Mechanical Fault Diagnosis of High Voltage Circuit Breakers Based on Variational Mode Decomposition and Multi-Layer Classifier

Nantian Huang [1,*], Huaijin Chen [1], Guowei Cai [1], Lihua Fang [1] and Yuqiang Wang [2]

[1] School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China; chjin1990@126.com (H.C.); caigw@mail.nedu.edu.cn (G.C.); zang_fang0412@163.com (L.F.)
[2] State Grid Jibei Electric Power Co., Ltd. Maintenance Branch, Beijing 102488, China; wangyq181@163.com
[*] Correspondence: huangnantian@126.com; Tel.: +86-135-7855-5676

Abstract: Mechanical fault diagnosis of high-voltage circuit breakers (HVCBs) based on vibration signal analysis is one of the most significant issues in improving the reliability and reducing the outage cost for power systems. The limitation of training samples and types of machine faults in HVCBs causes the existing mechanical fault diagnostic methods to recognize new types of machine faults easily without training samples as either a normal condition or a wrong fault type. A new mechanical fault diagnosis method for HVCBs based on variational mode decomposition (VMD) and multi-layer classifier (MLC) is proposed to improve the accuracy of fault diagnosis. First, HVCB vibration signals during operation are measured using an acceleration sensor. Second, a VMD algorithm is used to decompose the vibration signals into several intrinsic mode functions (IMFs). The IMF matrix is divided into submatrices to compute the local singular values (LSV). The maximum singular values of each submatrix are selected as the feature vectors for fault diagnosis. Finally, a MLC composed of two one-class support vector machines (OCSVMs) and a support vector machine (SVM) is constructed to identify the fault type. Two layers of independent OCSVM are adopted to distinguish normal or fault conditions with known or unknown fault types, respectively. On this basis, SVM recognizes the specific fault type. Real diagnostic experiments are conducted with a real SF_6 HVCB with normal and fault states. Three different faults (i.e., jam fault of the iron core, looseness of the base screw, and poor lubrication of the connecting lever) are simulated in a field experiment on a real HVCB to test the feasibility of the proposed method. Results show that the classification accuracy of the new method is superior to other traditional methods.

Keywords: mechanical fault diagnosis; high voltage circuit breakers; acceleration sensor; variational mode decomposition; local singular value; one-class support vector machines

1. Introduction

As an integral part of the power system, high-voltage circuit breakers (HVCBs) are responsible for the control and protection of the system. HVCB faults will directly harm system reliability, causing significant outage costs. Therefore, the study of fault diagnostic methods for HVCBs is urgent. An inquiry about HVCB faults by the International Council on Large Electric Systems (CIGRE) showed that 39% of minor faults and 44% of major faults are of mechanical origin [1]. Hence, the research on mechanical fault diagnosis of HVCBs has practical significance. Vibration signals generated during the opening/closing operations of HVCBs contain certain important information associated with the mechanical state of breakers. Runde et al. [2] demonstrated through an extensive HVCB diagnostic test that vibration analysis is a suitable and reliable noninvasive diagnostic method for HVCBs. Analysis

of HVCB vibration signals collected by acceleration sensor has been widely used in the state detection and fault diagnosis of HVCBs [3–9].

HVCB vibration signals have a strong transient and wide frequency distribution. The signal acquisition equipment must have a high sampling rate. An acceleration sensor has high accuracy, wide frequency and amplitude responses, small size, and is easy to install; thus, it is widely used in vibration data acquisition [3–11].

Signal processing is conducted after vibration signals are obtained by the acceleration sensor, extracting the signal features for fault diagnosis. The HVCB vibration signals during operation are nonstationary and nonlinear. Traditional signal-processing methods, such as Fourier transform (FT) [12], are unsuitable for vibration signal analysis and processing. Conversely, time-frequency analysis methods, including wavelet packet decomposition (WPD) [3], empirical mode decomposition (EMD) [6,7], and local mean decomposition (LMD) [9], can analyze HVCB vibration signals well. Wavelet analysis can represent the local characteristics of signals both in time and frequency domains; thus, it is widely used in mechanical fault diagnosis [13]. However, wavelet transform is essentially an adjustable windowed FT with the limitation of energy leakage [14]. Wavelet basis function and decomposition scale are difficult to select in practical applications. The EMD proposed by Huang et al. is a completely adaptive signal analysis method, which is suitable for nonstationary signals analysis [15]. However, the EMD algorithm has a problem of mode aliasing. EMD is sensitive to noise and sampling, and its algorithmic nature lacks mathematical theory [16]. The LMD algorithm is similar to EMD; consequently, it also has the above-mentioned disadvantages of the EMD algorithm.

Variational mode decomposition (VMD) is a new adaptive signal-processing method proposed by Dragomiretskiy et al. (2014) [16]. This method introduces an entirely non-recursive VMD model and translates the decomposition problem into a variational one. Each mode and corresponding center frequency are continuously updated by solving the optimal solutions of the variational problem. VMD method has a solid theoretical foundation and good noise robustness. It has been successfully applied to the propagating mode extraction of microwave waveguide [17], the classification of power quality events [18], speech signals, and mechanical fault detection [19,20]. The features of vibration signals can be easily extracted from the IMFs of VMD.

After the feature extraction of fault vibration signals, a classifier should be used for fault type identification. Neural networks (NNs) [3] and SVM [6,7] achieve good classification accuracy in HVCB fault recognition. NNs have better capacities of self-learning and non-linear pattern recognition [21]. However, the determination of various parameters of NNs is difficult, and finding the optimal configuration of NNs is time consuming [22]. SVM algorithm is based on statistical learning theory and structural risk minimizing principle. It is suitable for classification problems with small sample size [23]. Sufficient fault sample data with all fault types are difficult to obtain because HVCB operations are seldom. SVM cannot correctly identify a new fault type because of lack of training samples. Consequently, the sample of an unknown fault type is recognized as a normal sample. In this case, the SVM-based classifier hardly meets the reliability requirements. The occurrence of a new fault is also unpredictable in an HVCB operation. Recent research results show that only a few types of HVCB mechanical fault with training samples can be recognized [3,6,7]. Thus, a new fault type in the mechanical system cannot be identified successfully.

OCSVM [24] is a classifier that can be trained by using only one type of samples. It is widely used in the field of fault diagnosis and detection [25–28]. The classification boundary of OCSVM is closer to the object samples than that of SVM. Hence, OCSVM has a lower false acceptance rate, i.e., lower possibility of no-object samples misrecognized as object samples. Accordingly, OCSVM has a superior fault detection capability for HVCBs.

This paper proposes a new method based on VMD and MLC for diagnosing HVCB mechanical faults. An acceleration sensor is used to acquire HVCB vibration data. The vibration data are then decomposed by VMD to obtain the corresponding IMFs. On this basis, local singular value decomposition (LSVD) is utilized to extract the vibration features. The MLC used for fault recognition

is constructed by two OCSVMs and an SVM. The first OCSVM (OCSVM1) trained by normal samples determines whether a test sample is in the fault state. The second OCSVM (OCSVM2) trained by all available fault samples identifies whether the type of the fault samples is new. SVM is adopted to identify the known fault type. Comparative experiments are designed with the measured fault data of real HVCBs to validate the new method.

2. Vibration Data Acquisition and Fault Diagnosis Process

2.1. Acceleration Sensor

Acceleration is a physical quantity that characterizes an object's movement. Vibration essence is the reciprocating movement of an object. Thus, vibration data can be obtained by measuring the acceleration with an acceleration sensor. Integrated electronics piezo electric (IEPE) acceleration sensor is widely used and can obtain HVCB vibration signals well. It also has the following advantages: small size, light weight, low noise, and anti-interference capability. This paper adopts a CA-YD-182A piezoelectric acceleration sensor to measure HVCB vibration data. The main technical indicators of the CA-YD-182A include ± 250 g (g = 9.8 m/s^2) measuring range, 20 mV/g sensitivity, 40 kHz natural frequency, 10 kHz frequency response, a maximum output voltage of 6 V, and a weight of 9 g.

2.2. Data Acquisition System

In this paper, the CA-YD-182A acceleration sensor and an NI 9234 data acquisition card are applied to build the vibration signal acquisition system for HVCBs. The measuring object is the LW9-72.5 series, which is an outdoor high-voltage SF$_6$ circuit breaker. The acquisition system of HVCB vibration signals and its block diagram are shown in Figure 1. The acceleration sensor is used to measure the vibrational state of HVCBs and produce the corresponding voltage signals. The voltage signals are digitized by using the NI 9234. When the circuit breaker receives an opening command, the system starts sampling. The sampling rate is 25.6 kS/s, and the sampling period is 150 ms.
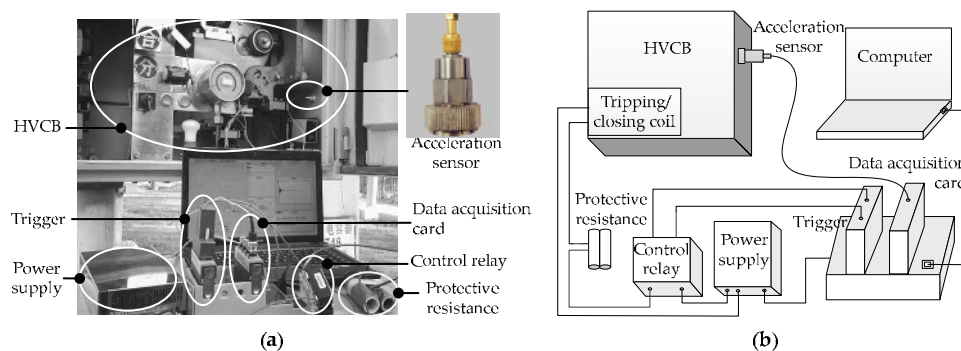

Figure 1. (**a**) The vibration signal acquisition system of HVCBs; (**b**) The block diagram of the acquisition system.

In an actual measurement, the installation location and the method of the acceleration sensor affect the performance of the acquisition system. The principle for selecting measurement position is that the sensor does not affect the normal operation of the measured object, and the position is close to the object or the most concerned point of the object. In this paper, the sensor is installed on the mechanism box near the operating mechanism. Acceleration sensor installation methods mainly include handheld magnetic adsorption, glue bonding, and screw fixation. An adhesive mounting is selected according to the actual demand of the diagnosis of HVCB mechanical fault.

2.3. Fault Diagnosis Process

The new method proposed in this paper consists of three parts: feature extraction, state detection, and fault recognition. In feature extraction, the features of vibration signals are extracted by using VMD and LSVD methods. In state detection, the normal or fault state of the HVCB is determined by OCSVM1. In fault recognition, the fault type is recognized using OCSVM2 and SVM. The fault diagnosis process is shown in Figure 2, in which OCSVM1 is trained by the normal samples, and OCSVM2 is trained by all available fault samples (fault samples with known types). For a test sample, the new method recognizes a fault as a normal condition, known with a specific fault type, or unknown without specific fault type.
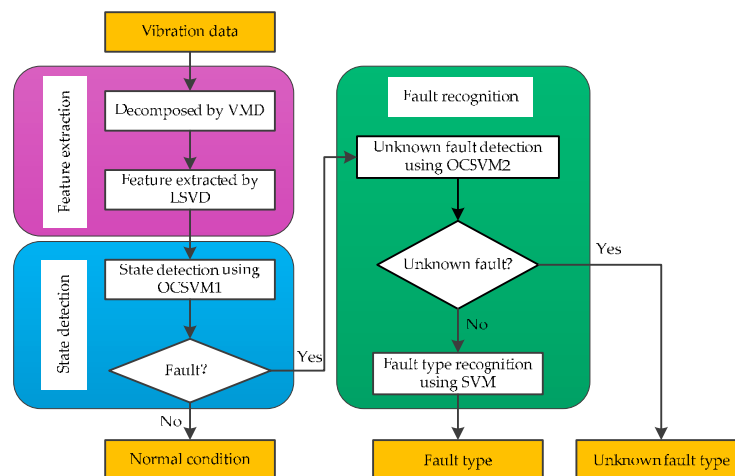

Figure 2. Fault diagnosis process of the proposed method.

3. VMD

3.1. VMD Theory

VMD process is the solving of a variational problem. Therefore, this algorithm can be divided into the construction and solution of the variational problem. VMD involves three key concepts: classic Wiener filtering, Hilbert transform, and frequency mixing.

- Construction of the variational problem

  The VMD turns an input signal h into K modes. Each mode m_k is mostly compact around a center frequency ω_k. The variational problem can be described as seeking the K modes to make the sum of all bandwidths of the modes minimum. The constraint condition is that the sum of each mode is equals to the input signal h. The detailed construction scheme is as follows: (1) The associated analytic signal of each mode m_k is computed by the Hilbert transform to obtain the unilateral frequency spectrum; (2) The frequency spectrum of each mode is tuned to the respective estimated center frequency by mixing with the exponential $e^{-j\omega_k t}$; (3) The bandwidth is estimated through the squared L^2-norm of the gradient of the demodulated signal. The constrained variational problem is written as:

$$\begin{cases} \min\limits_{\{m_k\},\{\omega_k\}} \left\{ \sum\limits_{k=1}^{K} \| \partial_t \left[\left(\delta\left(t\right) + \frac{j}{\pi t} \right) * m_k\left(t\right) \right] e^{-j\omega_k t} \|_2^2 \right\} \\ \text{s.t.} \sum\limits_{k=1}^{K} m_k = h \end{cases}, \tag{1}$$

  where $\{m_k\} = \{m_1, m_2, \cdots, m_K\}$ is the set of all modes, $\{\omega_k\} = \{\omega_1, \omega_2, \cdots, \omega_K\}$ are the corresponding center frequencies, $\delta\left(t\right)$ is the Dirac function, and * denotes the convolution.

- Solution of the variational problem

A constrained variational problem can become unconstrained by introducing a Lagrange multiplier α and a quadratic penalty factor η. The Lagrange multiplier enforces constraints strictly; and the quadratic penalty factor guarantees the reconstruction fidelity of the signal with Gaussian noise. The augmented Lagrange expression is as follows [29]:

$$L\left(\{m_k\},\{\omega_k\},\alpha\right) = \eta \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta\left(t\right) + \frac{j}{\pi t} \right) * m_k\left(t\right) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| h\left(t\right) - \sum_{k=1}^{K} m_k\left(t\right) \right\|_2^2 + \left\langle \alpha\left(t\right), h\left(t\right) - \sum_{k=1}^{K} m_k\left(t\right) \right\rangle, \quad (2)$$

The alternating direction method of multipliers (ADMM) solves the saddle point of the augmented Lagrange. m_k^{n+1}, ω_k^{n+1}, and α^{n+1} are alternately updated using the ADMM approach. The updates of m_k^{n+1}, ω_k^{n+1}, and α^{n+1} are as follows (see Appendix A for the detailed solution process):

$$\hat{m}_k^{n+1}\left(\omega\right) = \frac{\hat{h}\left(\omega\right) - \sum\limits_{i \neq k} \hat{m}_i\left(\omega\right) + \frac{\hat{\alpha}\left(\omega\right)}{2}}{1 + 2\eta\left(\omega - \omega_k\right)^2}, \quad (3)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left|\hat{m}_k\left(\omega\right)\right|^2 d\omega}{\int_0^\infty \left|\hat{m}_k\left(\omega\right)\right|^2 d\omega}, \quad (4)$$

$$\alpha^{n+1} = \alpha^n + \tau\left(h - \sum_{k=1}^{K} m_k^{n+1}\right), \quad (5)$$

where $\hat{\cdot}$ denotes the FT of $\cdot$, and τ is the update parameter of the Lagrange multiplier. The mode m_k^{n+1} can be obtained as the real part of the inverse FT of $\hat{m}_k^{n+1}$. VMD estimates the mode m_k and center frequency ω_k constantly through an iteration. For a given convergence tolerance $e > 0$, the termination condition of this iteration is:

$$\sum_k \| m_k^{n+1} - m_k^n \|_2^2 / \| m_k^n \|_2^2 < e, \quad (6)$$

3.2. Simulated Vibration Signal Analysis Based on VMD

The vibration signal of HVCBs consists of a series of vibration events. It can be described by a set of exponentially decaying sinusoidal signals, which is as follows [5]:

$$V\left(t\right) = \sum_{i=1}^{n} A_i e^{-\mu_i\left(t - t_i\right)} \sin\left[2\pi f_i\left(t - t_i\right)\right] \varepsilon\left(t - t_i\right), \quad (7)$$

where n is the number of vibration events, $\varepsilon(t)$ is the unit step function, A_i is the amplitude of the ith vibration event, μ_i is attenuation coefficient, f_i is oscillation frequency, and t_i is the starting time of vibration. The vibration events V_1 to V_5 generated by MATLAB compose the simulated vibration signal for HVCBs. The parameter of each vibration event is shown in Table 1. The waveforms of the simulated vibration signal and each vibration event with a signal-to-noise ratio (SNR) of 20 dB are shown in Figure 3, in which the sampling rate is 25.6 kS/s.

Table 1. The parameter of each vibration event.

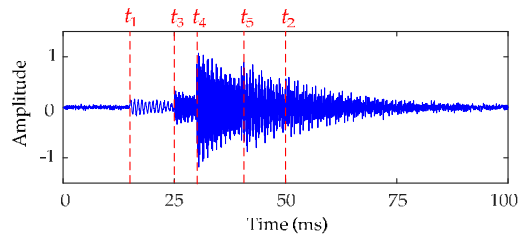| Vibration Events	A_i	t_i (ms)	f_i (Hz)	μ_i
V_1	0.15	15	1200	80
V_2	0.2	50	3000	50
V_3	0.3	25	4500	95
V_4	1.0	30	5500	75
V_5	0.5	40	7000	60

Figure 3. The waveform of the simulated HVCB vibration signal.

EMD has been proven to be a suitable method for the vibration signal processing of HVCBs. We mainly compare the performances of VMD and EMD to decompose this simulated vibration signal (with an SNR of 20 dB). In addition, VMD performance is also compared with a few new and improved EMD-related methods, i.e., LMD [9], ensemble EMD (EEMD) [30], and complete EEMD (CEEMD) [31]. The original vibration events and IMFs decomposed by these five methods are shown in Figure 4.
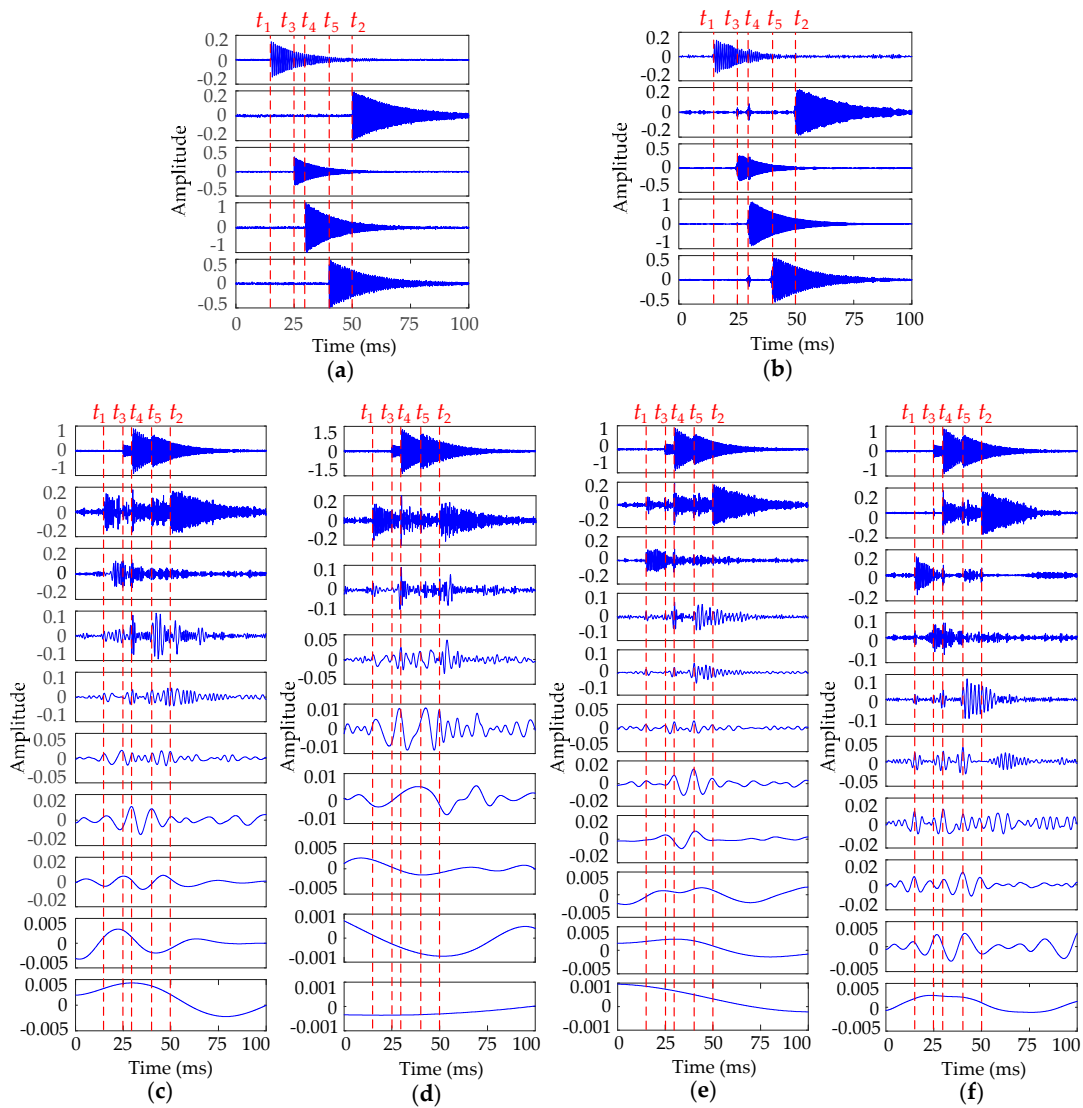

Figure 4. (**a**) Original vibration events; (**b**) IMFs decomposed by VMD method; (**c**) IMFs decomposed by EMD method; (**d**) PFs decomposed by LMD method; (**e**) IMFs decomposed by EEMD method; (**f**) IMFs decomposed by CEEMD method.

When the number of vibration events and corresponding parameters of the simulated signal are known, the performance of each signal-processing method can be determined by comparing the correlation degrees of their modes and the original vibration events. Figure 4b shows that the signal is decomposed into five IMFs by VMD, and each IMF is mostly the same as the corresponding vibration event in Figure 4a. That is, the VMD approach can decompose vibration signals thoroughly. Conversely, we obtain approximately 10 IMFs through EMD, LMD, EEMD, and CEEMD approaches. In Figure 4c, the modes decomposed by EMD have a serious mode aliasing problem, especially for the second mode. Although LMD is better than EMD in some aspects, such as the endpoint effect suppression and algorithm speed, it shows almost the same performance as EMD with modal aliasing in this study. Both EEMD and CEEMD can eliminate modal aliasing to a certain extent, but the latter has a better effect. EMD and its derivation algorithms cannot effectively separate the vibration events from the composite vibration signal because of the limitation of its algorithmic nature. Consequently, the characteristics (such as starting time and spectrum) of each mode obtained by EMD and other similar methods are almost irrelevant with the original signal characteristics; thus, these modes fail to reflect the physical significance of each vibration event, i.e., existence of false modes. Therefore, the VMD method is more suitable for the feature extraction of HVCB vibration signals.

3.3. Determining the Number of K Modes of VMD

The number of *K* modes should be predefined in VMD method. Each mode component of VMD contains local features of the original signal at a center frequency and different time scales. A great number of *K* modes suggests that VMD has abundant frequency components. The reconstructed signals by *K* modes will be highly similar to the original signal. The measured vibration signals of HVCBs contain a large number of vibration components; thus, the analysis should focus on the main vibration event rather than all vibration components. Therefore, we determine the number of *K* modes by comparing the similarity measure between the reconstructed and original signals.

Distance measure is a common measure of pattern similarity. Normalized distance (ND) is selected to evaluate the similarity between the original and reconstructed signals using different mode numbers. The ND of two discrete signals $p = (p_1, p_2, \ldots p_n)$ and $q = (q_1, q_2, \ldots q_n)$ is defined as:

$$d(p,q) = \frac{\| p - q \|}{\| p \| + \| q \|} = \frac{\left[\sum\limits_{i=1}^{n} (p_i - q_i)^2 \right]^{1/2}}{\left(\sum\limits_{i=1}^{n} p_i^2 \right)^{1/2} + \left(\sum\limits_{i=1}^{n} q_i^2 \right)^{1/2}}, \tag{8}$$

VMD is used to decompose the simulated vibration signal with different *K* and compute the corresponding reconstructed signals. The NDs between the reconstructed and original signals with different *K* are shown in Figure 5.
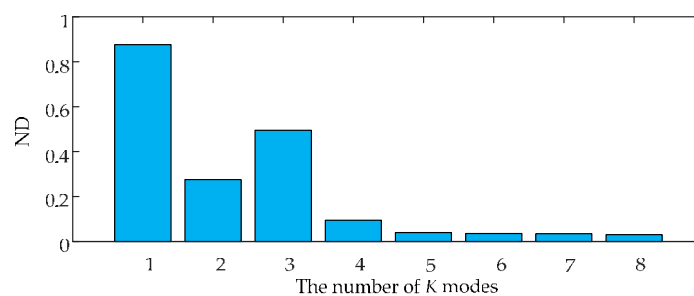

Figure 5. The NDs between the reconstructed signals and the original signals with different *K*.

Figure 5 shows that the ND almost does not change when *K* becomes greater than 5 and remains at a near-zero value. In this case, the similarity between the original and reconstructed signals is

maximized, i.e., the reconstructed signal contains all the main information characteristics of the original signal. Hence, the optimal number of modes of VMD is set at 5, which is consistent with the number of vibration events contained in the original vibration signal. Accordingly, ND method is effective for mode number selection.

4. Principles of SVM and OCSVM

4.1. SVM

SVM, proposed by Vapnik in 1995, has many advantages in solving small-sample, high-dimensional, and nonlinear pattern recognition problems [32]. The basic principles of SVM are mapping the data samples from a low-dimensional space to a high-dimensional one and making the indivisible low-dimensional data become linearly separable. A linear partition is then used to determine the classification boundary. The classification principle of SVM is shown in Figure 6.
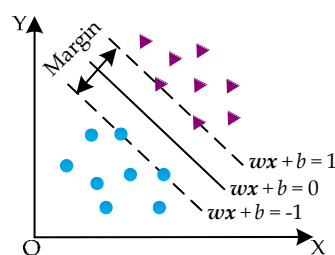

Figure 6. The classification principle of SVM.

We suppose that the training sample set $(x_i, y_i)(i = 1, 2, \cdots, l; x_i \in R^d, y_i \in \{-1, 1\})$ is composed of two different sample classes. The samples are linearly separable when a hyperplane $w \cdot x + b = 0$ can correctly divide them into two classes, i.e., when they satisfy:

$$\begin{cases} w \cdot x_i + b \geq 1, & y_i = 1 \\ w \cdot x_i + b \leq -1, & y_i = -1 \end{cases}, \quad i = 1, 2, \cdots, l, \tag{9}$$

The samples that satisfy $|w \cdot x_i + b| = 1$ are called support vectors. The distance between two classes of support vectors is $2/\|w\|$, i.e., the classification margin is $2/\|w\|$. The goals of SVM are to seek the optimal hyperplane under the constraints in Equation (9), and make $2/\|w\|$ as maximum and $\|w\|^2/2$ as minimum:

$$\begin{cases} \min_{w,b} \frac{1}{2}\|w\|^2 \\ \text{s.t. } y_i(w \cdot x_i + b) \geq 1, \ i = 1, 2, \cdots, l \end{cases}, \tag{10}$$

For most situations, the samples in the training set are linearly inseparable. SVM introduces a slack variable ξ_i to reduce the constraint to $y_i(w \cdot x_i + b) \geq 1 - \xi_i$. Meanwhile, penalty factor C is introduced to control the degree of punishment to error-classifying samples. Thus, the objective function becomes:

$$\begin{cases} \min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l} \xi_i \\ \text{s.t. } y_i(w \cdot x_i + b) \geq 1 - \xi_i, \ i = 1, 2, \cdots, l \end{cases}, \tag{11}$$

This problem can be solved through saddle point of the Lagrange function, which is constructed as:

$$L(w, b, \alpha_i) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{l} \alpha_i [y_i(w \cdot x_i + b) - 1], \tag{12}$$

where $\alpha_i > 0$ is Lagrange coefficient. Equation (12) is converted into the following dual problem according to dual theory:

$$\begin{cases} \max Q\left(\boldsymbol{\alpha}\right) = \sum\limits_{i=1}^{l} \alpha_i - \frac{1}{2}\sum\limits_{i=1}^{l}\sum\limits_{j=1}^{l} \alpha_i\alpha_j y_i y_j \left(\boldsymbol{x}_i \cdot \boldsymbol{x}_j\right) \\ \text{s.t. } \sum\limits_{i=1}^{l} \alpha_i y_i = 0,\ 0 \le \alpha_i \le C \end{cases}, \tag{13}$$

The optimal solution of the quadratic programming problem $\boldsymbol{\alpha} = \left[\alpha_1, \alpha_2, \cdots, \alpha_l\right]^T$ can be obtained, followed by optimal $\boldsymbol{w}$ and b. The optimal decision function is:

$$f\left(\boldsymbol{x}\right) = \text{sgn}\left[\sum\limits_{i=1}^{l} \alpha_i y_i \left(\boldsymbol{x}_i \cdot \boldsymbol{x}\right) + b\right], \tag{14}$$

where sgn(z) is sign function, which equals +1 for $z \ge 0$ and -1 otherwise.

For a nonlinear classification problem, SVM uses kernel function $\phi(\boldsymbol{x})$ to map the sample data from a low-dimensional space to a high-dimensional, making these samples linearly separable. The kernel function is defined as follows:

$$K\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = \phi\left(\boldsymbol{x}_i\right) \cdot \phi\left(\boldsymbol{x}_j\right), \tag{15}$$

After introducing the kernel function, Equation (13) becomes:

$$\begin{cases} \max Q\left(\boldsymbol{\alpha}\right) = \sum\limits_{i=1}^{l} \alpha_i - \frac{1}{2}\sum\limits_{i=1}^{l}\sum\limits_{j=1}^{l} \alpha_i\alpha_j y_i y_j K\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) \\ \text{s.t. } \sum\limits_{i=1}^{l} \alpha_i y_i = 0,\ 0 \le \alpha_i \le C \end{cases}, \tag{16}$$

The decision function becomes:

$$f\left(\boldsymbol{x}\right) = \text{sgn}\left[\sum\limits_{i=1}^{l} \alpha_i y_i K\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) + b\right], \tag{17}$$

4.2. OCSVM

OCSVM also maps the training data into a high-dimensional feature space by using the kernel function. OCSVM aims to separate sample data from the origin with a maximum margin, which is different from SVM. The object and no-object samples are located on either side of the hyperplane. The classification principle of OCSVM is shown in Figure 7. For convenience, we still use $\{\boldsymbol{x}_i\}\ (i = 1, 2, \cdots, l; \boldsymbol{x}_i \in \boldsymbol{R}^d)$ to represent the training sample set.
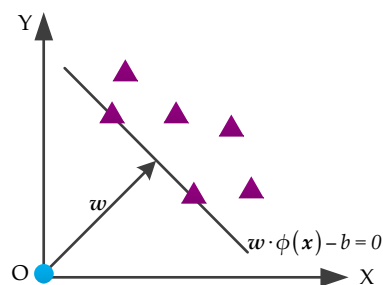

Figure 7. The classification principle of OCSVM.

Similar to SVM, the classification hyperplane of OCSVM is expressed as $\boldsymbol{w} \cdot \phi(\boldsymbol{x}) - b = 0$. OCSVM solves the following quadratic programming problem:

$$\begin{cases} \min \frac{1}{2}\| \boldsymbol{w} \|^2 + \frac{1}{vl}\sum_i \xi_i - b \\ \text{s.t. } \boldsymbol{w} \cdot \phi(\boldsymbol{x}_i) \geq b - \xi_i, \ \xi_i \geq 0 \end{cases}, \tag{18}$$

where $v \in (0,1]$ is the margin of error that controls the number of outliers. The decision function is as follows:

$$f(\boldsymbol{x}) = \text{sgn}(\boldsymbol{w} \cdot \phi(\boldsymbol{x}) - b), \tag{19}$$

The value of decision function $f(\boldsymbol{x})$ is +1 or -1 along with Equation (19). $f(\boldsymbol{x})$ is considered as the object sample when it takes the value of +1 in a test sample. Therefore, once $\boldsymbol{w}$ and b are solved, we can determine the sample class.

Lagrange multipliers are introduced to solve the above quadratic programming problem. The Lagrange function is as follows:

$$L(\boldsymbol{w},\xi,b,\alpha,\beta) = \frac{1}{2}\| \boldsymbol{w} \|^2 + \frac{1}{vl}\sum_i \xi_i - b - \sum_i \alpha_i(\boldsymbol{w} \cdot \phi(\boldsymbol{x}_i) - b + \xi_i) - \sum_i \beta_i \xi_i, \tag{20}$$

where α_i, $\beta_i \geq 0$ are Lagrange multipliers. We set the partial derivatives of variables $\boldsymbol{w}$, ξ, b in Equation (20) equal to zero, yielding:

$$\begin{cases} \boldsymbol{w} = \sum_i \alpha_i \phi(\boldsymbol{x}_i) \\ \alpha_i = \frac{1}{vl} - \beta_i \leq \frac{1}{vl}, \ \sum_i \alpha_i = 1 \end{cases}, \tag{21}$$

Combined with the kernel function in Equation (15), the dual form of this optimization problem is described as:

$$\begin{cases} \min_\alpha \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j K(\boldsymbol{x}_i, \boldsymbol{x}_j) \\ \text{s.t.}\sum_i \alpha_i = 1, \quad 0 \leq \alpha_i \leq \frac{1}{vl} \end{cases}, \tag{22}$$

The support vector is located on the hyperplane; thus b can be found by support vector $\boldsymbol{x}_i$ and the corresponding α_i:

$$b = \boldsymbol{w} \cdot \phi(\boldsymbol{x}_i) = \sum_{j=1}^{l} \alpha_j K(\boldsymbol{x}_j, \boldsymbol{x}_i), \tag{23}$$

The decision function together with Equation (15) can be transformed into a kernel expansion form:

$$f(\boldsymbol{x}) = \text{sgn}\left(\sum_{i=1}^{l} \alpha_i K(\boldsymbol{x}_i, \boldsymbol{x}_j) - b\right), \tag{24}$$

Figures 6 and 7 illustrate that the support vectors of OCSVM are on the classification hyperplane, whereas those of SVM are on both sides of the hyperplane with a certain distance. Accordingly, OCSVM can identify the non-target samples more accurately and has higher capability of fault identification than SVM in the fault diagnosis area of HVCBs.

5. Feature Extraction of Vibration Signal

5.1. Singular Value Decomposition (SVD)

SVD [33] is an important matrix decomposition method that is widely used in feature extraction. According to SVD theory, for an $m \times n$ matrix A ($A \in R^{m \times n}$), there must exist two orthogonal matrices $U_{m \times m}$ and $V_{n \times n}$, and a diagonal matrix Λ, satisfying:

$$\begin{cases} A = U \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix} V^T \\ \Lambda = \text{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_r)\,, \ r = \text{rank}\,(A) \end{cases}, \tag{25}$$

where $\lambda_i(i = 1, 2, \cdots, r)$ is the singular value of matrix A, and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq 0$. The singular value tends to correspond to the important information implied in the matrix, and the importance is in positive correlation with the value.

The SVD of a matrix has the following property:

We assume matrices $A, B \in R^{m \times n}$, and the singular values of A and B are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_R \geq 0$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_R \geq 0$, respectively, where $R = \min(m, n)$. Then:

$$|\lambda_i - \sigma_i| \leq \| A - B \|_2, \ i = 1, 2, \cdots, R, \tag{26}$$

This property indicates that when matrix A has slight disturbance, the changes in singular values are not greater than the spectral radius of the perturbation matrix. Hence, the singular values of a matrix are insensitive to the changes in matrix elements.

5.2. Feature Extraction Based on LSVD

In the feature extraction of the vibration signal of circuit breakers, a few energy-based features, such as the time segmentation energy entropy (TSEE), are often used as signal features [9]. However, the energy feature of the signal is sometimes not enough to reflect the fault characteristics of the signal accurately. SVD is an effective method to extract the algebraic feature of a matrix, which can better reflect the changes in the internal characteristics of the signal.

LSVD method is used in this study to extract HVCB vibration features to improve the disturbance detection capability of SVD. A sample sequence of length N can be decomposed into K IMFs by VMD. The data length of each IMF is also N. Hence, the size of the IMF matrix is $K \times N$. The research in [34] showed that the singular values of the entire matrix cannot indicate the local and detailed features of the matrix. For some faults of HVCBs, such as time delay fault, the singular values of the entire matrix tend not to reflect the fault characteristic information. Therefore, more detailed local information in the time domain is required. The local information of HVCB vibration signals at different time periods is obtained using LSVD method, which is as follows:

(1) VMD is used for decomposing HVCB vibration signals to obtain the IMF matrix.
(2) The IMF matrix is equally divided into 30 submatrices along the time axis. The size of each submatrix is $K \times (N/30)$.
(3) The 30 submatrices are decomposed by a series of SVDs, obtaining 30 singular value sequences.
(4) The singular values of each submatrix attenuate rapidly; thus the largest singular value of each submatrix $\lambda_{i\text{max}}$ is selected to construct the feature vector $F = [\lambda_{1\text{max}}, \lambda_{2\text{max}}, \cdots, \lambda_{30\text{max}}]$.

6. Experimental Results

6.1. Feature Analysis of Measured Vibration Signal

HVCB vibration data are collected using the acquisition system in Figure 1. Three types of fault are simulated in field experiments: (1) jam fault of the iron core (Fault I); (2) looseness of the base

screw (Fault II); and (3) poor lubrication of the connecting lever (Fault III). Excessive opening/closing operations will damage the circuit breaker; thus 40 samples of normal condition and 40 samples per fault type are collected through several experiments. Typical waveforms of four different types of vibration signals are shown in Figure 8.
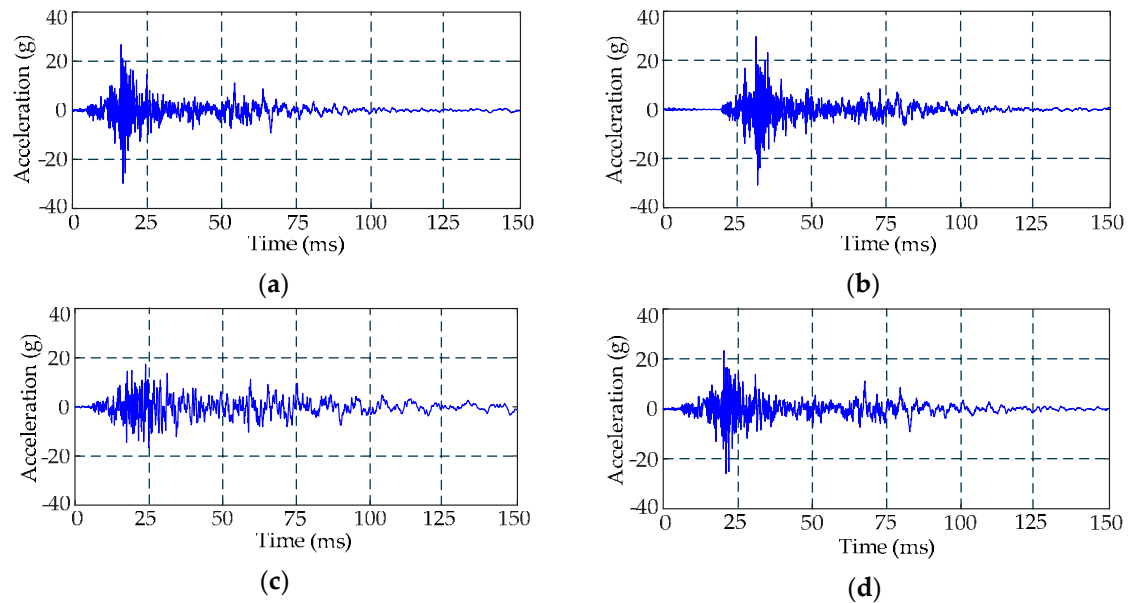

(**a**)

(**b**)

(**c**)

(**d**)

Figure 8. Waveforms of four types of vibration signals. (**a**) Normal condition; (**b**) Fault I; (**c**) Fault II; (**d**) Fault III.

As mentioned previously, the number of VMD modes should be predefined. According to the abovementioned method for determining the number of modes, we use VMD to decompose the four types of vibration signals with different K. The NDs between their corresponding reconstructed and measured signals are then computed, which are shown in Figure 9.
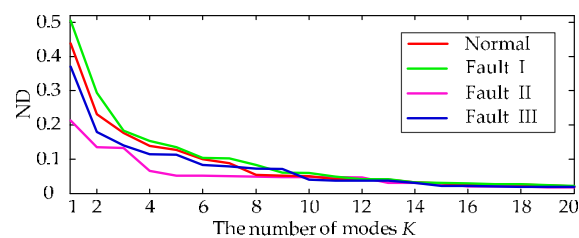

Figure 9. The NDs between the reconstructed signals and the measured signals with different K.

Figure 9 shows that the NDs of the four signal types decrease with the increase of K. When K is greater than 10, the changes in ND values show signs of leveling off. The number of K modes is set to 10 to guarantee that the four signal types can be effectively decomposed.

The normal and fault vibration signals are decomposed by VMD, and the corresponding IMFs are shown in Figure 10. Ten modes of each signal type are arranged from top to bottom based on the increase in center frequencies, and the red dashed line indicates the starting time t_s of a normal vibration signal.

Figure 10 indicates some characteristics of fault signals in the time or frequency domain. Compared with the normal state, the vibration of Fault I has a significant time delay. The amplitudes of the last seven modes of Fault II are significantly smaller than the normal state, i.e., the vibration

focuses on a lower-frequency area. The vibration time duration in different modes of Fault III is longer than other types of signals because of the poor lubrication of the connecting lever.
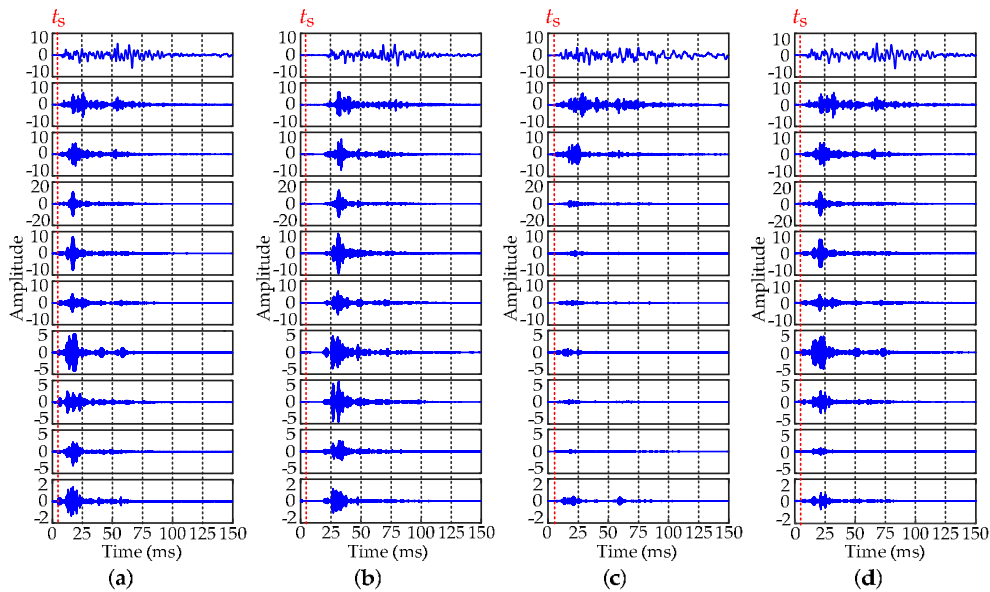

Figure 10. IMFs of the four types of vibration signals obtained by VMD. (**a**) Normal condition; (**b**) Fault I; (**c**) Fault II; (**d**) Fault III.

The LSVD method is adopted to extract the features of vibration signals. The LSV feature vectors of the normal and three types of fault conditions are shown in Figure 11. For clarity, each type only displays three feature vectors.
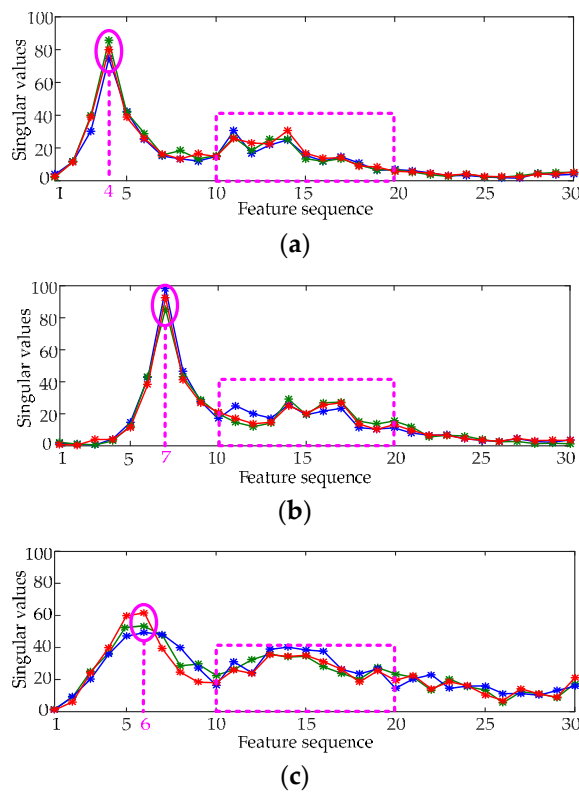

Figure 11. *Cont.*
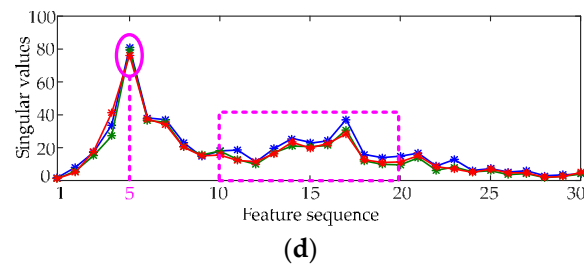
(**d**)

Figure 11. The LSV feature vectors of normal and fault signals. (**a**) Normal condition; (**b**) Fault I; (**c**) Fault II; (**d**) Fault III.

Figure 11 presents that the feature vectors of different types of vibration signals have significant differences. The peak of the feature vector of normal condition appears around the fourth feature, whereas that of Fault I appears around the seventh feature, that of Fault II appears around the sixth feature, and that of Fault III appears around the fifth feature. The variations in the 10th to 20th features of the four signals are also different. The classifier can make a good classification according to the differences among these feature vectors. These feature vectors roughly reflect the energy distributions of the corresponding vibration signals in the time domain from Figures 8 and 11.

We use the whole SVD (WSVD) method to extract the features of vibration signals, validating the LSV feature vectors. The entire matrix is directly decomposed into *K* (*K* = 10 here) singular values by SVD [35]. The whole singular value (WSV) feature vectors of the four types of vibration signals are shown in Figure 12.
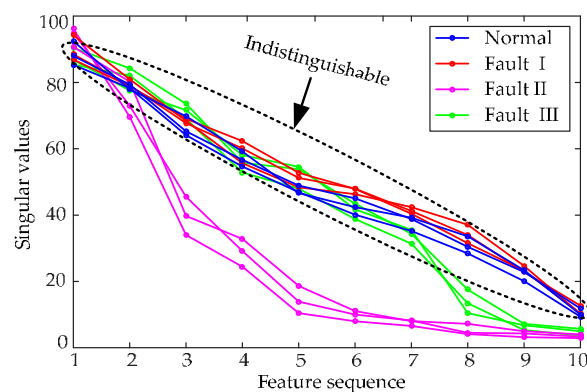

Figure 12. The WSV feature vectors of different types of vibration signals.

The WSVD method may not distinguish normal from Fault I signals, as presented in Figure 12. Fault I is essentially a time delay fault that contains the same vibration rules as normal signals. Thus, almost all the major elements of the IMF matrix of Fault I are the same as those of the normal signal. Consequently, the WSV feature of Fault I tends to be nearly equal to the normal condition. WSVD method cannot directly reflect the vibration laws of the original signal over time, unlike LSVD. Thus, LSVD approach is more suitable for the feature extraction of HVCB vibration signals.

6.2. Fault Classification Using MLC

The LSV feature vectors are entered into the MLC to achieve the relevant classification results. The MLC consists of three classifiers: OCSVM1, OCSVM2, and SVM. These classifiers need to be trained first. For each type of vibration signals, 40 vibration data are included. We select 20 data randomly as the training samples and the other 20 data as test samples. OCSVM1 is trained using normal training samples, whereas OCSVM2 and SVM are trained by fault training samples. SVM is the most widely used classifier in HVCB fault diagnosis and has achieved a good classification effect.

We compare the classification performances of MLC and SVM. The experiment results are shown in Table 2. "New Fault" in the Table refers to the new type of fault that has not been recorded before, i.e., the unknown fault type.

Table 2. Diagnosis results using MLC and SVM.

Classifier	Test Sample	Diagnosis Results					Accuracy
		Normal	Fault I	Fault II	Fault III	New Fault	
MLC	Normal	18	0	0	2	0	90%
	Fault I	0	20	0	0	0	100%
	Fault II	0	0	20	0	0	100%
	Fault III	0	0	0	20	0	100%
SVM	Normal	19	0	0	1	-	95%
	Fault I	0	20	0	0	-	100%
	Fault II	0	0	20	0	-	100%
	Fault III	3	0	0	17	-	85%

According to the results in Table 2, three types of fault states are correctly recognized by the MLC method, and their classification accuracies are 100%. Conversely, three samples of Fault III are recognized as normal samples by SVM, and the corresponding classification accuracy is 85%. This comparison shows that the new approach has a higher capability of fault identification. For normal state, two samples are wrongly classified by MLC and one by SVM. For HVCBs, normal samples that are recognized as fault samples will not cause accidents and outage cost. Moreover, the operational reliability of the device is not reduced by the new method. Therefore, the new method improves the accuracy of fault diagnosis while ensuring the reliability of HVCBs. When the WSV is selected as the input feature vector of the classifier in this paper, the corresponding classification results using MLC and SVM are shown in Table 3.

Table 3. Diagnosis results using MLC and SVM with the WSV feature.

Classifier	Test Sample	Diagnosis Results					Accuracy
		Normal	Fault I	Fault II	Fault III	New Fault	
MLC	Normal	14	6	0	0	0	70%
	Fault I	5	15	0	0	0	75%
	Fault II	0	0	19	0	1	95%
	Fault III	0	3	0	17	0	85%
SVM	Normal	13	7	0	0	-	65%
	Fault I	8	11	0	1	-	55%
	Fault II	0	0	20	0	-	100%
	Fault III	3	2	0	15	-	70%

The accuracy of fault diagnosis using WSVD method is lower than that using the LSVD method, as shown in Tables 2 and 3. It illustrates that the WSVD approach is unsuitable for the feature presentation of HVCB vibration signals. Besides, the entire classification accuracy of MLC remains higher than that of SVM in such a situation.

A new fault type without training sample appearing in test samples is also considered. We assume that Fault III is the new fault, and the training samples of Fault III do not participate in the training of OCSVM2 and SVM. The classification results are shown in Table 4. The test samples of Fault III are selected as the test sample set. The classification results of MLC and SVM are compared under this situation and are shown in Table 4.

Table 4 shows that when a new fault type occurs, SVM cannot accurately identify the fault samples because of the lack of corresponding training. All fault samples are recognized as the normal state to reduce the fault diagnosis accuracy of SVM significantly. Conversely, MLC can identify a fault state with 100% accuracy. Thus, the new method has higher accuracy for the diagnosis of unknown new

fault types. When a new fault is recognized, we can determine its specific fault type according to the overall report made by the maintenance personnel. In this way we can continue to accumulate fault samples and get more fault types.

Table 4. Diagnosis results of the new type of fault using MLC and SVM.

Classifier	Diagnosis Results				Accuracy
	Normal	**Fault I**	**Fault II**	**New Fault**	
MLC	0	0	0	20	100%
SVM	20	0	0	-	0

7. Conclusions

This paper proposes a diagnosis method for HVCB mechanical faults based on VMD and MLC. The simulation and practical tests demonstrate the following advantages of the new approach:

(1) Compared with EMD, the mode decomposed by VMD has a clearer physical meaning. The latter can reduce the influence of false modes for feature extraction and has a better property of feature presentation for vibration signals.
(2) LSV can characterize the local and detailed features of vibration signals accurately, and the fault signatures can be extracted more precisely using the LSVD method, especially for delay fault.
(3) MLC uses OCSVM to improve the ability to detect fault conditions. This method can identify unknown fault types. The diagnosis accuracy and the reliability of MLC are significantly enhanced compared with those of the SVM method.

Author Contributions: N.H. conceived and designed the experiments; G.C. provided important guidance; H.C. performed the experiments and analyzed the data; F.L. wrote the paper; Y.W. revised the paper.

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

Appendix A

As mentioned in the main text, the variational problem is written as:

$$
\begin{cases}
\min\limits_{\{m_k\},\{\omega_k\}} \left\{ \sum\limits_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * m_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\
\text{s.t.} \sum\limits_{k=1}^{K} m_k = h
\end{cases}
, \tag{A1}
$$

The corresponding augmented Lagrange is constructed as:

$$
L(\{m_k\},\{\omega_k\},\alpha) = \eta \sum\limits_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * m_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| h(t) - \sum\limits_{k=1}^{K} m_k(t) \right\|_2^2 + \left\langle \alpha(t), h(t) - \sum\limits_{k=1}^{K} m_k(t) \right\rangle, \tag{A2}
$$

The variational problem (A1) can be solved through the saddle point of the augmented Lagrange (A2). In this paper, ADMM approach is used to solve the saddle point, in which m_k^{n+1}, ω_k^{n+1}, and α^{n+1} are alternately updated.

The m_k^{n+1} is updated in the following way:

$$
m_k^{n+1} \leftarrow \arg\min\limits_{m_k} L\left(\left\{ m^{n+1} \right\}, \left\{ m_{i \geq k}^n \right\}, \left\{ \omega_i^n \right\}, \alpha^n \right), \tag{A3}
$$

The value of m_k^{n+1} can be expressed as:

$$m_k^{n+1} = \underset{m_k}{\arg\min} \left\{ \eta \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * m_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| h(t) - \sum_i m_i(t) + \frac{\alpha(t)}{2} \right\|_2^2 \right\}, \quad \text{(A4)}$$

For simplicity, $\cdot^{n+1}$ and $\cdot^n$ are omitted for the fixed directions $m_{i \neq k}$ and ω_k, respectively. They represent the most recent available updates. Problem (A4) can be solved in the frequency domain based on the Parseval-Plancherel theorem, which is as follows:

$$\hat{m}_k^{n+1} = \underset{\hat{m}_k, m_k}{\arg\min} \left\{ \eta \| j\omega \left[(1 + \text{sgn}(\omega + \omega_k)) \hat{m}_k(\omega + \omega_k) \right] \|_2^2 + \left\| \hat{h}(\omega) - \sum_i \hat{m}_i(\omega) + \frac{\hat{\alpha}(\omega)}{2} \right\|_2^2 \right\}, \quad \text{(A5)}$$

We use the variable $\omega - \omega_k$ instead of ω in the first term:

$$\hat{m}_k^{n+1} = \underset{\hat{m}_k, m_k}{\arg\min} \left\{ \eta \| j(\omega - \omega_k) \left[(1 + \text{sgn}(\omega)) \hat{m}_k(\omega) \right] \|_2^2 + \left\| \hat{h}(\omega) - \sum_i \hat{m}_i(\omega) + \frac{\hat{\alpha}(\omega)}{2} \right\|_2^2 \right\}, \quad \text{(A6)}$$

Equation (A6) can be written as the integral over the non-negative frequencies using Hermitian symmetry, which is as follows:

$$\hat{m}_k^{n+1} = \underset{\hat{m}_k, m_k}{\arg\min} \left\{ \int_0^\infty 4\eta (\omega - \omega_k)^2 |\hat{m}_k(\omega)|^2 + 2 \left| \hat{h}(\omega) - \sum_i \hat{m}_i(\omega) + \frac{\hat{\alpha}(\omega)}{2} \right|^2 d\omega \right\}, \quad \text{(A7)}$$

The solution of the quadratic optimization problem is:

$$\hat{m}_k^{n+1}(\omega) = \frac{\hat{h}(\omega) - \sum_{i \neq k} \hat{m}_i(\omega) + \frac{\hat{\alpha}(\omega)}{2}}{1 + 2\eta (\omega - \omega_k)^2}, \quad \text{(A8)}$$

Equation (A8) shows that $\hat{m}_k^{n+1}$ is the Wiener filtering of the current residual with the signal prior of $1/(\omega - \omega_k)^2$. $\hat{m}_k(\omega)$ can be transformed into mode $m_k(t)$ using inverse FT.

Similarly, ω_k^{n+1} is updated as follows:

$$\omega_k^{n+1} \leftarrow \underset{\omega_k}{\arg\min} L\left(\left\{ m_i^{n+1} \right\}, \left\{ \omega^{n+1} \right\}, \left\{ \omega_{i \geq k}^n \right\}, \alpha^n \right), \quad \text{(A9)}$$

The center frequencies ω_k appear only in the first term of Equation (A2). Thus, the relevant problem can be written as:

$$\omega_k^{n+1} = \underset{\omega_k}{\arg\min} \left\{ \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * m_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}, \quad \text{(A10)}$$

This optimization problem is transformed into the Fourier domain and eventually turns into the following form:

$$\omega_k^{n+1} = \underset{\omega_k}{\arg\min} \left\{ \int_0^\infty (\omega - \omega_k)^2 |\hat{m}_k(\omega)|^2 d\omega \right\}, \quad \text{(A11)}$$

The solution of this quadratic problem is:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{m}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{m}_k(\omega)|^2 d\omega}, \quad \text{(A12)}$$

It shows that the new ω_k is the center of gravity of the power spectrum of the recent mode.

Finally, the update of α^{n+1} is as follows:

$$\alpha^{n+1} = \alpha^n + \tau \left(h - \sum_{k=1}^{K} m_k^{n+1} \right) \tag{A13}$$

where τ is the update parameter of the Lagrange multiplier.

References

1. CIGRE Working Group. *Final Report of the Second International Enquiry on High Voltage Circuit Breaker Failures and Defects in Service*; CIGRE Report No. 83; CIGRE: Paris, France, 1994.
2. Runde, M.; Ottesen, G.E.; Skyberg, B.; Ohlen, M. Vibration analysis for diagnostic testing of circuit-breakers. *IEEE Trans. Power Deliv.* **1996**, *11*, 1816–1823. [CrossRef]
3. Lee, D.S.; Lithgow, B.; Morrison, R.E. New fault diagnosis of circuit breakers. *IEEE Trans. Power Deliv.* **2003**, *22*, 61–61. [CrossRef]
4. Landry, M.; Leonard, F.; Landry, C.; Beauchemin, R. An improved vibration analysis algorithm as a diagnostic tool for detecting mechanical anomalies on power circuit breakers. *IEEE Trans. Power Deliv.* **2008**, *23*, 1986–1994. [CrossRef]
5. Meng, Y.P.; Jia, S.L.; Shi, Z.Q.; Rong, M.Z. The detection of the closing moments of a vacuum circuit breaker by vibration analysis. *IEEE Trans. Power Deliv.* **2006**, *21*, 652–658. [CrossRef]
6. Huang, J.; Hu, X.G.; Yang, F. Support vector machine with genetic algorithm for machinery fault diagnosis of high voltage circuit breaker. *Measurement* **2011**, *44*, 1018–1027. [CrossRef]
7. Huang, J.; Hu, X.G.; Geng, X. An intelligent fault diagnosis method of high voltage circuit breaker based on improved EMD energy entropy and multi-class support vector machine. *Electr. Power Syst. Res.* **2011**, *81*, 400–407. [CrossRef]
8. Høidalen, H.K.; Runde, M. Continuous monitoring of circuit-breakers using vibration analysis. *IEEE Trans. Power Deliv.* **2005**, *20*, 2458–2465. [CrossRef]
9. Huang, N.; Fang, L.; Cai, G.; Xu, D.; Chen, H.; Nie, Y. Mechanical Fault Diagnosis of High Voltage Circuit Breakers with Unknown Fault Type Using Hybrid Classifier Based on LMD and Time Segmentation Energy Entropy. *Entropy* **2016**, *18*, 322. [CrossRef]
10. Sekiya, H.; Kimura, K.; Miki, C. Technique for determining bridge displacement response using MEMS accelerometers. *Sensors* **2016**, *16*, 257. [CrossRef] [PubMed]
11. Flett, J.; Bone, G.M. Fault detection and diagnosis of diesel engine valve trains. *Mech. Syst. Signal Process.* **2015**, *20*, 316–327. [CrossRef]
12. Harris, F.J. On the use of harmonic analysis with the discrete Fourier transform. *Proc. IEEE* **1978**, *66*, 51–83. [CrossRef]
13. Zhu, K.; Wong, Y.S.; Hong, G.S. Wavelet analysis of sensor signals for tool condition monitoring: A review and some new results. *Int. J. Mach. Tool. Manuf.* **2009**, *49*, 537–553. [CrossRef]
14. Peng, Z.K.; Jackson, M.R.; Rongong, J.A.; Chu, F.L.; Parkin, R.M. On the energy leakage of discrete wavelet transform. *Mech. Syst. Signal Process.* **2009**, *23*, 330–343. [CrossRef]
15. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A-Math. Phys. Eng.* **1998**, *454*, 903–995. [CrossRef]
16. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [CrossRef]
17. Yin, A.J.; Ren, H.J. A propagating mode extraction algorithm for microwave waveguide using variational mode decomposition. *Meas. Sci. Technol.* **2015**, *26*. [CrossRef]
18. Abdoos, A.A.; Mianaei, P.K.; Ghadikolaei, M.R. Combined VMD-SVM based feature selection method for classification of power quality events. *Appl. Soft Comput.* **2015**, *38*, 637–646. [CrossRef]
19. Upadhyay, A.; Pachori, R.B. Instantaneous voiced/non-voiced detection in speech signals based on variational mode decomposition. *J. Frankl. Inst.* **2015**, *352*, 2679–2707. [CrossRef]

20. Wang, Y.X.; Markert, R.; Xiang, J.W.; Zheng, W.G. Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system. *Mech. Syst. Signal Process.* **2015**, *60*, 243–251. [CrossRef]

21. Tayarani-Bathaie, S.S.; Khorasani, K. Fault detection and isolation of gas turbine engines using a bank of neural networks. *J. Process Control* **2015**, *36*, 22–41. [CrossRef]

22. Paliwal, M.; Kumar, U.A. Neural networks and statistical techniques: A review of applications. *Expert Syst. Appl.* **2009**, *36*, 2–17. [CrossRef]

23. Xian, G.M.; Zeng, B.Q. An intelligent fault diagnosis method based on wavelet packer analysis and hybrid support vector machines. *Expert Syst. Appl.* **2009**, *37*, 12131–12136. [CrossRef]

24. Schölkopf, B.; Platt, J.C.; Shawe-Taylor, J.; Smola, A.J.; Williamson, R.C. Estimating the support of a high-dimensional distribution. *Neural Comput.* **2001**, *13*, 1443–1471. [CrossRef] [PubMed]

25. Huang, N.T.; Chen, H.J.; Zhang, S.X.; Cai, G.W.; Li, W.G.; Xu, D.G.; Fang, L.H. Mechanical Fault Diagnosis of High Voltage Circuit Breakers Based on Wavelet Time-Frequency Entropy and One-Class Support Vector Machine. *Entropy* **2016**, *18*, 7. [CrossRef]

26. Mahadevan, S.; Shah, S.L. Fault detection and diagnosis in process data using one-class support vector machines. *J. Process Control* **2009**, *19*, 1627–1639. [CrossRef]

27. Fernández-Francos, D.; Martínez-Rego, D.; Fontenla-Romero, O.; Alonso-Betanzos, A. Automatic bearing fault diagnosis based on one-class *ν*-SVM. *Comput. Ind. Eng.* **2013**, *64*, 357–365. [CrossRef]

28. Yin, S.; Zhu, X.P.; Jing, C. Fault detection based on a robust one class support vector machine. *Neurocomputing* **2014**, *145*, 263–268. [CrossRef]

29. Bertsekas, D.P. Multiplier methods: A survey. *Automatica* **1976**, *12*, 133–145. [CrossRef]

30. Wu, Z.H.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2011**, *1*, 1–41. [CrossRef]

31. Torres, M.E.; Colominas, M.A.; Schlotthauer, G.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011.

32. Vapnik, V.N. *The Nature of Statistical Learning Theory*, 1st ed.; Springer: New York, NY, USA, 1995; pp. 127–140.

33. Klema, V.C.; Laub, A.J. The singular value decomposition: its computation and some applications. *IEEE Trans. Autom. Control* **1980**, *25*, 164–176. [CrossRef]

34. Tian, Y.; Tan, T.N.; Wang, Y.H.; Fang, Y.C. Do singular values contain adequate information for face recognition. *Pattern Recognit.* **2003**, *36*, 649–655. [CrossRef]

35. Tian, Y.; Ma, J.; Lu, C.; Wang, Z.L. Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine. *Mech. Mach. Theory* **2015**, *90*, 175–186. [CrossRef]