined mathematical functions, while the Learning Dictionary is designed to learn the most representative atoms directly from the data. This article used the K-SVD Learning Dictionary algorithm to carry out PD signal denoising.

The K-SVD algorithm was introduced by Aharon, Elad, and Bruckstein in 2006, and it is generalized by the K-Means algorithm [28]. In the initial phase of sparse coding, the dictionary D can proceed under the assumption that it is static. This stage of the optimization process focuses on seeking sparse representations, with the associated coefficients aggregated in the matrix α. The term encapsulating the sparsity constraint can thus be reformulated as follows:

$$min_{D,\,\alpha}||X - D\alpha||_F^2 = \sum_{i=1}^{N} ||x_i - D\alpha_i||_2^2 \tag{11}$$

Therefore, Formula (11) can be decoupled to N distinct problems as follows:

$$\underset{\alpha_i}{\text{minimize}} \left\{ \|\|x_i - D\alpha_i\|\|_2^2 \right\} \quad \text{subject to} \quad \|\alpha_i\|_0 \leq T_0, \quad \text{for } i = 1, 2, ..., N \tag{12}$$

Here, this problem can be addressed by the pursuit algorithms such as OMP.

Secondly, both D and α can be assumed fixed, and the process of updating the dictionary is with the nonzero. If only one column in the dictionary d_k is put in the question and

the coefficients that correspond to it, the k-th row in α, denoted as x_T^k, the penalty term can be rewritten as

$$\begin{aligned} \|\mathbf{X} - \mathbf{D}\alpha\|_F^2 &= \|\mathbf{X} - \sum_{j=1}^{k} d_j x_j^T\|_F^2 \\ &= \|(\mathbf{X} - \sum_{j \neq k} d_j x_j^T) - d_k x_k^T\|_F^2 \\ &= \|\mathbf{E}_k - d_k x_k^T\|_F^2 \end{aligned} \tag{13}$$

Here, the SVD algorithm can be employed to find alternatives for α and d_k; SVD is adept at finding the closest rank-1 matrix to E_k in the Frobenius norm, effectively minimizing the error as defined in Equation (13). By applying SVD decomposition $E_k^R = U\Delta V^T$, it can select the updated dictionary column d_k to be the first column of U. The coefficient vector x_T^k can be updated to be the first column of V multiplied by the leading singular value $\Delta(1,1)$. This method maintains the structure of the dictionary while optimizing the approximation of the error matrix E_k.

2.2.3. Denoising Algorithm Based on AO-VMD and K-SVD Flowchart

As shown in Figure 2, the integral process of the combined PD signal denoising method proposed in this paper is as follows:

Step 1: Input the original PD signal.

Step 2: Introduce noise into the original PD signal to simulate a realistic noisy environment.

Step 3: Use the AO method to find optimal VMD settings. Apply VMD to break down the noisy signal into modes and then reconstruct it.

Step 4: Start with an initial learning dictionary. Refine this dictionary using the K-SVD algorithm for a better representation of the signal.

Step 5: Use the OMP algorithm with the refined dictionary for further signal decomposition and reconstruction after VMD processing.

Step 6: Output the Reconstructed PD Signal.

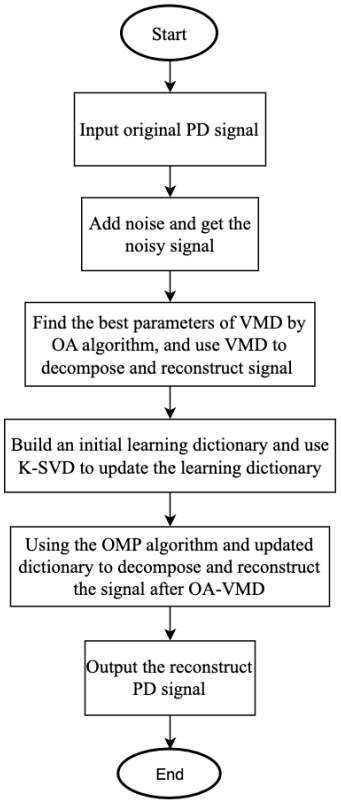


Figure 2. Flowchart of PD noise elimination.

3. Simulation Analysis of PD Signal Denoising

3.1. Simulation Signal Generation

According to research, the original noise-free PD signal can be simulated by four distinct models with parameters in Tables 1 and 2 [29]:

Table 1. Decay models and their equations.

Model	Equation
Single exponential decay model:	$f(t) = Ae^{-\frac{t}{\tau}}$
Double exponential decay model:	$f(t) = A\left(e^{-1.3\frac{t}{\tau}} - e^{-2.2\frac{t}{\tau}}\right)$
Single exponential oscillation decay model:	$f(t) = Ae^{-\frac{t}{\tau}}\sin(2\pi f_c t)$
Double exponential oscillation decay model:	$f(t) = A\left(e^{-1.3\frac{t}{\tau}} - e^{-2.2\frac{t}{\tau}}\right)\sin(2\pi f_c t)$

Table 2. Parameters of PD signal.

Sequences of Signal	A (mV)	τ (μs)	f_c (MHz)	Sample Frequency (MHz)
I	0.8	0.02	50	1
II	0.3	0.03	50	1

where τ is the attenuation constant and f_c is the oscillation constant.

The most prevalent types of noise encountered in on-site PD signals are narrowband noise and white noise [30]. This simulation includes the −20 dB Gaussian white noise and narrowband noise to bring the PD simulation signal closer to the actual signal. Narrowband noise in the engineering environment consists primarily of carrier communication interference and broadcast communication (medium band 0.5–1.6 MHz, short band 2.3–25 MHz, FM band 88–108 MHz). Combining sinusoidal signals with varying amplitudes allows us to simulate narrowband noise. In this paper, the frequency of simulating narrowband interference is set to 9 MHz, 14 MHz, and 96 MHz. Consequently, the original noiseless and polluted PD signal can be successfully simulated in Figure 3:

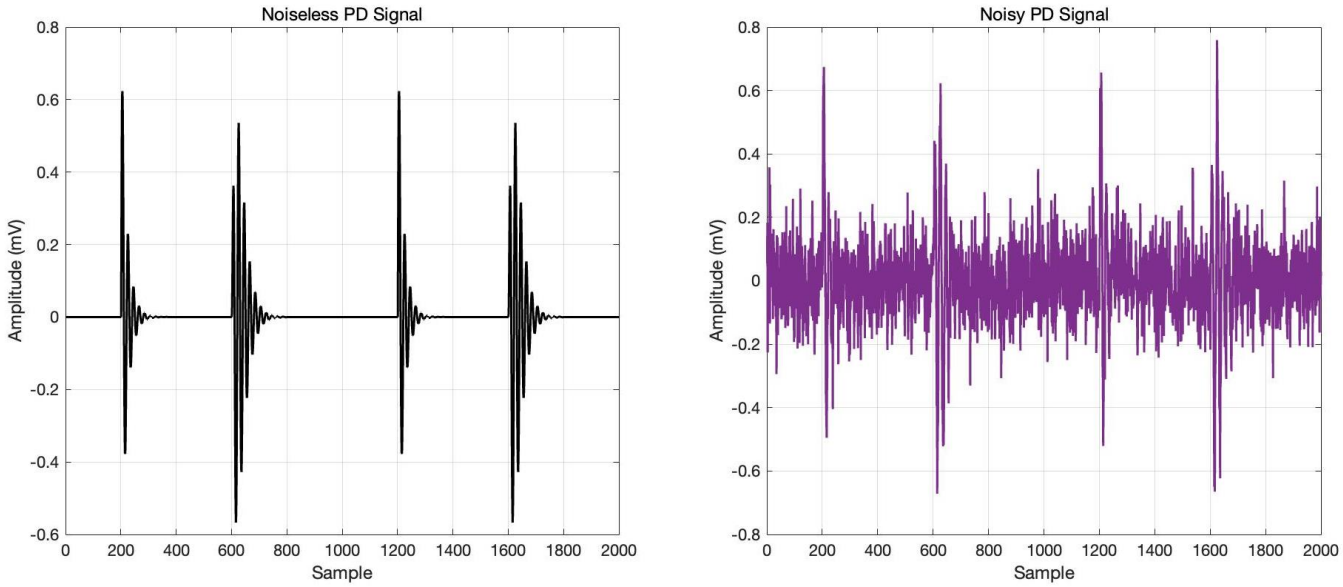


Figure 3. Simulated noiseless and polluted PD signal.

Apparently, the original PD signal has been totally submerged in noise.

3.2. AO-VMD Simulation

Subsequently, preprocessing is needed on the noisy PD signal to remove some noise and allow the characteristics of the PD signal. Common preprocessing techniques for PD signals include Fast Fourier Transform(FFT), Short-Time Fourier Transform(STFT) [31], and EMD [32]. However, the effectiveness of these preprocessing methods is generally suboptimal, as illustrated in Figure 4:

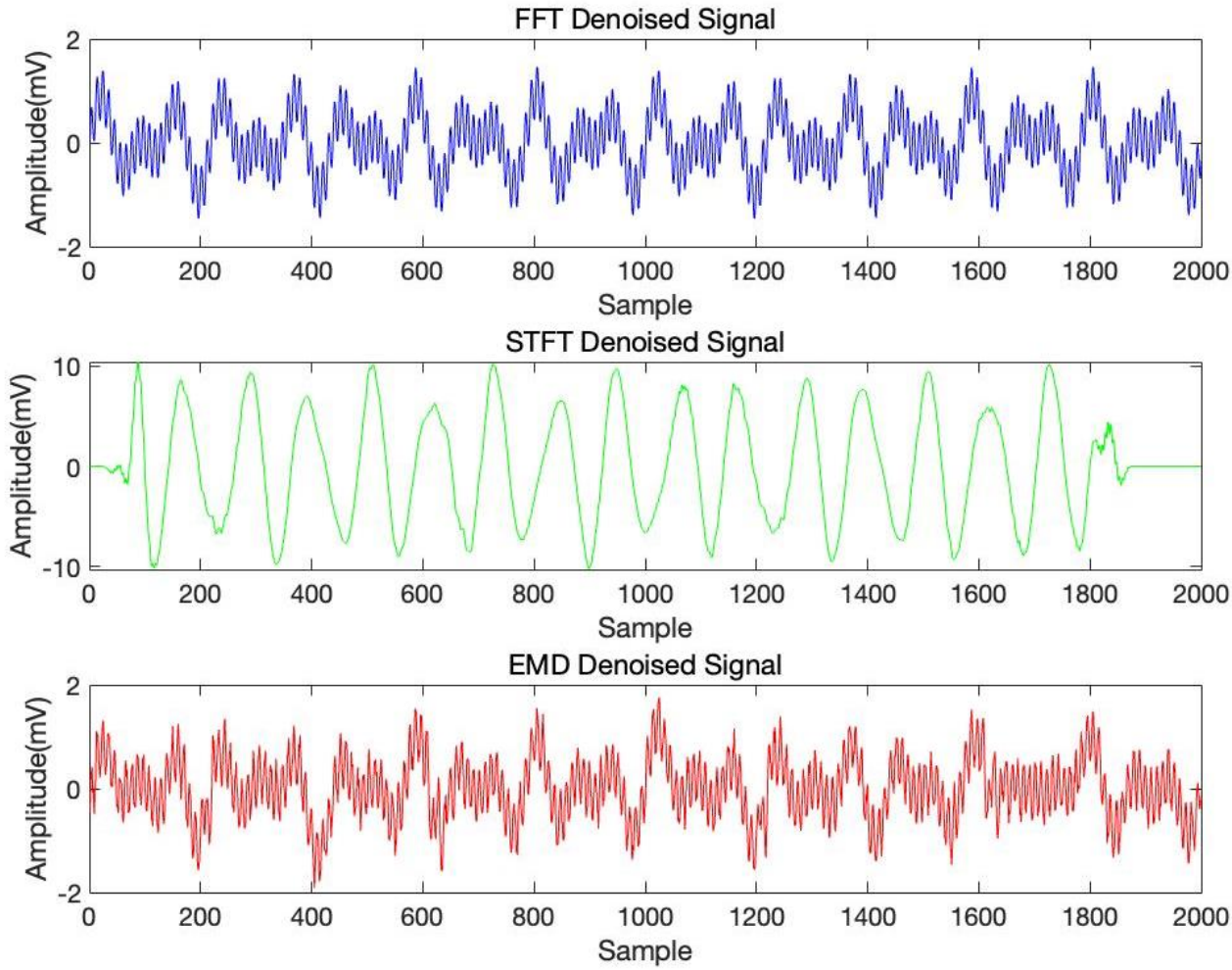


Figure 4. FFT, STFT, and EMD denoising on noisy PD signal.

Among them, FFT has almost no denoising results, while SFTF filters out both useful signals and noise. The EMD algorithm has a certain effect in the range of useful signals, such as around the sampling point 200 and 1200, but it is not enough to show the characteristics of the original signal.

In this simulation, the VMD algorithm is used to carry out preprocessing. To enhance the effectiveness of the VMD preprocessing, the AO algorithm is employed for calibrating parameters critical to the VMD process. Through a series of iterations, the AO algorithm refines a set of parameters that minimize the objective function. In each cycle, the solution space is progressively explored and exploited, facilitating convergence towards a global optimum. This process ultimately determines the optimal parameter configuration for the VMD algorithm. Figure 5 depicts the entire parameter optimization process executed by the AO algorithm, where the inflection points on the curve represent the optimal parameters for VMD.

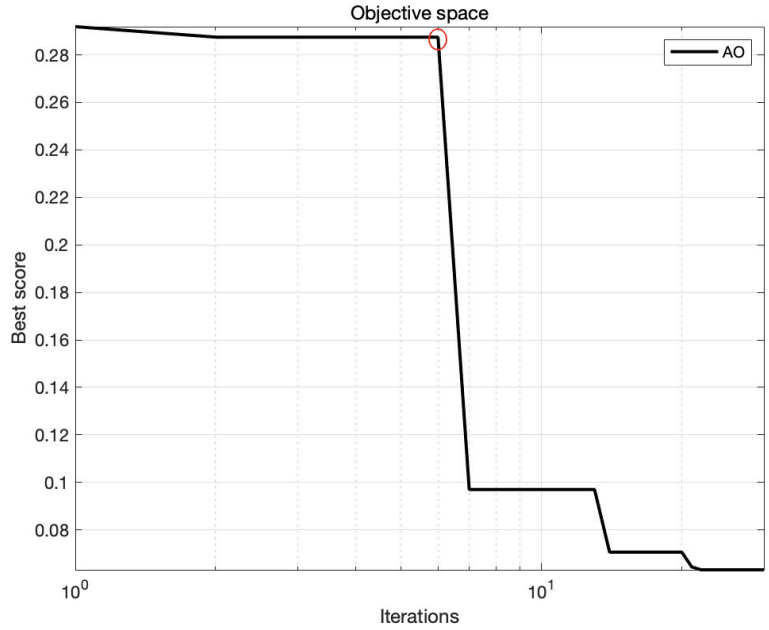


Figure 5. Searching the best parameters of VMD by AO.

After the AO-VMD decomposition of the noisy PD signal, the IMFs are shown in Figures 6 and 7 and the selection of modes is based on the kurtosis shown in Table 3.

Table 3. The optimal parameters of VMD.

Number of Decomposed Modes	Penalty Factor
8	1344

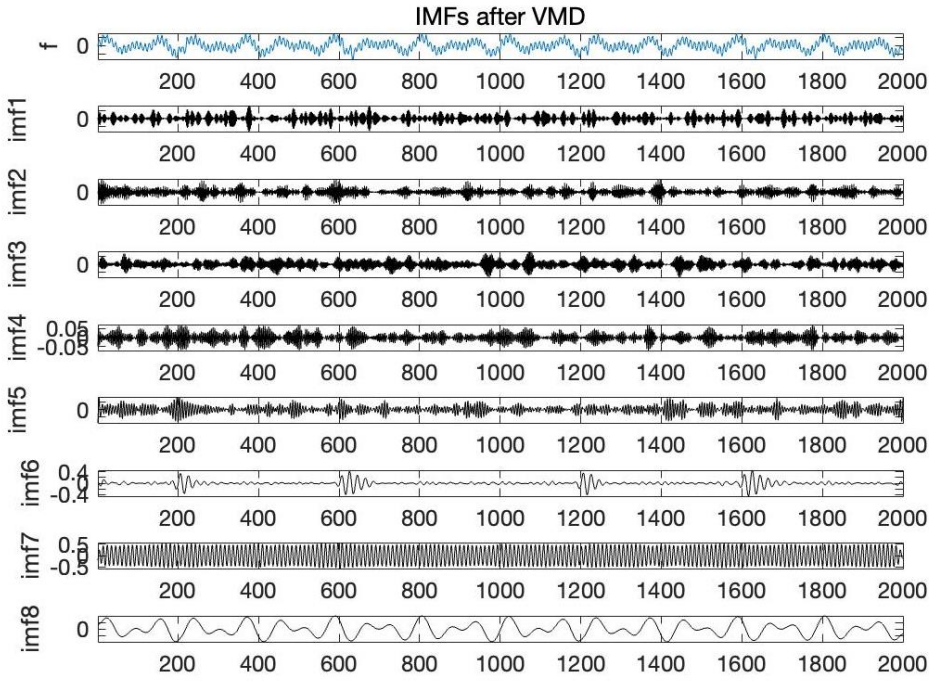


Figure 6. IMFs after VMD.

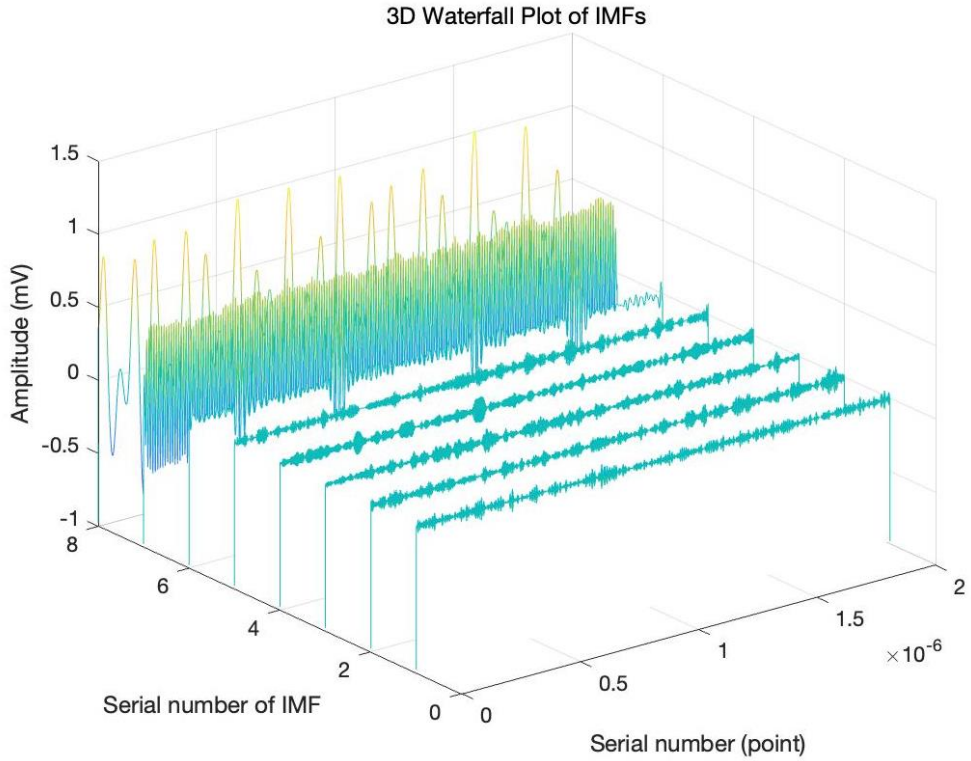


Figure 7. Waterfall IMFs after VMD.

The PD signal reconstructed based on the kurtosis showned in Table 4 criterion is displayed in Figure 8:

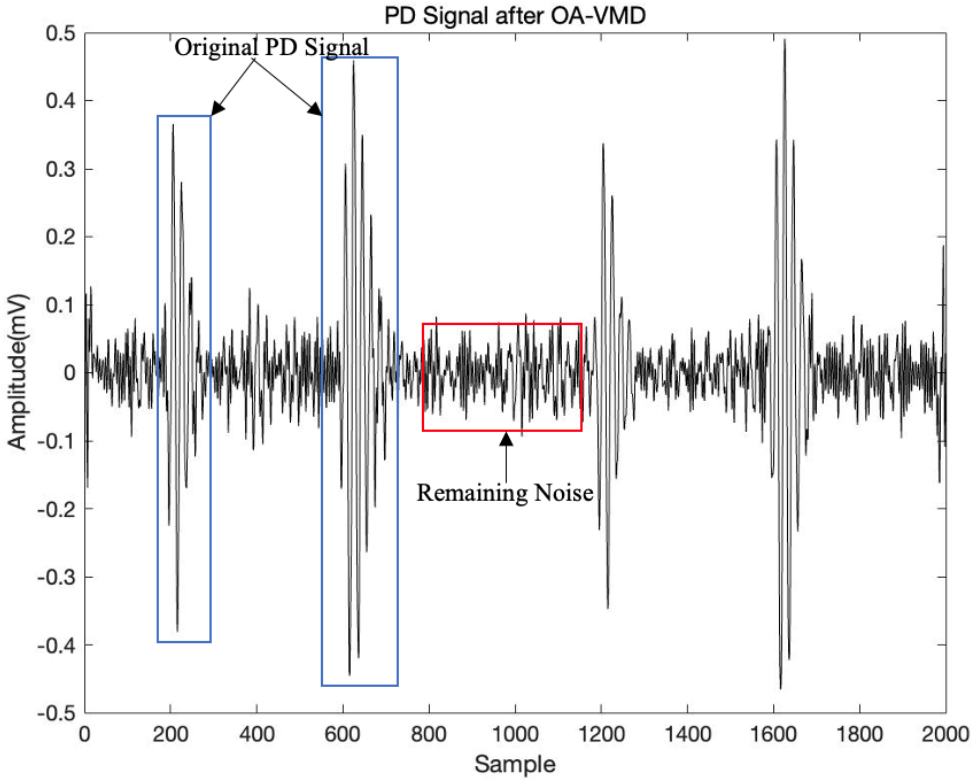


Figure 8. PD signal after AO-VMD.

Table 4. The kurtosis of IMFs.

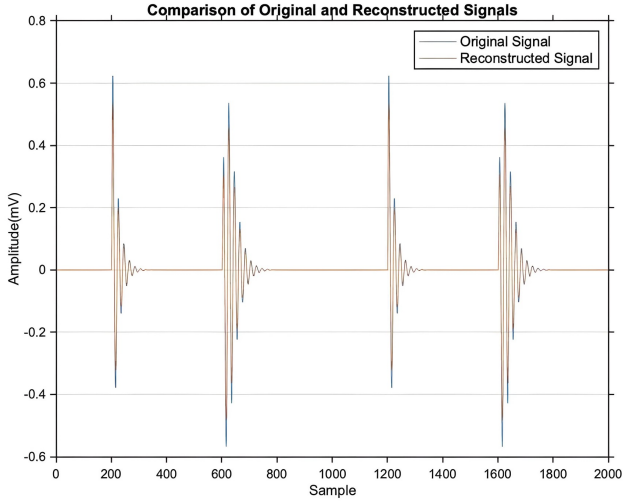
IMF Component	**IMF1**	**IMF2**	**IMF3**	**IMF4**
Kurtosis	2.7609	3.2024	2.9276	2.8780
IMF Component	**IMF5**	**IMF6**	**IMF7**	**IMF8**
Kurtosis	3.2310	12.3905	1.5268	2.1598

The preprocessing results demonstrate that VMD preprocessing, in contrast to other methods, is capable of eliminating a substantial portion of noise, thereby accentuating the intrinsic characteristics of the PD signal. However, residual noise still exists in the time domain, which may lead to misjudgment when extracting information such as the start and end points of the original signal. Therefore, further precise denoising is necessary.

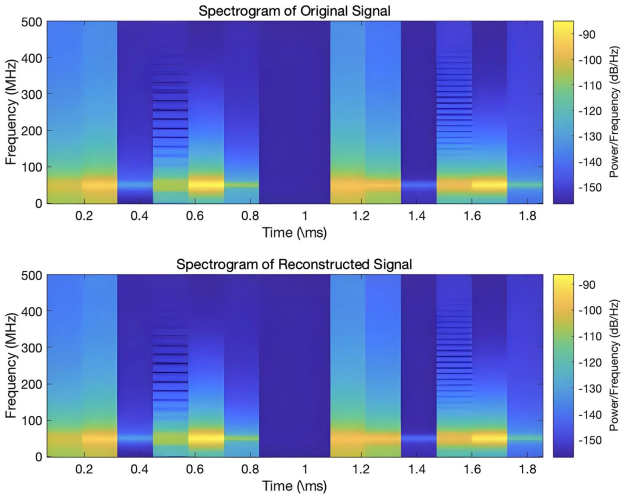
3.3. K-SVD Simulation

In this paper, an initial dictionary is constructed as a zero matrix with dimensions corresponding to the signal's sampling length multiplied by a configurable dictionary size N, ensuring dictionary redundancy as long as N is greater than 1. Subsequently, samples from the original PD data are randomly populated into the dictionary matrix to establish the initial learning dictionary.

Subsequently, based on the K-SVD and the OMP algorithm described in the previous section, the dictionary is updated column by column, which constitutes the entire process of dictionary learning. The learned dictionary is then used for sparse decomposition and reconstruction of the original signal, with the results depicted in Figure 9; at the same time, it displays the comparison of the original noiseless PD signal and the signal reconstructed using the AO-VMD and K-SVD algorithms, in both the time domain and time-frequency spectrum.



(a) Time domain comparison



(b) Time spectrum comparison

Figure 9. Comparison of original and reconstructed signal.

The comparison demonstrates that the noise in the PD signal has been effectively removed, and the original noiseless PD signal almost overlaps with the reconstructed signal, validating the efficacy of the proposed method in this paper.

In the evaluation criteria for dictionaries, sparsity and reconstruction error are two important indicators. Sparsity demonstrates the dictionary's ability to represent signals with a minimal number of non-zero coefficients. A higher sparsity indicates that the signal can

be represented with fewer elements, thus achieving a more efficient representation. On the other hand, the reconstruction error reflects the accuracy with which the original signal can be reconstructed using the dictionary. A lower reconstruction error signifies a more accurate representation of the original signal, indicating the effectiveness of the dictionary in preserving signal integrity during compression and reconstruction processes [33]. Figure 10 compares the performance of the K-SVD and Discrete Cosine Transform (DCT) dictionaries in denoising the same noisy PD signal using these two metrics.

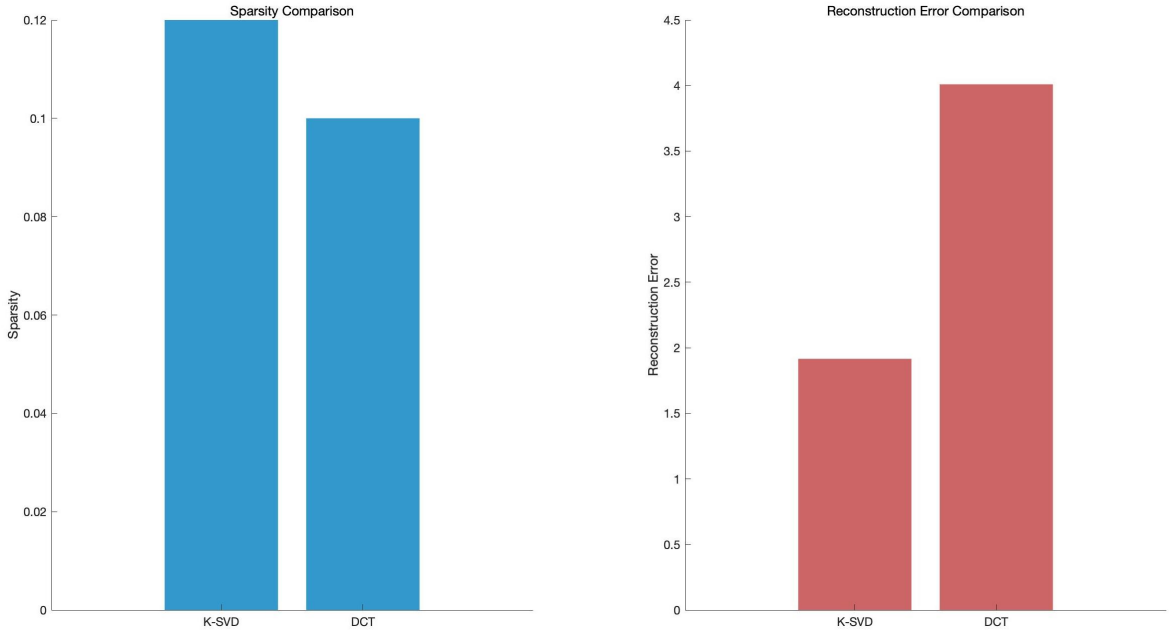


Figure 10. The comparison of K-SVD and DCT dictionaries.

The results demonstrate that the K-SVD learning dictionary is not only markedly effective in removing noise from PD signals but also exhibits superior performance compared to traditional analytical dictionaries such as DCT. This underscores the applicability and effectiveness of K-SVD in the denoising of PD signals.

4. Comparation of the Denoising Effect

To quantitatively evaluate the denoising effects of the methods, this study utilized four key evaluative metrics: the Root-Mean-Square Error (RMSE), Signal-to-Noise Ratio (SNR), Normalized Cross-Correlation (NCC), and Noise Reduction Ratio (NRR). The RMSE quantifies signal distortion, with lower scores indicating less distortion. The SNR assesses the denoising performance, with higher values denoting better noise suppression. The NCC measures the similarity between waveforms, with values nearing 1 implying a greater resemblance to the original, noise-free signal. Lastly, the NRR evaluates the extent of noise suppression achieved, with higher values representing more effective noise elimination [34]. The formulas for these metrics are delineated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|x(t) - y(t)|^2} \tag{14}$$

$$\mathrm{SNR} = 10 \cdot \log_{10}\left(\frac{\sum_{i=1}^{N} s(i)^2}{\sum_{i=1}^{N}(y(i) - s(i))^2}\right) \tag{15}$$

$$\mathrm{NCC} = \frac{\sum_{i=1}^{N} s(i) \cdot y(i)}{\sqrt{\sum_{i=1}^{N} s(i)^2} \cdot \sqrt{\sum_{i=1}^{N} y(i)^2}} \tag{16}$$

$$\mathrm{NRR} = 10 \cdot \left(\log_{10}(\sigma_1^2) - \log_{10}(\sigma_2^2) \right) \tag{17}$$

This paper compares Method A with the other three common PD denoise methods—AEEMD [10], the integration of SVD with VMD algorithm [11], and the Adaptive Wavelet Multilevel Soft Threshold algorithm [12]—on different input SNR PD signals to demonstrate the noise reduction effect. For example, when the input SNR is 7.76 dB, the signals after de-noising can be obtained as in Figure 11.

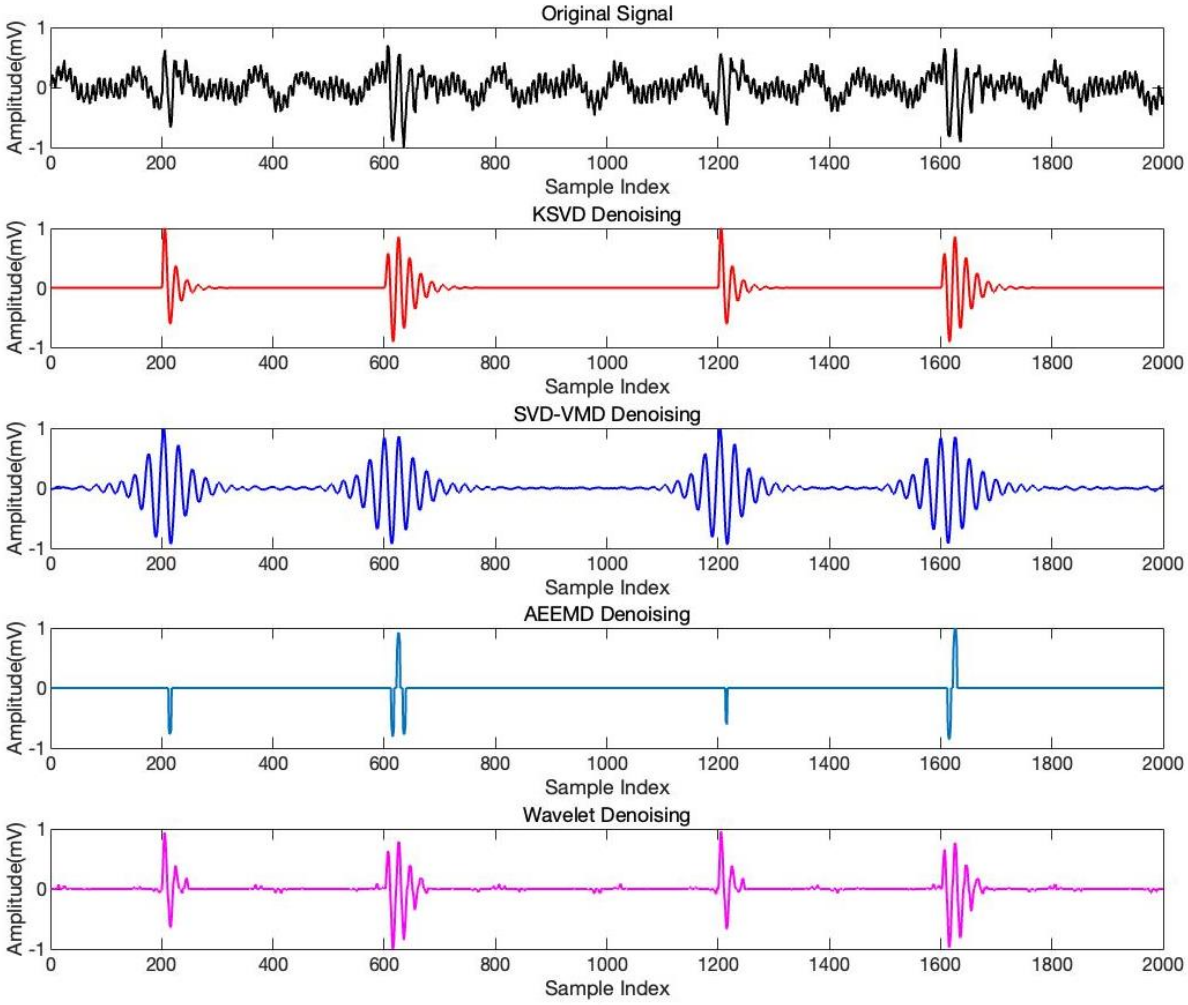


Figure 11. Comparison of PD signal denoising among four methods.

From the denoising results, it can be seen that the method improved by this paper can remove noise efficiently, and the characteristics of the original signal remain entirely. The SVD-VMD method can remove most noise, but it can be seen that the signal is excessively reconstructed, resulting in localized distortions, maybe because it is hard for this algorithm to fully identify signal characteristics, leading to imprecise reconstructions. The denoising result of the AEEMD algorithm indicates that the method over-denotes under these conditions, filtering out some components of the useful signal. The result of the Wavelet Multilevel Soft Thresholding denoising algorithm is better; although, some noise still remains in the time domain.

Subsequently, this paper compares the performance of these four methods under various signal-to-noise ratios and illustrates the outcomes in Figure 12.

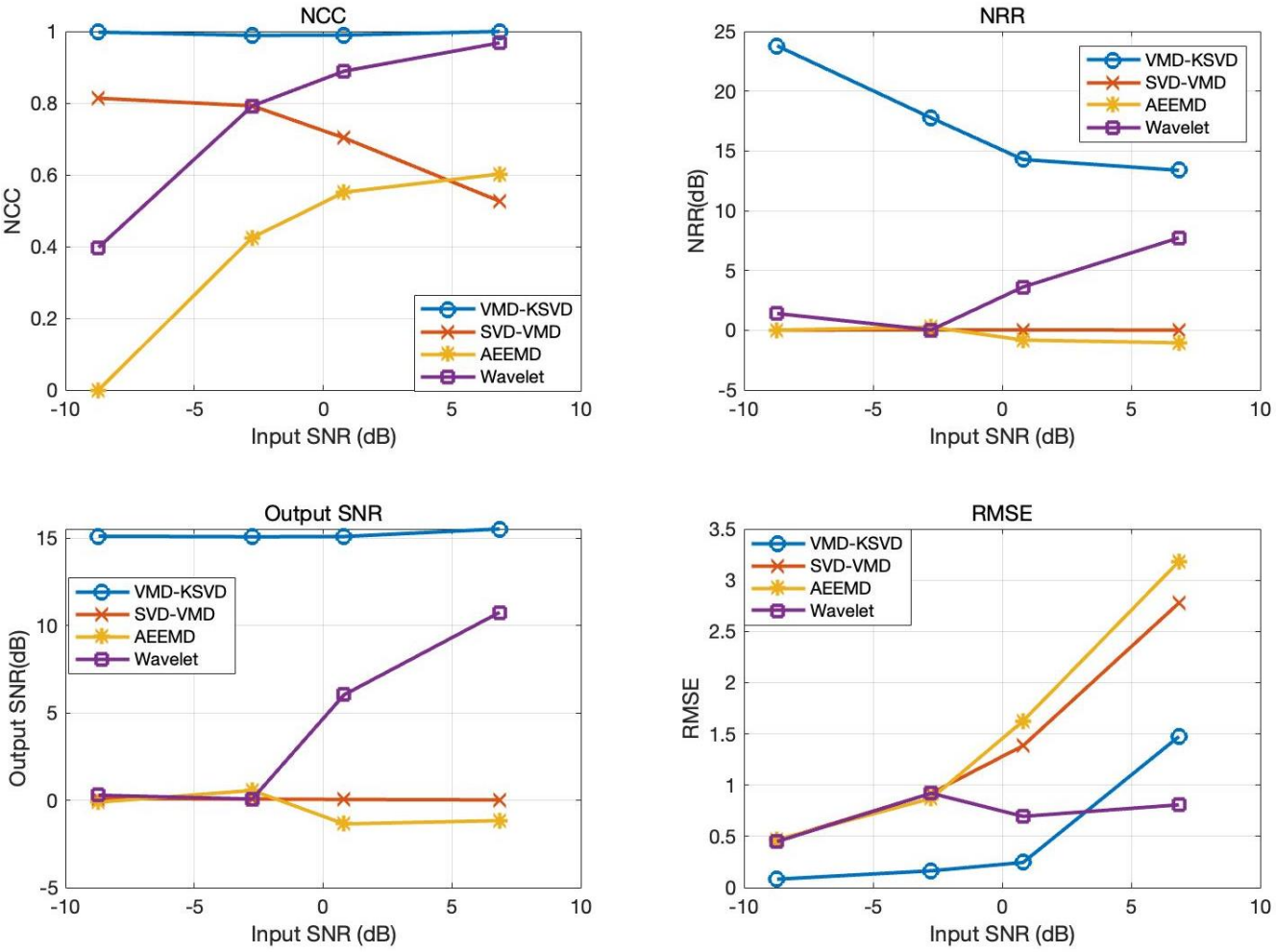


Figure 12. The comparison results in different input SNR.

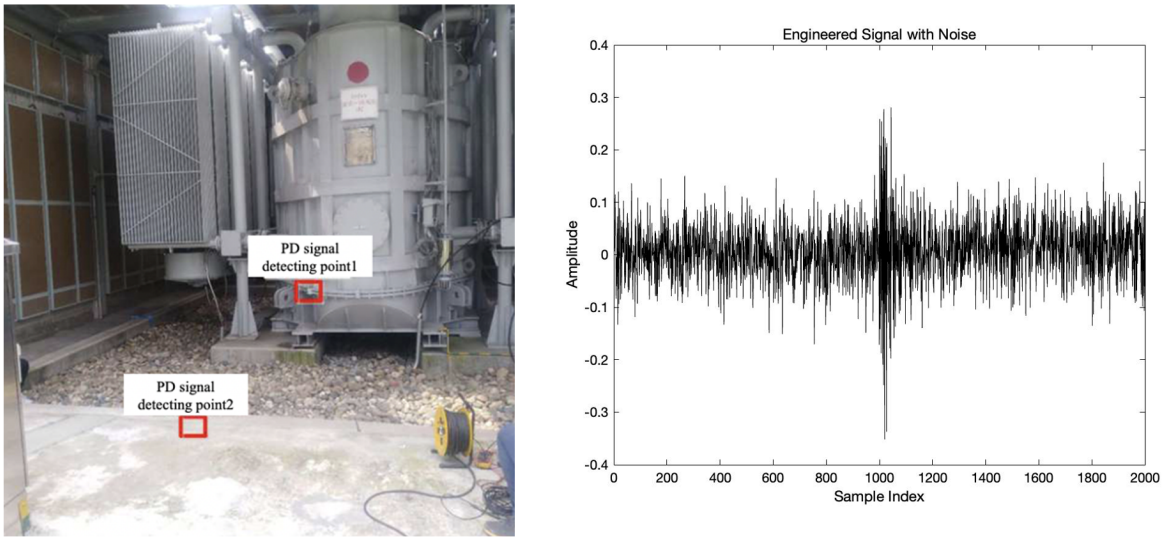
In the above figure, the blue lines represent the method proposed in this paper. It can be observed that this method maintains good performance metrics across various signal-to-noise ratios. This demonstrates the effectiveness and superiority of the proposed method, capable of stable noise reduction and signal extraction in various noise environments.

5. Denoising Analysis of the Actual PD Signal

The engineering scenario for detecting PD signals at a 500 kV high-voltage shunt reactor is shown in Figure 13. Its sensor is an ultrahigh-frequency sensor. In the detection mode, the signal is modulated, and the sampling frequency is 400 MHz.

By using the proposed method, the SVD-VMD algorithm, AEEMD algorithm, and Wavelet Multilevel Soft Thresholding algorithm, the measured signals were denoised as shown in Figure 14.

In the provided figure, the denoising result by the method improved by this paper demonstrates superior noise reduction with a notably smooth signal profile, suggesting effective noise suppression. The SVD-VMD denoising technique demonstrates limited effectiveness, as evident from the residual noise and the lack of enhancement in the useful signal. The AEEMD method shows pronounced peaks in the denoised signal, which could either denote preserved signal characteristics or insufficient noise reduction. The Wavelet method, while successful in removing a significant portion of noise and preserving signal characteristics, still leaves residual noise spikes in the time domain, which could potentially impact the accuracy of feature extraction and subsequent signal analysis.



(a) Points of the detection (b) Actual PD signal with noise

Figure 13. Actual PD signal.

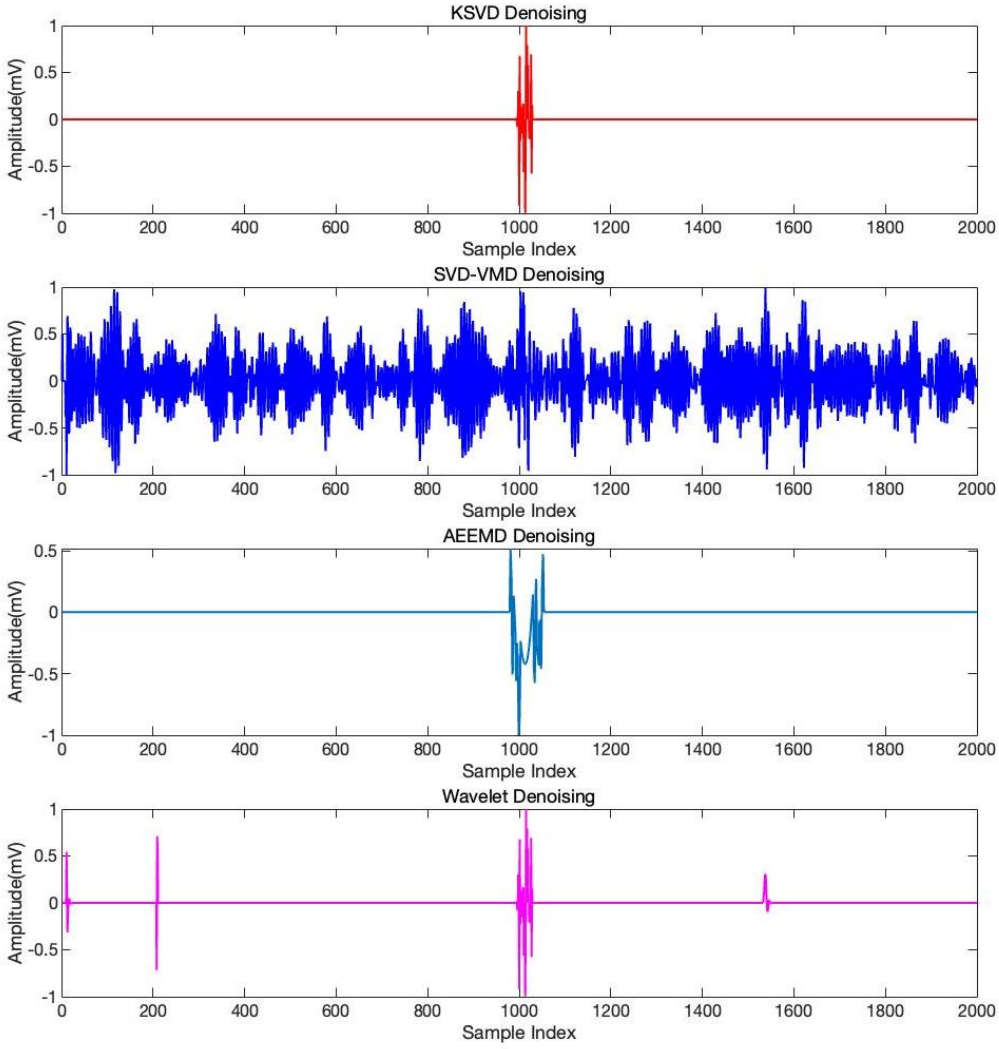


Figure 14. Comparison of denoising results for actual PD signals using four methods.

6. Conclusions

This paper proposes a PD signal denoising method which is based on the AO-VMD and K-SVD algorithms. The optimization algorithm is employed to fine-tune the parameters of the VMD, facilitating an initial noise reduction stage that accentuates the intrinsic characteristics of the original PD signal. Subsequently, the K-SVD algorithm is applied to eliminate residual noise, yielding a purified or significantly noise-reduced PD signal. The efficacy of the proposed algorithm is demonstrated through simulations and actual PD signals, where it consistently excels in noise reduction. Comparative analyses with other denoising strategies—SVD-VMD, AEEMD, and Wavelet Multilevel Soft Thresholding—reinforce the superior noise reduction capabilities of the proposed approach, confirming its potential for improved PD signal analysis.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app14072755/s1, Supplementary Materials Data.

Author Contributions: J.Z.: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing—review and editing; Z.L.: Formal analysis; Investigation; Methodology, Software, Validation, writing—original draft, Writing—review and editing; X.B.: Resources, Formal analysis, Methodology, Visualization. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article and Supplementary Materials.

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

Abbreviations

The following abbreviations are used in this manuscript:

PD	Partial Discharge
VMD	Variation Mode Decomposition
OMP	Orthogonal Matching Pursuit
EMD	Empirical mode decomposition
SVD	Singular Value Decomposition
K-SVD	K-Singular Value Decomposition
SNR	Signal-to-Noise Ratio
RMSE	Root Mean Square Error
NCC	Normalized Cross-Correlation
NRR	Noise Reduction Ratio
AEEMD	Adaptive Ensemble Empirical Mode Decomposition
AWMST	Adaptive Wavelet Multilevel Soft Threshold
AO	Adaptive Optimization

References

1. Fang, W.; Chen, G.; Li, W.; Xu, M.; Xie, W.; Chen, C.; Wang, W.; Zhu, Y. A PRPD-Based UHF Filtering and Noise Reduction Algorithm for GIS Partial Discharge. *Sensors* **2023**, *23*, 6763. [CrossRef] [PubMed]
2. Wang, Y.; Chen, P.; Zhao, Y.; Sun, Y. A Denoising Method for Mining Cable PD Signal Based on Genetic Algorithm Optimization of VMD and Wavelet Threshold. *Sensors* **2022**, *22*, 9386. [CrossRef] [PubMed]
3. Ma, C.; Li, H.; Zhou, W.; Yu, J.; Wang, L.; Yang, S.; Hu, S. Background Noise of Partial Discharge Detection and Its Suppression in Complex Electromagnetic Environment. In Proceedings of the 2018 IEEE International Conference on High Voltage Engineering and Application (ICHVE), Athens, Greece, 10–13 September 2018.
4. Jahangir, H.; Hajipour, E.; Vakilian, M.; Akbari, A.; Blackburn, T.; Phung, B.T. A method to capture and de-noise partial discharge pulses using discrete wavelet transform and ANFIS. *Int. Trans. Electr. Energy Syst.* **2014**, *25*, 2696–2712. [CrossRef]

5. Lin, M.-K.; Tai, C.; Tang, Y.-W.; Su, C.-C. Partial discharge signal extracting using the empirical mode decomposition with wavelet transform. In Proceedings of the 2011 7th Asia-Pacific International Conference on Lightning, Chengdu, China, 1–4 November 2011. [CrossRef]
6. Xu, Y.; Yan, J.; Wang, Y.; He, R.; Wang, J.; Geng, Y. Novel denoizing method for partial discharge signals using singular value decomposition and spectral subtraction. *IET Sci. Meas. Technol.* **2022**, *17*, 105–114. [CrossRef]
7. Yang, X.; Huang, H.; Shu, Q.; Zhang, D.; Chen, B. Partial Discharge Signal Extraction Method Based on EDSSV and Low Rank RBF Neural Network. *IEEE Access* **2021**, *9*, 9744–9752. [CrossRef]
8. Jin, T.; Li, Q.; Mohamed, M.A. A Novel Adaptive EEMD Method for Switchgear Partial Discharge Signal Denoising. *IEEE Access* **2019**, *7*, 58139–58147. [CrossRef]
9. Onuki, M.; Tanaka, Y. SVD for Very Large Matrices: An Approach with Polar Decomposition and Polynomial Approximation. In Proceedings of the 2018 IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, 17–20 November 2018; pp. 954–963. [CrossRef]
10. Sun, C.; Ju, P.; Li, D. Research on Partial Discharge Signal Denoising Based on Adaptive Ensemble Empirical Mode Decomposition Algorithm. *Electr. Technol.* **2022**, *23*, 67–72 + 78.
11. Lei, Z.; Wang, F.; Li, C. A Denoising Method of Partial Discharge Signal Based on Improved SVD-VMD. *IEEE Trans. Dielectr. Electr. Insul.* **2023**, *30*, 2107–2116. [CrossRef]
12. Sun, S.; Sun, Y.; Li, Y.; Ma, S.; Zhang, L.; Hu, Y. An Adaptive Wavelet Multilevel Soft Threshold Algorithm for Denoising Partial Discharge Signals. In Proceedings of the 2019 IEEE Sustainable Power and Energy Conference (iSPEC), Beijing, China, 21–23 November 2019; pp. 1874–1878. [CrossRef]
13. Liu, H.; Xiang, J. Autoregressive model-enhanced variational mode decomposition for mechanical fault detection. *IET Sci. Meas. Technol.* **2019**, *13*, 843–851. [CrossRef]
14. Zhang, Y.; Li, B.; Li, J.; Xing, Z. Rail Corrugation Identification Method Based on Parameter optimization VMD and SPWVD. In Proceedings of the 2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC), Banff, AB, Canada, 9–11 May 2019; pp. 110–115. [CrossRef]
15. Zhang, J.; Wu, J. A New Feature Extraction for Rolling Bearing Using Sparse Representation Based on Improved K-singular Value Decomposition and VMD. In Proceedings of the 2021 7th International Conference on Condition Monitoring of Machinery in Non-Stationary Operations (CMMNO), Guangzhou, China, 20–22 May 2021; pp. 27–31. [CrossRef]
16. De Azevedo Silva, F.T.; De Oliveira Mota, H. Partial discharge signal processing using overcomplete dictionaries and sparse representations. In Proceedings of the 2017 IEEE Electrical Insulation Conference (EIC), Baltimore, MD, USA, 11–14 June 2017; pp. 388–391. [CrossRef]
17. Avelar, A.S.O.; Vasconcelos, F.H.; Mota, H.D. Partial Discharge Signal Denoising Using Morphological Component Analysis. In Proceedings of the 2019 4th International Symposium on Instrumentation Systems, Circuits and Transducers (INSCIT), Sao Paulo, Brazil, 26–30 August 2019, pp. 1–6. [CrossRef]
18. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [CrossRef]
19. Zhang, D.; Zhang, Z.; Chen, Z.; Zhou, Y.; Li, F.; Chi, C. Wind power interval prediction based on variational mode decomposition and the fast gate recurrent unit. *Front. Energy Res.* **2023**, *10*, 1–13. [CrossRef]
20. Abualigah, L.; Yousri, D.; Abd Elaziz, M.; Ewees, A.A.; Al-qaness, M.A.A.; Gandomi, A.H. Aquila Optimizer: A novel meta-heuristic optimization algorithm. *Comput. Ind. Eng.* **2021**, *157*, 107250. [CrossRef]
21. Di, X.; Liu, S.; Li, D.; Ma, W.; Mei, X. AC Series Arc Fault Diagnosis Method Based on AO-VMD Multidimensional Feature Extraction. *J. Phys.* **2023**, *2636*, 012024. [CrossRef]
22. Shen, Y.; Zheng, W.; Yin, W.; Xu, A.; Zhu, H. Feature Extraction Algorithm Using a Correlation Coefficient Combined with the VMD and Its Application to the GPS and GRACE. *IEEE Access* **2021**, *9*, 17507–17519. [CrossRef]
23. Cai, C.; Li, Y.; Su, Z.; Zhu, T.; He, Y. Short-Term Electrical Load Forecasting Based on VMD and GRU-TCN Hybrid Network. *Appl. Sci.* **2022**, *12*, 6647. [CrossRef]
24. Chan, J.Q.; Illias, H.A.; Othman, M. Partial Discharge Localization Techniques: A Review of Recent Progress. *Energies* **2023**, *16*, 2863. [CrossRef]
25. Mathis, H. On the kurtosis of digitally modulated signals with timing offsets. In Proceedings of the 2001 IEEE Third Workshop on Signal Processing Advances in Wireless Communications, Taiwan, China, 20–23 March 2001. [CrossRef]
26. Morales-Perez, C.; Rangel-Magdaleno, J.; Peregrina-Barreto, H.; Martínez-Carballido, J. Technique for Signal Noise Reduction based on Sparse Representation. In Proceedings of the 2018 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), Ixtapa, Mexico, 14–16 November 2018; pp. 1–5. [CrossRef]
27. Sofuoglu, S. E.; Aviyente, S. Sparse Discriminative Tensor Dictionary Learning for Object Classification. In Proceedings of the 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Anaheim, CA, USA, 26–29 November 2018; pp. 1341–1345. [CrossRef]
28. Aharon, M.; Elad, M.; Bruckstein, A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **2006**, *54*, 4311–4322. [CrossRef]

29. Guojun, L.; Yong, W.; Le, L.; Shaofeng, G.; Haiqing, N.; Xuemei, W. Suppressing white noise in PD signal based on wavelet entropy and improved threshold function. In Proceedings of the 2017 IEEE 11th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED), Tinos, Greece, 29 August–1 September 2017, pp. 475-479. [CrossRef]
30. Yang, X.; Cai, Y.; Zhang, J.; Cheng, L.; Wang, F. Periodic Narrow-Band Noise Denoising of Partial Discharge Signal Through Time-Frequency Analysis. In Proceedings of the 2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2), Wuhan, China, 30 October–1 November 2020; pp. 2248–2251. [CrossRef]
31. Wang, Y.; Ji, S.; Xu, H. Non-stationary Signals Processing Based on STFT. In Proceedings of the 2007 8th International Conference on Electronic Measurement and Instruments, Xi'an, China, 16–18 August 2007; pp. 3-301–3-304. [CrossRef]
32. Khayam, U.; Surandaka, Y.A. Design, implementation, and testing of partial discharge signal processing system. In Proceedings of the 2016 2nd International Conference of Industrial, Mechanical, Electrical, and Chemical Engineering (ICIMECE), Yogyakarta, Indonesia, 6–7 October 2016; pp. 175–179. [CrossRef]
33. Rajesh, K.; Negi, A. Heuristic Based Learning of Parameters for Dictionaries in Sparse Representations. In Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, 18–21 November 2018; pp. 1013–1019. [CrossRef]
34. Ahmed, Z.; Lehtonen, M.; Ghulam, M.; Hussain, A. Analysis and De-Noising of Partial Discharge Signals in Medium Voltage XLPE Cables. Available online: https://core.ac.uk/download/pdf/80718741.pdf (accessed on 16 May 2016).