High-impedance Fault Section Location for Distribution Networks Based on t-distributed Stochastic Neighbor Embedding and Variable Mode Decomposition

Zhihua Yin, Yuping Zheng, Zhinong Wei, Guoqiang Sun, Sheng Chen, and Haixiang Zang

***Abstract*—When high-impedance faults (HIFs) occur in resonant grounded distribution networks, the current that flows is extremely weak, and the noise interference caused by the distribution network operation and the sampling error of the measurement devices further masks the fault characteristics. Consequently, locating a fault section with high sensitivity is difficult. Unlike existing technologies, this study presents a novel fault feature identification framework that addresses this issue. The framework includes three key steps: ① utilizing the variable mode decomposition (VMD) method to denoise the fault transient zero-sequence current (TZSC); ② employing a manifold learning algorithm based on t-distributed stochastic neighbor embedding (t-SNE) to further reduce the redundant information of the TZSC after denoising and to visualize fault information in high-dimensional 2D space; and ③ classifying the signal of each measurement point based on the fuzzy clustering method and combining the network topology structure to determine the fault section location. Numerical simulations and field testing confirm that the proposed method accurately detects the fault location, even under the influence of strong noise interference.**

***Index Terms*—High-impedance fault, noise interference, fault section location, t-distributed stochastic neighbor embedding (t-SNE), transient zero-sequence current.**

I. Introduction

AMONG all fault types of distribution networks, the probability of a single phase-to-ground (SPG) accounts for more than 70% [1], [2]. If they cannot be cleared in time, SPG faults may develop into short-circuit faults, which may cause greater harm to the safe operation of distribution networks [3]. Unfortunately, due to their weak fault characteristics, particularly when high-impedance faults (HIFs) occur, the relay protection devices cannot be activated promptly, making it difficult to detect these faults effectively and accurately for a long time [4], [5].

Currently, the fault detection technologies can be roughly categorized into two groups: methods based on steady-state signals and methods based on transient signals [6].

The fault detection methods based on steady-state signals are mainly applicable to isolated neutral distribution systems [7]. Because the neutral point is ungrounded, the fault current of the faulty feeder is approximately equal to the sum of the capacitive currents of the non-faulty feeders, making the fault characteristics of the faulty feeder relatively prominent [8], [9]. Reference [10] calculated the fault current and fault resistance by detecting the changes in the steady-state voltage and current before and after the fault, and determined whether to send a fault signal based on the threshold. Reference [11] calculated the zero-sequence admittance by measuring the zero-sequence voltage (ZSV) and zero-sequence current (ZSC) after a fault, which was utilized to construct the fault location criteria. However, for resonant grounded systems, the steady-state current of the faulty feeder no longer exhibits significant features because of the compensation effect of the arc-suppression coil on the ZSC, which renders all steady-state detection methods ineffective [12], [13].

The fault detection methods based on transient signals are not affected by the neutral grounding mode and contain rich fault information. Combined with appropriate signal analysis approaches, these methods have been widely favored by researchers [14], [15]. Reference [1] proposed a fundamental frequency component offset method to eliminate the steady-state components in transient ZSCs (TZSCs) to obtain a transient energy function. Then, they constructed a fault location criterion based on the cosine similarity. To determine the fault location, [16] compared the magnitudes of the peaks and occurrence times by measuring the local extreme values

Manuscript received: April 12, 2023; revised: July 19, 2023; accepted: October 7, 2023. Date of CrossChecked: October 7, 2023. Date of online publication: December 5, 2023.

This work was supported in part by the Science and Technology Program of State Grid Corporation of China (No. 5108-202218280A-2-75-XG), the Fundamental Research Funds for the Central Universities (No. B200203129), and the Postgraduate Research and Practice Innovation Program of Jiangsu Province (No. KYCX20_0432).

Z. Yin, Z. Wei (corresponding author), G. Sun, S. Chen, and H. Zang are with the College of Energy and Electrical Engineering, Hohai University, Nanjing 210098, China (e-mail: zhyincumt@126.com; wzn_nj@263.net; hhusunguoqiang@163.com; chenshenghhu@163.com; zanghaixiang@hhu. edu.cn).

Y. Zheng is with State Key Laboratory of Smart Grid Protection and Control, NARI Group Corporation, Nanjing 211106, China (e-mail: zhengyuping@sgepri.sgcc.com.cn).

DOI: 10.35833/MPCE.2023.000225

of the transient current. Reference [17] proposed a fault location technique based on transient zero-sequence admittance, which utilizes changes in transient zero-sequence admittance measured by feeder terminal devices upstream and downstream of the fault to locate the faulty segment. Reference [18] established an accurate zero-sequence model of distribution networks with variable frequency parameters based on the distribution characteristics of the TZSC, and they identified the fault location by monitoring the unique carrier frequency of each monitoring point along the feeder line.

A common issue encountered by these transient methods is that the fault location may fail because of the gradually increasing noise introduced by the measurement devices as the fault resistance increases, which eventually overwhelms the transient fault characteristic signal. Reference [19] indicated that sensors such as potential and current transformers can introduce measurement noise into the phase angle and current or voltage amplitude. By analyzing the sampled data of the distribution systems, [20] showed that measurement noise obeys the Gaussian distribution. Reference [21] systematically analyzed the mechanism of current sensor measurement noise generation, investigated the measurement accuracy level of advanced sensors in engineering, and provided a reliable basis for setting the measurement noise disturbance parameters.

To enhance the accuracy and reliability of transient methods, researchers have realized the significance of extracting and analyzing weak fault signals in the presence of background noise. Reference [2] proposed an HIF detection and location method based on feature selection, semi-supervised learning, and probability learning. In the verification phase, Gaussian white noise with a mean of 0 and a variance of 10^{-4} was added to the sampled data to simulate the effects of measurement errors. The results showed that the introduction of noise significantly deteriorated the fault localization performance of the algorithm. Reference [22] proposed a time-domain waveform distortion fault detection algorithm, which successfully extracted fault feature information under 10 dB noise interference by utilizing segmental linear fitting based on the least-squares method. However, the least-squares method is a linear estimator, and the actual fault propagation process is nonlinear. Reference [23] utilized a compressed sensing algorithm to estimate the fault location. This algorithm effectively detects noise signals within 1% by only requiring voltage sampling before and during faults on a small number of buses. This method exhibits good performance, but it requires a large number of voltage signals, which are difficult to obtain in practical distribution networks. Reference [24] utilized the Hilbert transform to compute the analytic signal of the TZSC, followed by the utilization of shift factors to remove the fundamental component. Finally, using a multi-objective evolutionary algorithm, they accurately analyzed the noise-induced fault signal, which enabled reliable HIF detection under 5 dB of noise interference. However, when the noise intensity increased to −5 dB, the method experienced false detections, indicating that the level of noise interference was a crucial limiting factor in determining the reliable detection range of fault diagnosis algorithms. Several drawbacks of the Hilbert transform algorithm such as mode mixing and end effects are potential factors that lead to inaccuracies.

To address this problem, this study proposes a fault detection method for HIF based on variable mode decomposition (VMD) and t-distributed stochastic neighbor embedding (t-SNE) with the aim of compensating for the insufficient sampling accuracy of current sensors using signal analysis methods. The main contributions include:

1) For constructing fault criteria, based on an analysis of the fault mechanism, we find that the capacitor current is not immediately compensated by the inductor current after an HIF occurs within a very short period. Therefore, a fault criterion is proposed based on the TZSC amplitude feature.

2) For signal preprocessing, the VMD method is proposed to decompose the noisy TZSC signals after the faults. Combined with the fault mechanism analysis, the optimal decomposition parameter is determined to be 3 to ensure the effective extraction of weak fault features under strong noise interference.

3) For fault feature extraction, to further reduce redundant information and better preserve HIF features, a manifold learning algorithm based on t-SNE and a fuzzy clustering method are used to visualize the denoised TZSC in 2D space and to classify the measurement points, respectively. Under the strong noise interference background of −3 dB, the precise fault section location of 3000 Ω fault resistance is achieved.

II. Transient Analysis of ZSC Circuit

Figure 1 shows a typical 10 kV resonant grounded distribution network with distributed generation (DG). T_1 is a 110 kV/10 kV main transformer, T_2 is a grounding transformer, and n feeders l_1-l_n radiate to supply power to load users. Because the winding type of the transformer connected to the DG is △/Y and its primary side is not grounded, ZSC can only form a loop in the △ winding [17], [25], [26]. Therefore, the connection of the DG does not change the zero-sequence fault network. The existing line selection and fault location methods that utilize transient zero-sequence signals can still be applicable in principle.

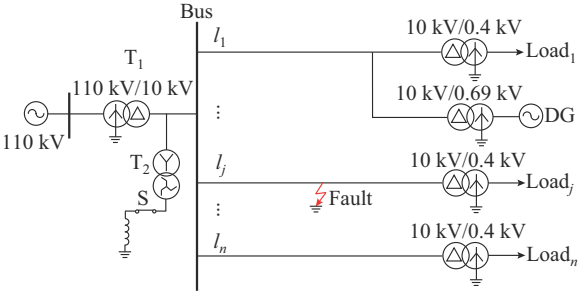

Fig. 1. Resonant grounding distribution network containing DG.

The distribution of ZSC in the entire network when HIF occurs on feeder j is shown in Fig. 2 [27], where $u_f = U_m \sin(\omega_0 t - \theta)$ is the virtual power source of the fault point; U_m is the amplitude of phase voltage; ω_0 is the fundamental

frequency, and θ is the initial phase angle of the fault; L_p is the zero-sequence equivalent inductance of the arc suppression coil; C_{0j} ($j=1,2,\ldots,n$) is the zero-sequence equivalent capacitance to ground of the feeder j; i_{0j} is the ZSC at the exit of the feeder j; $i_{C_{0j}}$ is the capacitance current of the feeder j to ground; i_L is the ZSC through L_p; u_0 is the zero-sequence voltage of the bus; and R is the equivalent resistance in the zero-sequence circuit, which is mainly composed of the feeder and fault resistances of the grounding point through which the grounding current flows. Under HIF conditions, R can be approximated to three times the fault resistance R_f (i.e., $R \approx 3R_f$).

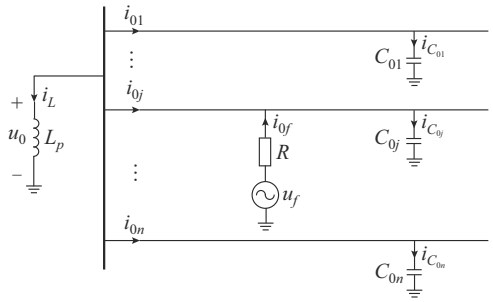

Fig. 2. Distribution of ZSC in entire network with HIF.

Because the initial stored energy of the dynamic component is zero, the following differential equation can be obtained from Fig. 2.

$$\begin{cases} RL_pC_{0\Sigma}\dfrac{\mathrm{d}^2 i_L}{\mathrm{d}t^2} + L_p\dfrac{\mathrm{d}i_L}{\mathrm{d}t} + Ri_L = U_m\sin(\omega_0 t-\theta) \\ i_L(0_+)=0 \\ \left.\dfrac{\mathrm{d}i_L}{\mathrm{d}t}\right|_{0+}=0 \end{cases} \tag{1}$$

where $C_{0\Sigma}=C_{0j}$ is the zero-sequence capacitance to ground of the entire distribution network.

For a practical HIF, R generally satisfies:

$$R > \frac{1}{2}\sqrt{\frac{L_p}{C_{0\Sigma}}} \tag{2}$$

In this case, the equivalent circuit corresponding to Fig. 2 is in an underdamped state. Solving (1) yields:

$$i_L = \frac{U_m}{\sqrt{\omega_0^2L_p^2+R^2(1-\omega_0^2L_pC_{0\Sigma})^2}}\left\{\sin(\omega_0 t-\beta)+\left[\sin\beta\cos\omega t+(\delta\sin\beta-\omega_0\cos\beta)\frac{\sin\omega t}{\omega}\right]\mathrm{e}^{-\delta t}\right\} \tag{3}$$

$$\begin{cases} \delta = \dfrac{1}{2RC_{0\Sigma}} \\ \omega^2 = \dfrac{1}{L_pC_{0\Sigma}} - \dfrac{1}{(2RC_{0\Sigma})^2} \end{cases} \tag{4}$$

$$\beta = \theta + \arctan\frac{\omega_0 L_p}{R(1-\omega_0^2L_pC_{0\Sigma})} \tag{5}$$

$i_{C_{0j}}$ can be obtained by:

$$i_{C_{0j}} = C_{0j}\frac{\mathrm{d}u_0}{\mathrm{d}t} = C_{0j}L_p\frac{\mathrm{d}^2 i_L}{\mathrm{d}t^2} = \frac{U_mC_{0j}L_p}{\sqrt{\omega_0^2L_p^2+R^2(1-\omega_0^2L_pC_{0\Sigma})^2}}\left\{-\omega_0^2\sin(\omega_0 t-\beta)+[2\delta\omega_0\cos\beta-(\delta^2+\omega^2)\sin\beta]\mathrm{e}^{-\delta t}\cos\omega t+\left[2\delta\omega\sin\beta+(\delta^2-\omega^2)\frac{\delta\sin\beta-\omega_0\cos\beta}{\omega}\right]\mathrm{e}^{-\delta t}\sin\omega t\right\} \tag{6}$$

Thus, i_{0j} can be obtained by:

$$i_{0j} = -i_L - \sum_{k\neq j}^{n} i_{C_{0k}} = \frac{U_m}{\sqrt{\omega_0^2L_p^2+R^2(1-\omega_0^2L_pC_{0\Sigma})^2}}\cdot\left\{[(C_{0\Sigma}-C_{0j})L_p\omega_0^2-1]\sin(\omega_0 t-\beta)-(\mathrm{e}^{-\delta t}\cos\omega t)\{(C_{0\Sigma}-C_{0j})L_p[2\delta\omega_0\cos\beta-(\delta^2+\omega^2)\sin\beta)]+\sin\beta\}-\mathrm{e}^{-\delta t}\sin\omega t\left\{(C_{0\Sigma}-C_{0j})L_p\left[2\delta\omega\sin\beta+(\delta^2-\omega^2)\cdot\frac{\delta\sin\beta-\omega_0\cos\beta}{\omega}\right]+\frac{\delta\sin\beta-\omega_0\cos\beta}{\omega}\right\}\right\} \tag{7}$$

When $t\to+\infty$, it holds that

$$i_{C_{0k}} = -\frac{U_m}{\sqrt{\omega_0^2L_p^2+R^2(1-\omega_0^2L_pC_{0\Sigma})^2}}C_{0k}L_p\omega_0^2\sin(\omega_0 t-\beta) \tag{8}$$

$$i_{0j} = \frac{U_m}{\sqrt{\omega_0^2L_p^2+R^2(1-\omega_0^2L_pC_{0\Sigma})^2}}[(C_{0\Sigma}-C_{0j})L_p\omega_0^2-1]\sin(\omega_0 t-\beta) \tag{9}$$

Without consideration of the overcompensation effect of the arc-suppression coil during the transient period, it can be assumed that:

$$\omega_0 L_p \approx \frac{1}{\omega_0 C_{0\Sigma}} \tag{10}$$

Combining (8)-(10), we can obtain:

$$\frac{i_{0j}}{i_{C_{0k}}} = \frac{(C_{0\Sigma}-C_{0j})L_p\omega_0^2-1}{-C_{0k}L_p\omega_0^2} \approx \frac{C_{0j}}{C_{0k}} > 0 \tag{11}$$

This implies that after a sufficiently long period of oscillation decay, the oscillation direction of the ZSC at the faulty feeder terminal tends to be the same as that at the non-faulty feeder terminal, and the ratio of their amplitudes depends on the zero-sequence capacitance ratio of the respective feeders. At this point, neither the oscillation direction nor the amplitude of the ZSC at each feeder can be used for fault discrimination.

Considering the case of $t\to 0_+$, we have:

$$i_{0j}(0_+) = -(C_{0\Sigma}-C_{0k})\frac{U_mL_p[2\delta\omega_0\cos\beta-(\delta^2+\omega^2-\omega_0^2)\sin\beta)]}{\sqrt{\omega_0^2L_p^2+R^2(1-\omega_0^2L_pC_{0\Sigma})^2}} = -\sum_{k\neq j}^{n} i_{C_{0k}}(0_+) \tag{12}$$

In fact, (12) can be not only obtained from (6) and (7), but also explained by the principle that the energy of the inductive element cannot undergo a sudden change. Because the electric current is a representation of the energy stored in

the inductive components, we have:

$$i_L(0_+)=i_L(0_-)\approx 0 \tag{13}$$

In some special cases such as when impulse excitation occurs in the circuit, (13) may not hold. However, for a practical distribution network, there is no need to consider such special cases. Therefore, we can assume that (13) always holds.

For the fault point shown in Fig. 2, according to Kirchhoff's current law, we have:

$$i_{0j}(0_+)=-i_L(0_+)-\sum_{k\neq j}^{n} i_{C_{0k}}(0_+)=-\sum_{k\neq j}^{n} i_{C_{0k}}(0_+) \tag{14}$$

Clearly, the conclusions of (12) and (14) are identical, both of which prove that at the transient moment of the HIFs, i.e., at time $t(0_+)$, the ZSC flowing out of the faulty feeder is the sum of all non-faulty feeder-to-ground capacitive currents, but with opposite directions. As the transient process progresses, the inductive current flowing through the arc-suppression coil gradually increases, slowly compensating for the capacitive current flowing through the fault point, which is manifested as the ZSC at the faulty feeder exits as it slowly decays from its maximum amplitude. This suggests that this transient process can be used to construct fault criteria, and that the decay factor δ determines the duration. Based on the actual distribution network feeder parameters, this paper sets that this conclusion holds true during the period of [0, T/2], where $T=0.02$ s represents the power system period. The physical essence of this process is that the feeder-to-ground capacitance of the entire system releases the stored energy, and the faulty feeder forcibly absorbs this part of the energy.

The aforementioned theoretical analysis and conclusions indicate that it is possible to directly utilize the TZSC for criterion construction, without the need to exclude the fundamental current components in (6) and (7) nor to obtain voltage signals from various measurement points. This method has obvious advantages over the existing ones [1].

III. Noise Reduction Method Based on VMD

A. Noise Reduction Method Based on VMD for Raw Signals

The VMD method demonstrates the excellent performance in noise reduction and processing of nonstationary signals [28], [29]. Its essence lies in transforming the signal decomposition problem into a constrained optimization problem, which can be divided into two parts: the construction and solution of the variational problem [30].

The VMD method constructs a variational problem as expressed by:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left\{ \sum\limits_k \left\| \partial_t \left[\left(\delta(t)+\dfrac{\mathrm{j}}{\pi t} \right) u_k(t) \right] \mathrm{e}^{-\mathrm{j}\omega_k t} \right\|_2^2 \right\} \\ \text{s.t.} \ \sum\limits_k u_k = f \end{cases} \tag{15}$$

where $\{u_k\}=\{u_1,u_2,\ldots,u_K\}$ represents the various mode components; $\{\omega_k\}=\{\omega_1,\omega_2,\ldots,\omega_K\}$ represents the central frequency of each mode component; $\delta(t)$ is the impulse function; ∂_t is the partial derivative with respect to t; and f is the original signal.

Introducing the secondary penalty factor α and Lagrange multiplier operator $\lambda(t)$, (15) is transformed into an unconstrained variational problem, where the secondary penalty factor ensures the reconstruction accuracy of the signal in the presence of Gaussian noise. The extended Lagrange expression is:

$$\begin{aligned} L(\{u_k\},\{\omega_k\},\lambda) := \alpha \sum_k \left\| \partial_t \left[\left(\delta(t)+\frac{\mathrm{j}}{\pi t} \right) u_k(t) \right] \mathrm{e}^{-\mathrm{j}\omega_k t} \right\|_2^2 + \\ \left\| f(t)-\sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t)-\sum_k u_k(t) \right\rangle \end{aligned} \tag{16}$$

The expression for each mode component and its central frequency can be obtained by utilizing the alternating direction method of multipliers to solve the extended Lagrange expression.

$$\hat{u}_k^{n+1}(\omega)=\frac{\hat{f}(\omega)-\sum\limits_{i\neq k}\hat{u}_i(\omega)+\dfrac{\hat{\lambda}(\omega)}{2}}{1+2\alpha(\omega-\omega_k)^2} \tag{17}$$

$$\omega_k^{n+1}=\frac{\int_0^\infty \omega \left|\hat{u}_k(\omega)\right|^2 \mathrm{d}\omega}{\int_0^\infty \left|\hat{u}_k(\omega)\right|^2 \mathrm{d}\omega} \tag{18}$$

When (17) and (18) are iteratively computed until the convergence is achieved, the Fourier inverse transform can be applied to $\hat{u}_k(w)$, and its real part is used to obtain an analytical solution for $u_k(t)$.

$$u_k(t)=\mathrm{Re}(F^{-1}(\hat{u}_k(w))) \tag{19}$$

B. Determination of Parameter K

The setting for parameter K directly affects the effect of VMD, and determining the optimal K value has always been a challenge when using the VMD method [31]. Most existing studies choose the parameter based on experience, which lacks persuasiveness. In the research field of this study, it is necessary to further examine (6) and (7) and simplify them as the following form:

$$i_{C_{0j}}(t)=A_1\sin(\omega_0 t-\beta)+A_2\mathrm{e}^{-\delta t}\cos\omega t+A_3\mathrm{e}^{-\delta t}\sin\omega t \tag{20}$$

$$i_{0j}(t)=A_4\sin(\omega_0 t-\beta)+A_5\mathrm{e}^{-\delta t}\cos\omega t+A_6\mathrm{e}^{-\delta t}\sin\omega t \tag{21}$$

where A_1-A_6 are the system-related constants of the distribution network.

According to (20) and (21), the central frequency of the time-domain waveform of ZSC for each feeder is significantly correlated with ω and ω_0. Combined with (4), we can infer that the attenuation factor δ during HIF is generally small, resulting in a longer transient attenuation period and a smoother attenuation process. Therefore, we can assume that the transient transition process of the HIF corresponds to two central frequencies ω and ω_0. In addition, when the high-frequency noise introduced by the current sensors during the sampling process is considered, an additional higher

central frequency can be introduced, and $K=3$ can be set to decompose the noisy ZSC. Note that for most HIFs in practical engineering, $\omega \approx \omega_0$ is approximately satisfied, which means that the main energy of each feeder is concentrated near the fundamental frequency. Therefore, in practical applications, the modal component corresponding to ω_0 can be utilized to reconstruct the denoised fault features, that is, the intrinsic mode function (IMF) associated with $u_1(t)$.

Figure 3(a) shows the original signal of typical ZSC of the HIF, whereas Fig. 3(b) shows the ZSC affected by noise interference, indicating that the characteristic pattern of the original signal is no longer obvious due to the influence of noise. The polluted signal is decomposed by VMD, and the decomposition result when $K=3$ is shown in Fig. 4. The figure shows that IMF1 retains the main wave pattern of the original signal, which is reflected in the similarity of the wave period and peak value to the original signal. IMF2 and IMF3 mainly represent high-frequency noise interference, which is consistent with the theoretical analysis results and also confirms the effective feature extraction ability of VMD for weak fault signals in noisy backgrounds.

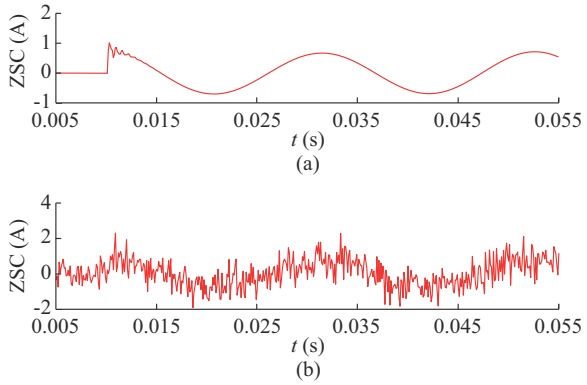

Fig. 3. Typical ZSC of HIF. (a) Original signal. (b) Polluted signal with noise interference.

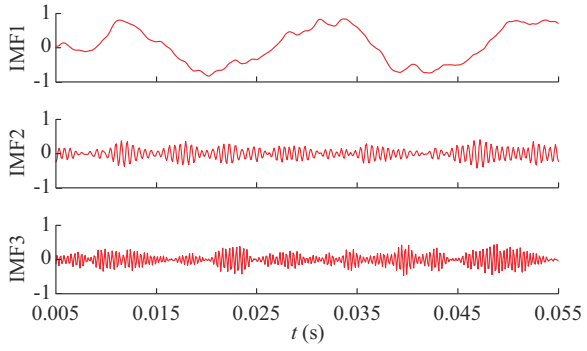

Fig. 4. Polluted signal decomposed by VMD.

IV. Fault Section Location Method Based on t-SNE

A. Manifold Learning Algorithm Based on t-SNE

According to the fault mechanism analysis in Section II, the transient duration of HIF is related to the attenuation factor δ, which can vary within a certain range due to the influence of fault conditions and system parameters. Because this paper focuses on the TZSC within $T/2$ after the fault, there will be corresponding levels of redundant data under specific fault conditions. Therefore, it is crucial to eliminate the redundant features within these data and simultaneously reduce the dimensionality of the high-dimensional fault dataset in order to extract valuable fault state information that reflects the HIF. This reduces the difficulty of fault classification and improves the accuracy of fault identification.

t-SNE is a nonlinear manifold deep learning algorithm that can effectively achieve the visualization and dimensionality reduction of high-dimensional data [32], [33]. The main steps are as follows.

For a high-dimensional dataset $\boldsymbol{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_N]\in \mathrm{R}^D$, the probability distribution $p_{j|i}$ is defined to represent the conditional probability that $\boldsymbol{x}_i$ will choose $\boldsymbol{x}_j$ as its neighbor:

$$p_{j|i}=\frac{\mathrm{e}^{-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_j\|^2}{2\sigma_i^2}}}{\sum_{k\neq i}\mathrm{e}^{-\frac{\|\boldsymbol{x}_i-\boldsymbol{x}_k\|^2}{2\sigma_i^2}}} \tag{22}$$

where σ_i is the Gaussian variance centered at $\boldsymbol{x}_i$, which is obtained based on the specified perplexity.

The perplexity is related to Shannon's information entropy, and its expression is given by:

$$Perp(P_i)=2^{-\sum_{j=1}^{N}p_{j|i}\log_2 p_{j|i}} \tag{23}$$

As $p_{j|i}\neq p_{i|j}$ holds here, a joint probability distribution satisfying the symmetry property is defined as:

$$p_{ij}=\frac{p_{j|i}+p_{i|j}}{2N} \tag{24}$$

To map the high-dimensional dataset $\boldsymbol{X}$ to a low-dimensional space, let $\boldsymbol{Y}=[\boldsymbol{y}_1,\boldsymbol{y}_2,\ldots,\boldsymbol{y}_N]\in \mathrm{R}^S$ be the initial coordinate set in the low-dimensional space, where S should satisfy $D>>S$. Here, the t-distribution is used to model the medium distances in the high-dimensional space as long distances in the low-dimensional space. As a result, the joint probability distribution of $\boldsymbol{Y}$ modeled in the low-dimensional space follows the Cauchy distribution and is given by:

$$q_{ij}=\frac{(1+\|\boldsymbol{y}_i-\boldsymbol{y}_j\|^2)^{-1}}{\sum_{k\neq l}(1+\|\boldsymbol{y}_k-\boldsymbol{y}_l\|^2)^{-1}} \tag{25}$$

To minimize the difference in probability distribution between the datasets $\boldsymbol{X}$ and $\boldsymbol{Y}$, a loss function $C(y)$ is constructed using the Kullback-Leibler divergence as:

$$C(y)=KL(P\|Q)=\sum_i\sum_j p_{ij}\log_2\frac{p_{ij}}{q_{ij}} \tag{26}$$

To obtain the target dataset $\boldsymbol{Y}$, $C(y)$ is minimized using gradient descent starting from a random initialization.

$$\frac{\partial C}{\partial \boldsymbol{y}_i}=4\sum_j(p_{ij}-q_{ij})(\boldsymbol{y}_i-\boldsymbol{y}_j)(1+\|\boldsymbol{y}_i-\boldsymbol{y}_j\|^2)^{-1} \tag{27}$$

The corresponding iterative update process is:

$$\boldsymbol{y}_i^r=\boldsymbol{y}_i^{r-1}+\xi\frac{\partial C}{\partial \boldsymbol{y}_i}+\eta(r)(\boldsymbol{y}_i^{r-1}-\boldsymbol{y}_i^{r-2}) \tag{28}$$

where ξ represents the learning rate; r is the iteration step; and $\eta(t)$ is the iteration coefficient.

B. Determination of Fault Section Location

Thus, we can obtain N S-dimensional vectors $[\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N]$ that correspond to N actual measurement points. To determine the fault section location, these N measurement points must be further divided into two categories: fault paths and non-fault paths. In this study, the FCM algorithm is used as a reference to construct the objective function:

$$\begin{cases} J(\boldsymbol{G}, \boldsymbol{c}_1, \boldsymbol{c}_2) = \sum_{i=1}^{2}\sum_{j=1}^{N} g_{ij}^{h} \|\boldsymbol{c}_i - \boldsymbol{y}_j\|^2 \\ \text{s.t.} \ \sum_{i=1}^{2} g_{ij} = 1 \end{cases} \tag{29}$$

where g_{ij} is the element of membership matrix $\boldsymbol{G}$; and h is the weighted index. The optimal solution to (29) can be obtained using the Lagrange multiplier method.

$$g_{ij} = \frac{1}{\sum_{k=1}^{2}\left(\frac{\|\boldsymbol{y}_j - \boldsymbol{c}_i\|}{\|\boldsymbol{y}_j - \boldsymbol{c}_k\|}\right)^{\frac{2}{h-1}}} \tag{30}$$

$$\boldsymbol{c}_i = \frac{\sum_{j=1}^{N} g_{ij}^{h} \boldsymbol{y}_j}{\sum_{j=1}^{n} g_{ij}^{h}} \tag{31}$$

When the difference between the new and old membership matrices is less than the set threshold value, the final classification result $\{\boldsymbol{c}_1, \boldsymbol{c}_2\}$ can be obtained, and the actual fault section can be determined by combining it with the structure of the distribution network.

C. Flow of Detection Method

To provide a clear illustration of the entire detection process, a flow chart of the HIF section location algorithm is presented in Fig. 5. This algorithm is triggered when zero-sequence voltage U_0 exceeds 0.15 times the phase voltage U_N. The denoising, fault detection, and output judgment modules are then executed in sequence, ultimately producing the fault section location result .

V. Simulation and Field Test Verification

A. Simulation Model

As shown in Fig. 6, this paper establishes a radial distribution network model in PSCAD/EMTDC. The transformer ratio is 110 kV/10 kV with a rated capacity of 40000 kVA. The equivalent inductance of the arc-extinguishing coil is $L_p = 0.71$ H. The system is a hybrid cable-overhead feeder structure that includes cable feeder model YJV22-3*400, overhead main feeder model JKLYJ-240, and overhead branch feeder model JKLYJ-150 with total lengths of 6700, 4350, and 3000 m, respectively. The wind power generation systems DG_1 and DG_2 are connected to the distribution network through 0.69 kV/10 kV step-up transformers.

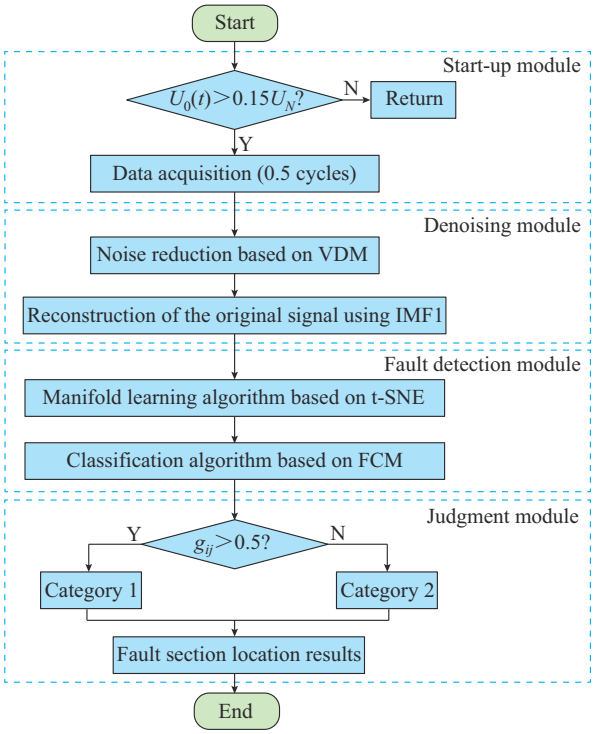

Fig. 5. Flow chart of HIF section location algorithm.

A 1 MW constant impedance load is used for all feeders. Thirteen groups of ZSC transformers are installed to obtain the ZSC at each measurement point (A-M) with a sampling frequency of 10 kHz, which is equal to 200 samples per cycle.

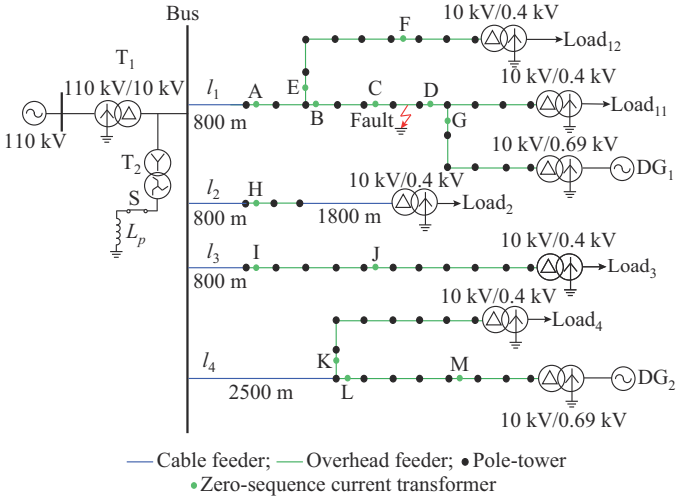

Fig. 6. Radial distribution network model established in PSCAD/EMTDC.

B. Fault Section Location Analysis

Based on the assumption that an SPG fault occurs in the section C-D at $t_f = 0.01$ s with the corresponding initial fault angle of $\theta = 90°$ and fault resistance of $R_f = 3000\ \Omega$, the system is in an under-damped state. Figure 7 shows the time-domain waveform of the faulty ZSCs under this fault condition, where measurement points A, D, and L are at the terminal of the faulty feeder, the downstream of the fault point, and the non-faulty feeder l_4, respectively. The following conclusions can be drawn.

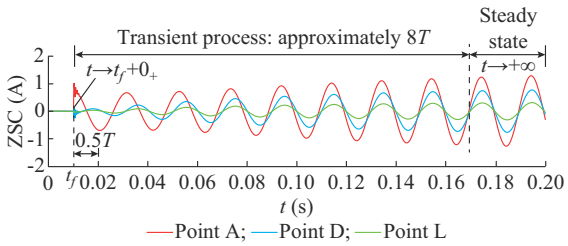

Fig. 7. Time-domain waveform of faulty ZSCs with R_f = 3000 Ω (original signals without considering noise interference).

1) When $t \to t_f + 0_+$, the value of ZSC at point A is significantly greater than those at points D and L, which is consistent with the mechanism analytical results derived from (12)-(14). At this time, the ZSCs at each measurement point are all capacitive.

2) As the fault progresses, the inductive current flowing through the arc-suppression coil begins to increase slowly. However, due to the relatively high value of R_f, the growth rate is relatively slow. This leads to the conclusion that within $T/2$ after the fault occurs, the ZSC at point A is assumed to be significantly greater than those at points D and L.

3) When $t_f + T/2 < t < t_f + 8T$, the system is still in a transient process, during which the ZSC at point A does not have a unchanged relationship with those at points D and L in terms of the wave amplitude and phase angle. This transient process also compensates for the arc-suppression coil inductive current of the system capacitive current.

4) When $t \to +\infty$, i.e., $t \geq t_f + 8T$, the system enters a steady state. During this period, the phases of the ZSCs at points A, D, and L are the same, and the relationship between the amplitudes is unchanged. This is consistent with the analysis and conclusions of the mechanisms derived from (8)-(11). During this period, the measured ZSCs cannot be used to construct an effective location criterion.

In practical distribution networks, ZSCs are not as clear as those shown in Fig. 7 because the accuracy of the measurement cannot be infinitely high, and the noise interference inevitably exists. In this paper, Gaussian white noise is used to simulate measurement error interference, and the signal-to-noise ratio (SNR) is used to evaluate noise intensity for assessing the immunity of the proposed algorithm against measurement noise.

Figure 8 shows the polluted signals with Gaussian white noise at an SNR of −3 dB to the original signals shown in Fig. 7. It can be observed that the original fault characteristics are overwhelmed by noise and are no longer significant. Therefore, this study proposes preprocessing these polluted signals using the VMD method.

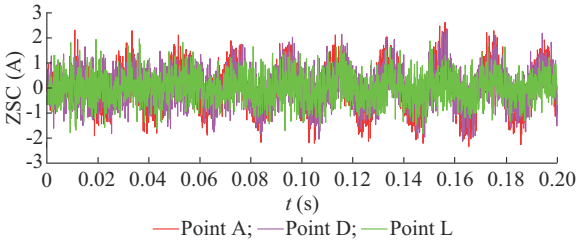

Fig. 8. Polluted signals with Gaussian white noise at an SNR of −3 dB.

Figure 9 shows the original signals, the polluted signals obtained through actual sampling, and the signals reconstructed using the VMD method (IMF1) of measurement points A, D, and L. It reveals that despite the strong randomness exhibited by the polluted signal, after being processed using the VMD method, IMF1 demonstrates a high degree of fitting with the original signal. This suggests that IMF1 can be directly used as the reconstructed signal and further validates the rationality of the proposed VMD method in Section III.

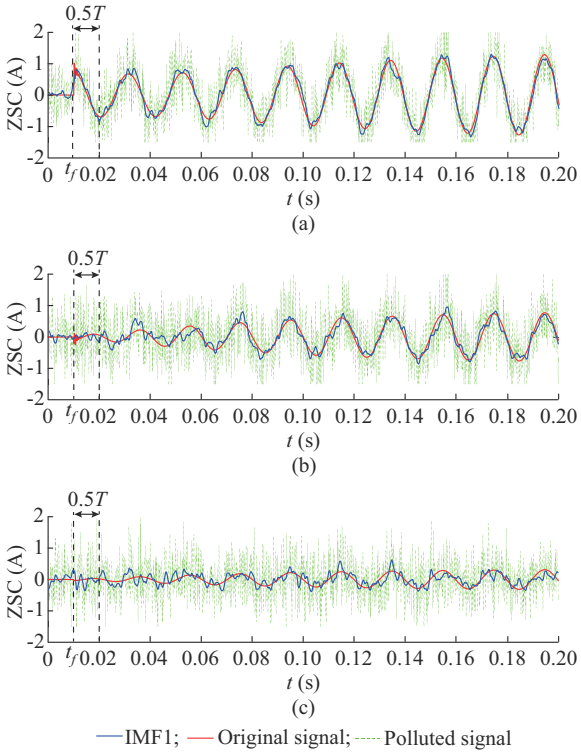

Fig. 9. Comparison among original signals, polluted signals, and reconstructed results. (a) Point A. (b) Point D. (c) Point L.

After the VMD denoising process is completed for all 13 ZSCs shown in Fig. 6, the IMF1 data from the $T/2$ time range shown in Fig. 9 are extracted as input for the t-SNE algorithm. Figure 10(a) shows the t-SNE visualization results after reducing these 13 high-dimensional vectors to two dimensions. Points A, B, and C indicate that the ZSC flows at the faulty feeder terminal are closely clustered. In addition, the other measurement points that do not cover the fault path are also closely clustered, forming two distinct clusters that are significantly distinguished by the FCM algorithm, as indicated by the green dashed line in Fig. 10.

In contrast, Fig. 10(b) shows the t-SNE visualization results without the VMD denoising step. It can be observed that the closeness of the two distinct clusters in this case is not as evident as Fig. 10(a). Moreover, the measurement point K is mistakenly grouped with the fault path measurement points. This highlights the necessity of using the VMD method to preprocess data.

Considering the actual network topology, the fault section location can be determined as the section C-D, as shown in Fig. 11.

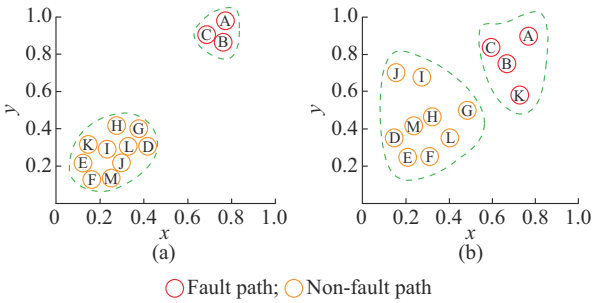

Fig. 10. t-SNE visualization results. (a) Under VMD denoising step. (b) Without VMD denoising step.

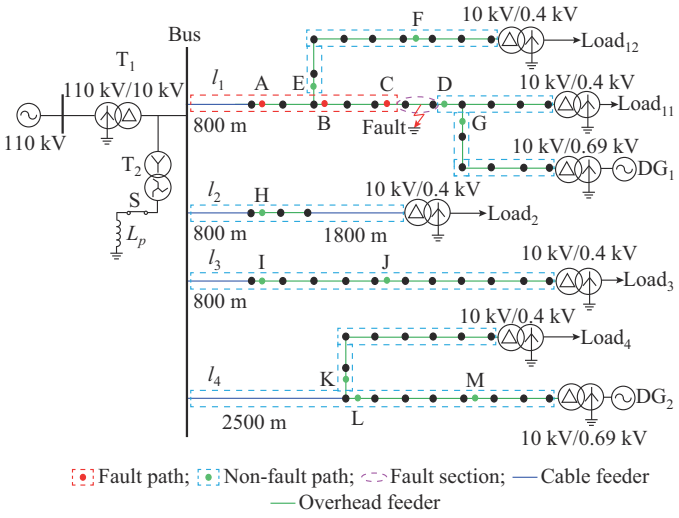

Fig. 11. Fault section location results of HIF.

Table I lists the fault section location results under different fault conditions.

TABLE I
FAULT SECTION LOCATION RESULTS

Fault section	Fault condition		SNR (dB)	Result
	θ (°)	R_f (Ω)		
B-C	0	500	30	B-C (√)
B-C	0	1000	20	B-C (√)
B-C	45	1500	10	B-C (√)
L-M	45	2000	5	L-M (√)
I-J	90	2500	0	I-J (√)
C-D	90	3000	−3	C-D (√)
E-F	0	3000	−10	Error (×)
C-D	0	2000	−10	Error (×)
C-D	90	4000	−6	Error (×)

Note: the symbols "√" and "×" represent that the results are correct and incorrect, respectively.

From Table I, we can conclude the following:

1) The proposed method can achieve accurate fault section location for all fault conditions within a −3 dB SNR, with a maximum transition resistance value of 3000 Ω.

2) When the SNR is too high such as −6 dB or −10 dB, it results in erroneous fault section location, indicating that the noise interference level is a major factor affecting the accuracy of the location results.

3) It should be noted that the actual sampling accuracy of the current sensors is limited, within a measurement error of approximately ±1 A [21]. This results in lower SNR values at higher transition resistances, leading to errors in the fault location, which highlights the importance of researchers paying close attention to the immunity of algorithms to noise interference.

C. Reliability Testing

During the verification phase, it is necessary to consider the influence of impulsive or non-normally distributed noise on the reliability of the algorithm. Based on the case shown in Fig. 9, Fig. 12(a) presents the VMD results simultaneously under five random impulsive noise pulses ①-⑤ and persistent Gaussian distributed noise interference. Compared with the case of only Gaussian noise interference, the impulsive noise has a certain impact on the VMD results, as clearly shown by the dashed circle in this figure.

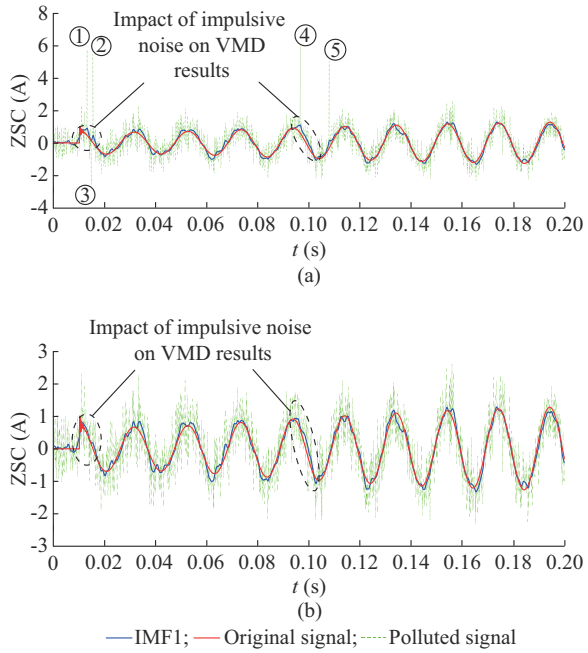

Fig. 12. Analysis of VMD results. (a) Under five random impulsive noise pulses and Gaussian distributed noise interference. (b) Under only Gaussian noise interference.

In addition, in practical applications, the sampling frequency has a significant impact on the data volume for faulty signal transmission. Therefore, we must also consider this possibility. Table II presents the algorithm testing results for different sampling frequencies and noise interferences under different probability distributions when an HIF at R_f = 3000 Ω occurs in fault section C-D. A small amount of random impulsive noises is added to the noise under different probability distributions, but the total SNR remains at −3 dB.

It can be observed that the proposed algorithm has a lower limit requirement for the sampling frequency, whereas higher sampling rates do not affect the reliability of the algo-

rithm. In terms of reducing faulty data transmission, the lower the sampling frequency, the higher the engineering practicality of the algorithm. According to test results, the actual sampling frequency should not be lower than 3 kHz. Regarding the different probability distributions of noise, although they may have a certain effect on the VMD results, accurate results can still be obtained after a series of processes using the proposed t-SNE is conducted. This indicates that when the SNR is constant, different probability distributions of noise do not significantly affect the computational results of the algorithm.

TABLE II
FAULT SECTION LOCATION RESULTS (R_f = 3000 Ω, SNR = −3 dB)

Sampling frequency (kHz)	Probability distribution of noise		
	Exponential distribution	Uniform distribution	Gaussian distribution
0.5	×	×	×
3.0	√	√	√
5.0	√	√	√
8.0	√	√	√
10.0	√	√	√
20.0	√	√	√
100.0	√	√	√

Considering the aforementioned analysis, the proposed method exhibits satisfactory performance in terms of its applicability in the sampling frequency range and its immunity to noise under different probability distributions.

D. Comparison with Existing Methods

A comprehensive comparison is presented in Table III showing the technical schemes of various HIF location methods and their corresponding anti-noise ability test results. All previous studies have acknowledged the significance of anti-noise analysis, as noise interference significantly affects the accuracy of fault location methods. Of all the technical schemes shown in the comparison, the proposed method demonstrates exceptional performance by being immune to noise interference with an intensity of as much as −3 dB.

TABLE III
COMPARISON OF VARIOUS HIF LOCATION METHODS

Reference	Technical scheme	Noise (dB)
This paper	VMD denoising followed by t-SNE manifold learning	30, 20, 10, 5, 0, and −3
[21]	Adaptive FDM and density-distance based FCM	32.76, 1.71, −0.21, and −0.32
[22]	Piecewise linear fitting based on least square	10, 8, and 5
[23]	Compressed sensing algorithm	60, 40, 35, and 30
[34]	Inner product transformation of extreme values	5
[35]	Optimized bistable system	−1

E. Field Data Test

To evaluate the adaptability of the proposed method in practical applications, a manual SGP test is conducted on a 10 kV urban distribution network in Wuxi, China. As shown in Fig. 13, the substation has 14 feeders, and the test is conducted on two of them (F1 and F2). The cable feeder lengths are indicated on the diagram according to engineering conventions. The overhead feeder lengths are converted based on the number of pole-towers. The average pole-tower spanning in this illustration is 50 m. In the test, the system capacitive current is 76.25 A, the working current is 81.60 A, and the detuning degree is −7.02%. Several leading domestic manufacturers of fault indicators participate in this test. F1 and F2 correspond to the two types of fault indicators under different models and measurement accuracies.

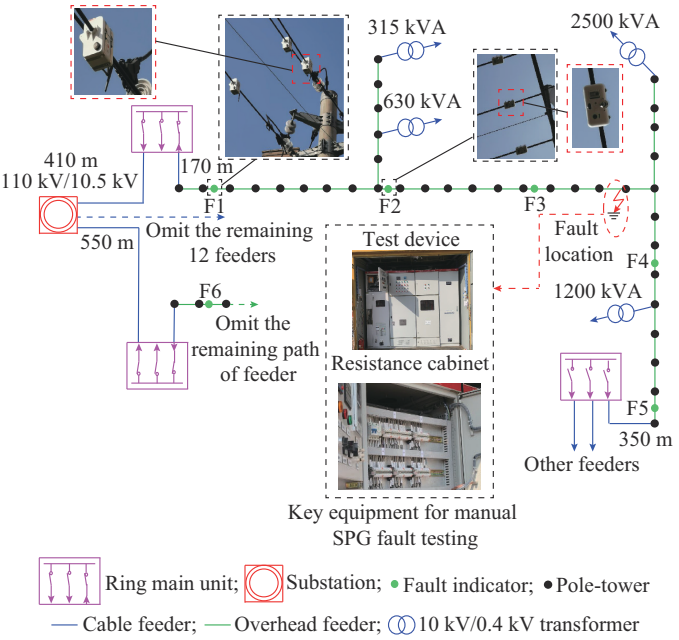

Fig. 13. Structure of distribution network for field tests.

Figure 14 shows the original recorded data of ZSC and its corresponding IMF1 obtained by VMD for points F1 and F2, respectively. The fault occurs at the 1000th sampling point under a sampling frequency of 10 kHz and a fault resistance value of 2000 Ω, indicating a typical HIF.

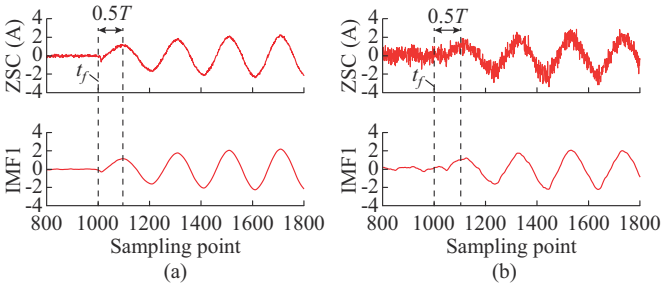

Fig. 14. Original recorded signal and its corresponding IMF1 obtained by VMD. (a) Point F1. (b) Point F2.

Clearly, the sampling accuracy of F1 is significantly higher than that of F2. Moreover, because both F1 and F2 be-

long to the same upstream fault point, their ZSC waveforms should be similar. Therefore, F1 could be regarded as the original fault signal and F2 could be considered as a noisy signal polluted by measurement errors. After the proposed method is applied to both signals and their respective IMF1 values are obtained, the decomposition results are found to be very similar to the original fault signal of F1, thus demonstrating the effectiveness of the proposed method.

Table IV presents the judgment results of the participating institutions 1-3 and the proposed method under different fault test conditions. It can be observed that all participating institutions can achieve reliable diagnosis when R_f is not greater than 500 Ω. However, when R_f increases to 2000 Ω, only the proposed method can achieve accurate diagnosis, indicating that it has greater reliability and can be well adapted to engineering practice.

TABLE IV
JUDGMENT RESULTS OF PARTICIPATING INSTITUTIONS 1-3 AND PROPOSED METHOD UNDER DIFFERENT FAULT TEST CONDITIONS

Method/ institution	Judgment result				Accuracy rate (%)
	200 Ω	500 Ω	1000 Ω	2000 Ω	
Proposed	√	√	√	√	100
Institution 1	√	√	√	×	75
Institution 2	√	√	√	×	75
Institution 3	√	√	×	×	50

VI. CONCLUSION

This study proposes a novel method for identifying noisy HIF features. An analysis of the fault mechanism reveals that a significant difference exists in the TZSC amplitude between faulty and non-faulty feeders shortly after fault occurrence.

To improve the signal processing, we utilize a VMD method to preprocess the sampled signal and a t-SNE method to remove redundant signals during $[t_f, t_f + T/2]$. The proposed method effectively localizes HIF sections under the operating conditions of $R_f = 3000$ Ω and SNR = −3 dB, and we are able to verify its reliability through successful field tests. This study provides a reference for the future exploration of the HIF mechanism and highlights the importance of studying the effects of noise interference on algorithm reliability.

REFERENCES

[1] X. Wei, X. Wang, D. Yang *et al*., "Faulty feeder detection for single-phase-to-ground fault in distribution networks based on transient energy and cosine similarity," *IEEE Transactions on Power Delivery*, vol. 37, no. 5, pp. 3968-3979, Oct. 2022.

[2] Q. Cui and W. Yang, "Enhance high impedance fault detection and location accuracy via μ-PMUs," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 797-809, Jan. 2022.

[3] J. Cen, E. Chu, Y. Li *et al*., "Faulty feeder identification and fault area localization in resonant grounding system based on wavelet packet and Bayesian classifier," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 4, pp. 760-767, Jul. 2020.

[4] L. Cheng, T. Wang, and Y. Wang, "A novel fault location method for distribution networks with distributed generations based on the time matrix of traveling-waves," *Protection and Control of Modern Power Systems*, vol. 7, no. 1, pp. 1-11, Nov. 2022.

[5] K. Zhang, Y. Zhu, and X. Liu, "A fault locating method for multi-branch hybrid transmission lines in wind farm based on redundancy parameter estimation," *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 5, pp. 1033-1043, Sept. 2019.

[6] X. Wang, H. Zhang, Q. Wu *et al*., "Location of single phase to ground faults in distribution networks based on synchronous transients energy analysis," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 774-785, Jan. 2020.

[7] K. Liu, B. Li, C. Zhang *et al*., "Flexible grounding system for single-phase to ground faults in distribution networks: a systematic review of developments," *IEEE Transactions on Power Delivery*, vol. 37, no. 3, pp. 1640-1649, Jun. 2022.

[8] B. Wang and X. Cui, "Nonlinear modeling analysis and arc high-impedance faults detection in active distribution networks with neutral grounding via petersen coil," *IEEE Transactions on Smart Grid*, vol. 13, no. 3, pp. 1888-1898, May 2022.

[9] X. Cheng, B. Cui, and S. Hou, "Fault line selection of distribution network based on modified CEEMDAN and GoogLeNet neural networkl," *IEEE Sensors Journal*, vol. 22, no. 13, pp. 13346-13364, Jul. 2022.

[10] X. Zeng, K. Li, W. Chan *et al*., "Ground-fault feeder detection with fault-current and fault-resistance measurement in mine power systems," *IEEE Transactions on Industry Applications*, vol. 44, no. 2, pp. 424-429, Mar. 2008.

[11] L. Zhang, B. Xu, Y. Xue *et al*., "Fault location method based on zero sequence admittance measurement in noneffectively earthed system," *Proceedings of the IEEE PES Innovation. Smart Grid Technology*, Tianjin, China, May 2012, pp. 1-4.

[12] J. Hu, W. Hu, J. Chen *et al*., "Fault location and classification for distribution systems based on deep graph learning methods," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 1, pp. 35-50, Jan. 2023.

[13] H. Hassani, R. Razavi-Far, and M. Saif, "Fault location in smart grids through multicriteria analysis of group decision support systems," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7318-7327, Dec. 2020.

[14] D. R. R. Penido, L. R. D. Araujo, V. T. S. Rodrigues *et al*., "An analytical zero sequence method to locate fault in distribution systems rich in DG," *IEEE Transactions on Smart Grid*, vol. 13, no. 3, pp. 1849-1859, May 2022.

[15] T. Sirojan, S. B. Lu, B. T. Phung *et al*., "Sustainable deep learning at grid edge for real-time high impedance fault detection," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 774-785, Jan. 2020.

[16] K. Łowczowski and B. Olejnik, "Monitoring, detection and locating of transient earth fault using zero-sequence current and cable screen earthing current in medium voltage cable and mixed feeders," *Energies*, vol. 15, no. 3, pp. 1-16, Feb. 2022.

[17] J. Li, Y. Liu, C. Li *et al*., "An FTU-based method for locating single-phase high-impedance faults using transient zero-sequence admittance in resonant grounding systems," *IEEE Transactions on Power Delivery*, vol. 37, no. 2, pp. 913-922, Apr. 2022.

[18] P. Wang, H. Zhou, B. Chen *et al*., "Fault location method in resonant grounded networks based on distributed modulation and compensation adjustment," *IEEE Transactions on Power Delivery*, vol. 34, no. 5, pp. 1938-1947, Oct. 2019.

[19] P. K. Ganivada and P. Jena, "Faulty feeder detection for single-phase-to-ground fault in distribution networks based on transient energy and cosine similarity," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 3000-3010, May 2022.

[20] M. Pignati, L. Zanni, P. Romano *et al*., "Fault detection and faulted line identification in active distribution networks using synchrophasors-based real-time state estimation," *IEEE Transactions on Power Delivery*, vol. 32, no. 1, pp. 381-392, Feb. 2017.

[21] Z. Yin, Z. Wei, G. Sun *et al*., "High sensitivity fault location technology for distribution networks considering measurement error," *International Journal of Electrical Power and Energy Systems*, vol. 140, pp. 1-12, Sept. 2022.

[22] B. Wang, J. Geng, and X. Dong, "High-impedance fault detection based on nonlinear voltage-current characteristic profile identification," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3783-3791, Jul. 2018.

[23] K. Jia, B. Yang, T. Bi *et al*., "An improved sparse-measurement-based fault location technology for distribution networks," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 1712-1720, Mar. 2021.

[24] X. Wei, D. Yang, B. Wang *et al*., "Faulty feeder detection based on fundamental component shift and multiple-transient-feature fusion in distribution networks," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1699-1711, Mar. 2021.

[25] D. Wang, V. Psaras, A. A. S. Emhemed *et al*., "A novel fault let-through energy based fault location for LVDC distribution networks," *IEEE Transactions on Power Delivery*, vol. 36, no. 2, pp. 966-974, Apr. 2021.
[26] C. A. Apostolopoulos, C. G. Arsoniadis, P. S. Georgilakis *et al*., "Unsynchronized measurements based fault location algorithm for active distribution systems without requiring source impedances," *IEEE Transactions on Power Delivery*, vol. 37, no. 3, pp. 2071-2082, Jun. 2022.
[27] Y. Xue, J. Li, and B. Xu, "Faulty feeder selection and transition resistance identification of high impedance fault in a resonant grounding system using transient signals," *Proceedings of the CSEE*, vol. 37, no.17, pp. 5037-5048, Sept. 2017.
[28] Y. Yang, H. Liu, L. Hang *et al*., "A feature extraction method using VMD and improved envelope spectrum entropy for rolling bearing fault diagnosis," *IEEE Sensors Journal*, vol. 23, no. 4, pp. 3848-3858, Feb. 2023.
[29] T. Chen, S. Gao, S. Zheng *et al*., "EMD and VMD empowered deep learning for radio modulation recognition," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 1, pp. 43-57, Feb. 2023.
[30] Z. Xing, Y. He, X. Wang *et al*., "VMD-IARIMA-based time-series forecasting model and its application in dissolved gas analysis," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 30, no. 2, pp. 802-811, Apr. 2023.
[31] Y. Yang, H. Liu, L. Ha *et al*., "Afeature extraction method using VMD and improved envelope spectrum entropy for rolling bearing fault diagnosis," *IEEE Sensors Journal*, vol. 23, no. 4, pp. 3848-3858, Feb. 2023.
[32] G. C. Linderman, M. Rachh, J. K. Hoskins *et al*., "Fast interpolation-based t-SNE for improved visualization of single-cell RNA-seq data," *Nature Methods*, vol. 16, no. 3, pp. 1-25, Mar. 2019.
[33] B. Ullah, M. Kamran, and Y. Rui, "Predictive modeling of short-term rockburst for the stability of subsurface structures using machine learning approaches: t-SNE, *K*-means clustering and XGBoost," *Mathematics*, vol. 10, no. 3, pp. 1-20, Feb. 2022.
[34] X. Wang, J. Gao, X. Wei *et al*., "Faulty feeder detection under high impedance faults for resonant grounding distribution systems," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 1880-1894, May 2023.
[35] X. Wang, Y. Wei, M. Chen *et al*., "Faulty line detection method based on optimized bistable system for distribution network," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1370-1381, Apr. 2018.

Zhihua Yin received the B.Sc. degree from China University of Mining and Technology, Xuzhou, China, in 2014, and the M.S. degree from Hohai University, Nanjing, China, in 2018, where he is currently pursuing the Ph.D. degree. His research interest includes fault diagnosis of power distribution networks.

Yuping Zheng received the B.Sc. degree from Hefei University of Technology, Hefei, China, in 1983, the M.S. degree from Nanjing Automation Research Institute, Nanjing, China, in 1986, and the Ph.D. degree from Wuhan University, Wuhan,China, in 2004. His research interests include relay protection and control of AC/DC hybrid power system.

Zhinong Wei received the B.Sc. degree from Hefei University of Technology, Hefei, China, in 1984, the M.S. degree from Southeast University, Nanjing, China, in 1987, and the Ph.D. degree from Hohai University, Nanjing, China, in 2004. His research interests include power system operation, analysis, and control.

Guoqiang Sun received the B.Sc., M.S., and Ph.D. degrees in electrical engineering from Hohai University, Nanjing, China, in 2001, 2005, and 2010, respectively. He was a Visiting Scholar at North Carolina State University, Raleigh, USA, from 2015 to 2016. His research interests include power system analysis, economic dispatch, and optimal control of integrated energy systems

Sheng Chen received the B.Sc. and Ph.D. degrees in electrical engineering from Hohai University, Nanjing, China, in 2014 and 2019, respectively. From January 2018 to 2019, he was a Visiting Scholar with The Ohio State University, Columbus, USA. His research interests include economic dispatch and security analysis of integrated energy systems.

Haixiang Zang received the B.S. and Ph.D. degrees in electrical engineering from Nanjing Normal University, Nanjing, China and Southeast University, Nanjing, China, in 2009 and 2014, respectively. His research interests include generation of renewable energy, operation, and control of power system.