



ARTICLE

Research on Asymmetric Fault Location of Wind Farm Collection System Based on Compressed Sensing

Huanan Yu¹, Gang Han^{1,*}, Hansong Luo² and He Wang¹¹Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education, Northeast Electric Power University, Jilin, 132012, China²Qinhuangdao Power Supply Company, State Grid Jibei Electric Power Co., Ltd., Qinhuangdao, 066099, China

*Corresponding Author: Gang Han. Email: 17865578539@163.com

Received: 14 December 2022 Accepted: 10 March 2023 Published: 03 August 2023

ABSTRACT

Aiming at the problem that most of the cables in the power collection system of offshore wind farms are buried deep in the seabed, which makes it difficult to detect faults, this paper proposes a two-step fault location method based on compressed sensing and ranging equation. The first step is to determine the fault zone through compressed sensing, and improve the data measurement, dictionary design and algorithm reconstruction: Firstly, the phase-locked loop trigonometric function method is used to suppress the spike phenomenon when extracting the fault voltage, so that the extracted voltage value will not have a large error due to the voltage fluctuation. Secondly, the λ -NIM dictionary is designed by using the node impedance matrix and the fault location coefficient to further reduce the influence of pseudo-fault points. Finally, the CoSaMP algorithm is improved with the generalized Jaccard coefficient to improve the reconstruction accuracy. The second step is to use the ranging equation to accurately locate the asymmetric fault of the wind farm collection system on the basis of determining the fault interval. The simulation results show that the proposed method is more accurate than the compressed sensing method and impedance method in fault section location and fault location accuracy, the relative error is reduced from 0.75% to 0.4%, and has a certain anti-noise ability.

KEYWORDS

Offshore wind farm; convergence system; compression sensing; ranging equation; fault location

Nomenclature

λ	Fault location coefficient
$I^{(2)}$	Fault negative sequence total current, kA
$I_p^{(2)}$	Negative sequence current injected at fault point, kA
\vec{I}	Equivalent vector group of fault negative sequence current
Z	Bus impedance matrix
ΔU	Fault negative sequence voltage vector group
$\Delta U_i^{(2)}$	Negative sequence voltage component of node i , kV
$U_i^{(a)}$	Negative sequence component of node i before system failure, kV
$U_i^{(2)}$	Negative sequence component of node i after system fault, kV
Z_{mn}	Negative sequence impedance between node m and node n , Ω



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

U_{mn}	Potential difference between node m and node n , kV
I_p	Fault negative sequence current source, kA
I_{p1}	Negative sequence current flowing from fault point to node m , kA
I_{p2}	Negative sequence current flowing from fault point to node n , kA
I_{pm}	Virtual equivalent current source of node m , kA
I_{pn}	Virtual equivalent current source of node n , kA
I'_{pm1}	Sum of all negative sequence currents outside node m , kA
I'_{pn1}	Sum of all negative sequence currents outside node n , kA
I_{pm1}	Negative sequence current from node m to node n , kA
I_{pn1}	Negative sequence current from node n to node m , kA
y	Measurement matrix
Φ	Sensing matrix
x	Fault sparse vector
$I_m^{(2)}$	Negative sequence current source of node m , kA
$I_n^{(2)}$	Negative sequence current source of node n , kA
$U_i^{(0)}$	Zero-sequence voltage at node i , kV
$U_0^{(0)}$	Zero sequence voltage value at collecting bus, kV
R	Resistance of grounding transformer, Ω
Λ_0	Index set
A_0	Support set
Z_m	Equivalent impedance of all lines before node m , Ω
Z_n	Equivalent impedance of all lines after node n , Ω

1 Introduction

In the 2020 ‘Wind Energy Beijing Declaration’, it is proposed to comprehensively consider the realistic feasibility of resource potential, technological progress trend, and grid-connected consumption conditions. In order to achieve the goal of connecting with the goal of carbon neutrality, in the ‘14th Five-Year Plan’, it is necessary to set the development space for wind power that is compatible with the national strategy of carbon neutrality: to ensure that the average annual increase in installed capacity is more than 50 million kilowatts. After 2025, the average annual increase in installed capacity of wind power in China should not be less than 60 million kilowatts, reaching at least 800 million kilowatts by 2030 and at least 3 billion kilowatts by 2060 [1]. The offshore wind farm collection system is composed of a large number of wind turbines. Most of the transmission lines in the collection system are buried deep in the seabed, which makes it difficult to locate the asymmetric fault and the fault recovery time is long. Therefore, when an asymmetric fault occurs in the collection system, its rapid and accurate positioning will help improve the utilization rate of wind power.

At present, the research methods for wind farm fault location mainly include the impedance method [2–5], the traveling wave method [6–8] and the deep learning method [9–14]. Reference [2] proposed a single-phase grounding fault location method based on successive search of fault sections, and introduced a logarithmic arc model describing the arc grounding of fault points to achieve high-ranging accuracy. In reference [3], the author constructed the impedance matching ranging algorithm from the first section to the section where the fault point is located by constructing a single interval ranging model and an equivalent injection current model of the fan branch. In reference [4], the author constructed a transmission equation containing the location of the fault point through the traveling wave characteristics. The overdetermined equations formed by the multi-point transmission equation are solved by the parameter estimation theory to obtain the fault location. Reference [5] proposed a

combined ranging method for wind farm collector lines using a combination of matrix method and impedance method. In reference [6], the traditional traveling wave method is optimized to improve the accuracy of traveling wave detection, and the traveling wave velocity is optimized and corrected to realize the accurate location of the fault distance of the wind farm collector line and the transmission line. Reference [7] proposed a fault location method based on decision coefficient and extreme-point symmetric mode decomposition-Teager energy operator (ESMD-TEO). Reference [8] deduced a calculation method of modified traveling wave velocity based on the principle of single-ended traveling wave and combined with double-ended asynchronous information. Reference [9] collected the fault data of the collecting line of the wind farm and trained the deep learning neural network to establish a prediction model for single terminal fault location. Reference [10] proposed a new spatio-temporal multiscale neural network (STMNN). The proposed STMNN model contains two parallel feature extraction modules: first, a multiscale deep echo state network module to extract temporal multiscale features; second, a multiscale residual network module to extract spatial multiscale features. Reference [11] proposed a new multi-level denoising autoencoder (MLD-AE) method, which enhances the representation learning ability by designing different multi-level noise addition schemes based on the multivariate data-driven fault detection (MDFD) framework. Reference [12] proposed a hybrid fault detection system based on single input generalized regression neural network ensemble (GRNN-ESI) algorithm, principal component analysis (PCA) and wavelet probability density function (PDF). Reference [13] proposed a monitoring method based on conditional convolution autoencoder, which has duality and can be used to identify the damage of wind turbine blades. Reference [14] proposed an anomaly detection method for wind turbine gearbox based on adaptive threshold and double support vector machine (TWSVM). Furthermore, some new but important work in fault diagnosis is worth investigating. Reference [15] proposed a class-imbalanced privacy-preserving federated learning framework for distributed wind turbine fault diagnosis that incorporates a separate gradient-based self-monitoring scheme to identify global imbalance information for class-imbalanced fault diagnosis. A novel framework for multi-domain vibration feature extraction, feature selection, and cost-sensitive learning methods was developed in reference [16]. Extensive experiments show that multi-domain feature extraction and feature selection can improve diagnostic accuracy significantly.

In summary, the impedance method uses the voltage and current data obtained at the measuring point after the fault occurs, and the total impedance of the fault circuit is calculated by the voltage equation constructed by the circuit, and then the fault distance is obtained [17]. However, the initial current of the collection system is not only affected by the fault point shape, but also by the output of the wind turbine on different branches. Therefore, it is difficult to list an accurate fault loop voltage equation only based on the measured value of the collector line. For the traditional traveling wave method, due to the complex lines in the wind farm and the short transmission lines between the wind turbines, the refraction and reflection of the traveling wave after the fault are complex and changeable. Therefore, it is difficult to use the traveling wave method for fault location, and the traveling wave device is expensive. The deep learning method has a long training time, and the algorithm requires a large amount of fault data for operation and learning. However, the wind farm has high real-time requirements and it is difficult to provide so much measured data.

Compressed sensing theory has been increasingly used in the field of power system fault location in recent years. Reference [18] achieved accurate fault location based on sparse measurement by combining the node high-frequency voltage equation with the compressed sensing theory based on the Bayesian algorithm. Reference [19] proposed a single-ended fault location method based on atomic decomposition and traveling wave natural frequency. The Dice coefficient atom matching criterion and Fourier transform were introduced in reference [20] to improve the regularized orthogonal matching

pursuit (ROMP) algorithm in the over-complete dictionary design to improve accuracy and achieve accurate fault location of HVDC lines.

Based on the above analysis, this paper analyzes the fault current sparsity characteristics in the wind farm collection system when asymmetric faults occur. Taking advantage of the fact that the negative sequence current generated by the wind turbine generator is almost zero when an asymmetric fault occurs in the system, the mathematical model of compression sensing is built through the node voltage equation. In the process of fault location, the phase-locked loop trigonometric function method is used to extract the fault negative sequence voltage, and the node impedance matrix is combined with the fault location coefficient to construct a new over-complete dictionary λ -NIM dictionary (λ -Node Impedance Matrix dictionary, λ -NIM). The algorithm is then improved by employing generalized Jaccard coefficients, yielding a generalized Jaccard Compressive Sampling matching pursuit (J-CoSaMP) algorithm. The algorithm is then used to locate the fault zone. Finally, using the ranging equation, the fault location is precisely determined based on the fault interval. The proposed method is anti-noisy and is less affected by fault location and transition resistance, allowing it to locate the fault accurately and quickly.

2 Fault Current Characteristic Analysis and Equivalent Replacement of Wind Farm

2.1 Analysis of Fault Current Sparse Characteristics

To be connected to the power grid, the wind turbines in the wind farm must pass through their respective box transformers, collection systems, transmission lines, and main transformers. Collection lines and collection buses are part of the collection system. Fig. 1 depicts the topology of the wind farm collection system.

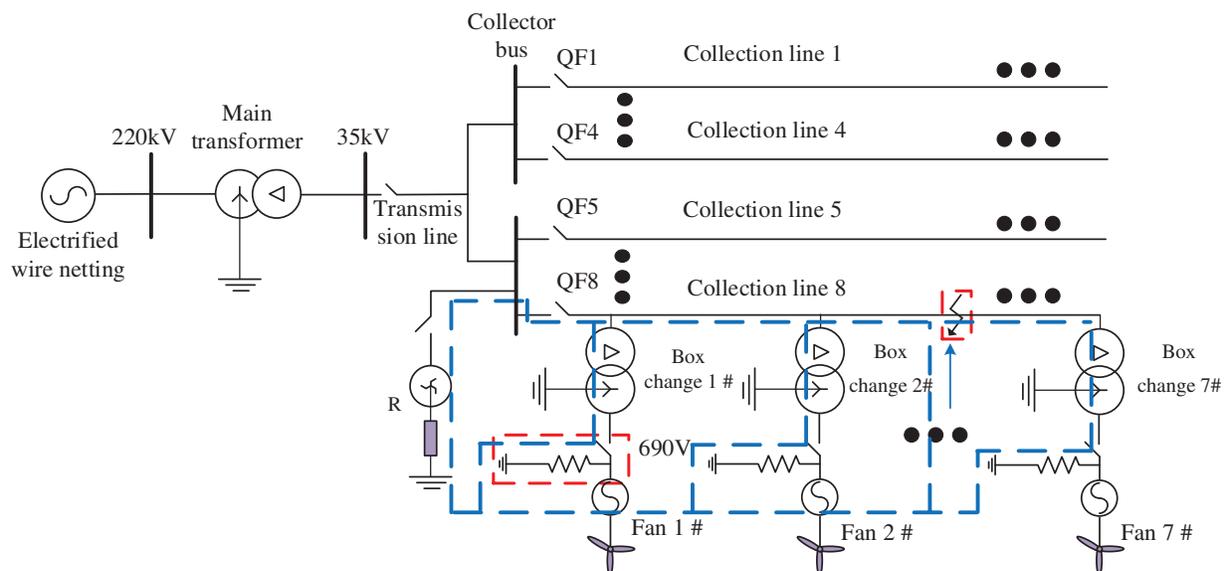


Figure 1: Topology of wind farm collection system

Because of the wind farm collection system's asymmetric faults, namely single-phase ground fault, phase-to-phase fault, and two-phase ground fault, the negative sequence current emitted by the wind turbine is very small, almost negligible in comparison to the negative sequence current injected at the

fault point. The negative sequence current flow path is shown in Fig. 2 when a single-phase ground fault, inter-phase fault, or two-phase ground fault occurs in the collection system.

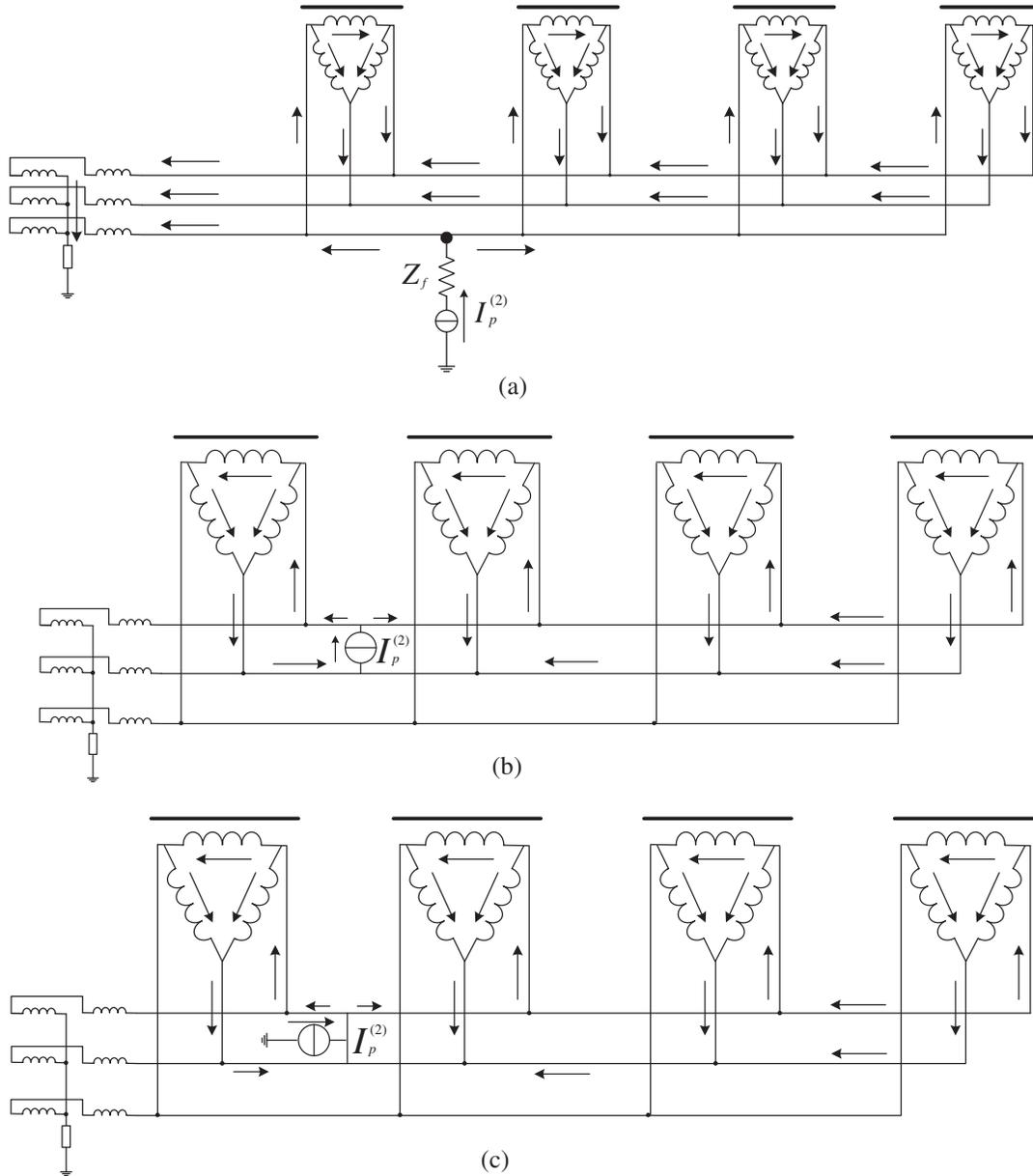


Figure 2: Negative sequence current flow path diagram. (a) Negative sequence current flow path in single-phase grounding fault, (b) Negative sequence current flow path in single-phase grounding fault, (c) Negative sequence current flow path in two-phase grounding fault

The short-circuit current negative sequence component $I^{(2)}$ in the wind farm collection system is only composed of the fault point negative sequence current, as shown in Fig. 2, and its expression is:

$$I^{(2)} = I_p^{(2)} \tag{1}$$

As a result, the fault negative sequence current is sparse, which meets the requirement of using compressed sensing theory. However, because underdetermined equations must be constructed in power system fault location, the compressed sensing theory can be introduced, and the negative sequence current just meets this condition. It flows into the earth from the three-phase load via the box-type transformer, resulting in the presence of the negative sequence current component in all transmission lines in the collection system. The node negative sequence voltage equations are written when the negative sequence voltage component and the node impedance matrix are combined. The underdetermined equations can be constructed by equivalently replacing the fault negative sequence current. Because single-phase ground faults account for the majority of asymmetric faults in wind farm collection systems, this paper chooses the single-phase ground fault experimental research object.

2.2 Equivalent Replacement of Fault Current

This paper selects the negative sequence current for equivalent substitution by analyzing the sparse characteristics of the acquisition system's fault current. Using the wind farm collection system with N nodes as an example, the fault component negative sequence network's node equivalent current substitution equation $\Delta U = Z \times I$ is constructed, as shown in Eq. (2). I is the equivalent component of the fault negative sequence current, Z is the node impedance matrix, and ΔU is the fault negative sequence voltage component.

$$\begin{bmatrix} \Delta U_1^{(2)} \\ \vdots \\ \Delta U_m^{(2)} \\ \Delta U_n^{(2)} \\ \vdots \\ \Delta U_N^{(2)} \end{bmatrix}_{N \times 1} = \begin{bmatrix} Z_{11} & \cdots & Z_{1m} & Z_{1n} & \cdots & Z_{1N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Z_{m1} & \cdots & Z_{mm} & Z_{mn} & \cdots & Z_{mN} \\ Z_{n1} & \cdots & Z_{nm} & Z_{nn} & \cdots & Z_{nN} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & \cdots & Z_{Nm} & Z_{Nn} & \cdots & Z_{NN} \end{bmatrix}_{N \times 1} \begin{bmatrix} I_1^{(2)} \\ \vdots \\ I_m^{(2)} \\ I_n^{(2)} \\ \vdots \\ I_N^{(2)} \end{bmatrix}_{N \times 1} \quad (2)$$

where $\Delta U_i^{(2)} = U_i^{(a)} - U_i^{(2)}$ is the negative sequence voltage component of node i , when the wind farm collection system does not fail, the negative sequence component and the zero sequence component in the system do not exist. As a result, the negative sequence component $U_i^{(a)}$ of node i prior to system failure has a value of 0, whereas the negative sequence component of node i following failure has a value of $U_i^{(2)}$.

A single-phase grounding fault develops between nodes m and n , as depicted in Fig. 3. Line impedance between nodes is Z_{mn} , and the fault location coefficient, which measures the proportion of the line's entire length (0, 1) that is traveled from node m to the fault point p , is λ .

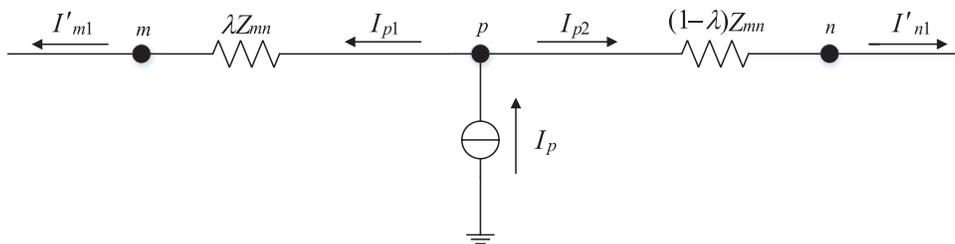


Figure 3: Schematic diagram of inter node fault

From Fig. 3:

$$\begin{cases} U_{mn} = -\lambda Z_{mn} I_{p1} + (1 - \lambda) Z_{mn} I_{p2} \\ I_p = I_{p1} + I_{p2} \end{cases} \quad (3)$$

where U_{mn} is the potential difference between node m and node n ; I_p is the fault negative sequence current source; I_{p1} and I_{p2} are negative sequence current flowing from fault point to node m and node n .

From Fig. 4:

$$\begin{cases} U_{mn} = Z_{mn} (I_{pm} - I'_{pm1}) \\ I_{pm} = I'_{pm1} + I_{pn} \\ I_{pn} = I'_{pn1} + I_{pm1} \end{cases} \quad (4)$$

where I_{pm} and I_{pn} are the virtual equivalent current sources of the fault negative sequence current source I_p replaced by node m and node n , respectively; I'_{pm1} and I'_{pn1} are the sum of all negative sequence currents outside nodes m and n after I_p replacement; I_{pm1} is the negative sequence current flowing from node m to node n ; I_{pn1} is the negative sequence current from node n to node m .

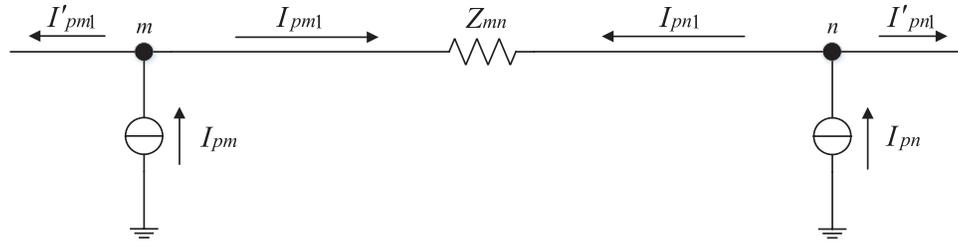


Figure 4: Schematic diagram of fault between nodes after equivalent

According to the KCL law, the nodes m and n , as well as the lines connecting them, can be considered as a whole, and the current of the outer line remains constant both before and after equivalence, that is:

$$\begin{cases} I'_{pm1} = I'_{m1} \\ I'_{pn1} = I'_{n1} \end{cases} \quad (5)$$

where I'_{m1} and I'_{n1} are the sum of all negative sequence currents outside nodes m and n before the fault negative sequence current source is replaced I_p .

The equivalent substitution relation of fault negative sequence current can be obtained by combining Eqs. (3)–(5):

$$\begin{cases} I_{pn} = \lambda I_p \\ I_{pm} = (1 - \lambda) I_p \\ I_{pm} + I_{pn} = I_p \end{cases} \quad (6)$$

When a fault occurs, a single fault negative sequence current source in Fig. 2 can be equivalent to a virtual equivalent current source of two adjacent nodes, as shown by Eq. (6). As shown in Fig. 3, the node's negative sequence voltage equation is:

$$\begin{bmatrix} \Delta U_1^{(2)} \\ \vdots \\ \Delta U_m^{(2)} \\ \Delta U_n^{(2)} \\ \vdots \\ \Delta U_N^{(2)} \end{bmatrix}_{N \times 1} = \begin{bmatrix} Z_{11} & \cdots & Z_{1m} & Z_{1n} & \cdots & Z_{1N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Z_{m1} & \cdots & Z_{mm} & Z_{mn} & \cdots & Z_{mN} \\ Z_{n1} & \cdots & Z_{nm} & Z_{nn} & \cdots & Z_{nN} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & \cdots & Z_{Nm} & Z_{Nn} & \cdots & Z_{NN} \end{bmatrix}_{N \times N} \begin{bmatrix} 0 \\ \vdots \\ I_m^{(2)} \\ I_n^{(2)} \\ \vdots \\ 0 \end{bmatrix}_{N \times 1} \quad (7)$$

Although the Eq. (7) is made up of N equations, in practice, only M ($M < N$) nodes are chosen to install the voltage measuring device in the entire collection system, and M equations are obtained as follows:

$$\begin{bmatrix} \Delta U_{i_1}^{(2)} \\ \Delta U_{i_2}^{(2)} \\ \vdots \\ \Delta U_{i_M}^{(2)} \end{bmatrix}_{M \times 1} = \begin{bmatrix} Z_{i_1 1} & Z_{i_1 2} & \cdots & Z_{i_1 N} \\ Z_{i_2 1} & Z_{i_2 2} & \cdots & Z_{i_2 N} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{i_M 1} & Z_{i_M 2} & \cdots & Z_{i_M N} \end{bmatrix}_{M \times N} \begin{bmatrix} 0 \\ \vdots \\ I_m^{(2)} \\ I_m^{(2)} \\ \vdots \\ 0 \end{bmatrix}_{N \times 1} \quad (8)$$

where $\Delta U_{i_1}^{(2)}$ denotes the negative sequence voltage component at node 1 following the failure of node i . Because Eq. (8) is an underdetermined system with a sparse fault quantity, the compressed sensing theory is used to accurately reconstruct the fault negative sequence current.

3 Fault Location Dictionary Design and Algorithm Improvement

3.1 Compressed Sensing Theory

By employing a lower sampling frequency algorithm, compressed sensing theory can recover the fault sparse vector. Its mathematical model is as follows:

$$\mathbf{y} = \Phi \mathbf{x} \quad (9)$$

where \mathbf{y} is the observation matrix; Φ is a sensing matrix; \mathbf{x} is the fault sparse vector.

The reconstruction algorithm can be used to recover the fault sparse vector \mathbf{x} from the measurement matrix \mathbf{y} as follows:

$$\min \|\mathbf{x}\|_0, \text{ s.t. } \mathbf{y} = \Phi \mathbf{x} \quad (10)$$

It is challenging to directly solve the l_0 norm in Eq. (10) and it is an NP-hard problem [21]. According to published research [22], if the sensing matrix meets the Restricted Isometry Property (RIP) and $\delta \in (0, 1)$:

$$(1 - \delta) \|\mathbf{x}\|_2^2 \leq \|\Phi \mathbf{x}\|_2^2 \leq (1 + \delta) \|\mathbf{x}\|_2^2 \quad (11)$$

The l_0 norm and l_1 norm can be transformed, that is:

$$\min \|\mathbf{x}\|_1, \text{ s.t. } \mathbf{y} = \Phi \mathbf{x} \quad (12)$$

3.2 Fault Voltage Extraction and λ -NIM Dictionary Design

The extraction of negative sequence fault voltage amplitude and the choice of an over-complete dictionary will impair the accuracy of fault location when compressed sensing is introduced for fault location in accordance with Eq. (8). In order to generate the λ -NIM dictionary, this study enhances

the fault voltage amplitude extraction scheme and incorporates the fault location coefficient in the node impedance matrix.

3.2.1 Fault Voltage Extraction Based on Phase-Locked Loop Trigonometric Function Method

The conventional voltage amplitude extraction method typically uses the instantaneous symmetrical component approach to extract the voltage component for the negative sequence fault. However, due to the peculiarities of the method, a non-model error will be produced at the voltage mutation point during the process of extracting the fault voltage, resulting in a spike occurrence, as illustrated in Fig. 5. The extracted negative sequence fault voltage amplitude deviation will cause reconstruction vector positioning inaccuracy, which is prone to erroneous fault points, and will also have an impact on the outcomes of the succeeding ranging when it is too great.

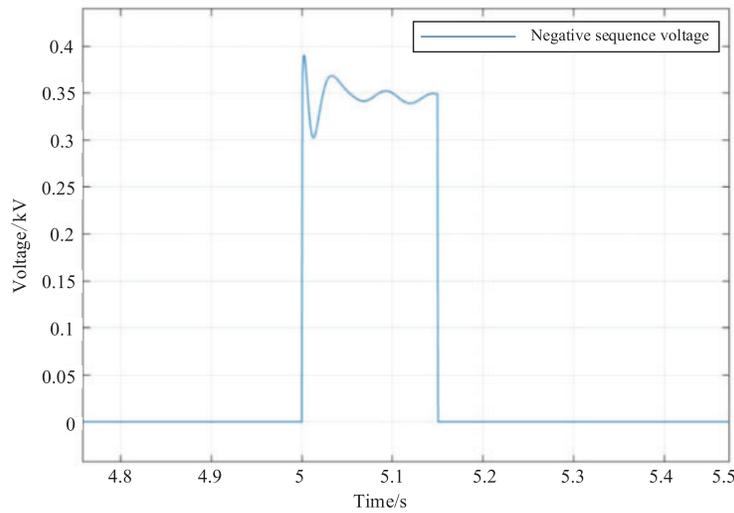


Figure 5: Extracting negative sequence voltage waveform by symmetrical component method

The conventional phase-locked loop trigonometric function method employs a low-pass filtering technique to reduce spike phenomena and improve fault location accuracy. When an asymmetric fault develops in the wind farm collection system, the collection line contains a large number of harmonic components. As a result, in this study [23], the active filter and phase-locked loop trigonometric function approaches are combined. The IGBT inverter module’s efficient filtering out of each harmonic improves the extraction accuracy of the fault voltage even more. Fig. 6 shows how well this strategy prevents the spike phenomenon. The overall flow chart is shown in Fig. 7.

3.2.2 Design Based on Fault Location Parameters λ -NIM Dictionary

The fault location coefficient is introduced in this paper to reduce the probability of false fault point and shorten the time of fault location. Eq. (7) is transformed, and the sparsity K of the fault negative sequence current is changed to 1, so that the fault location can be found with only one traversal in the subsequent algorithm reconstruction process. The value of fault current $I_p^{(2)}$ is the square of the sum of the arithmetic square roots of the virtual equivalent current sources $I_m^{(2)}$ and $I_n^{(2)}$. From Eq. (7):

$$\begin{cases} \Delta U_m^{(2)} = Z_{mm} \cdot I_m^{(2)} + Z_{mn} \cdot I_n^{(2)} \\ \Delta U_n^{(2)} = Z_{nm} \cdot I_m^{(2)} + Z_{nn} \cdot I_n^{(2)} \end{cases} \quad (13)$$

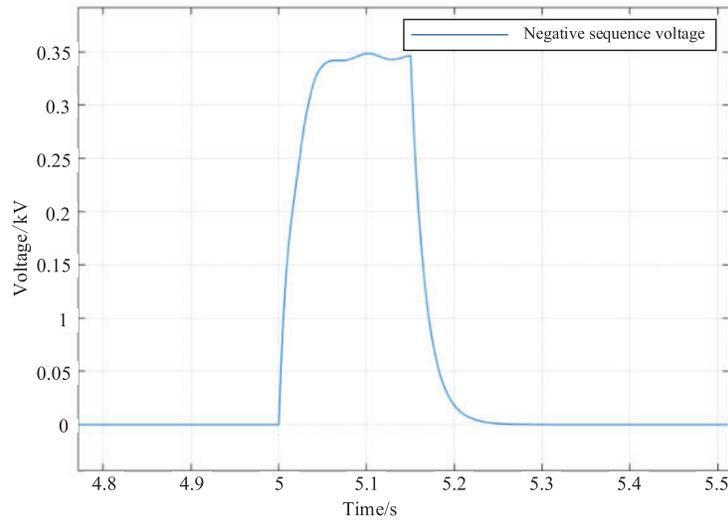


Figure 6: Extracting negative sequence voltage waveform by phase locked loop trigonometric function method

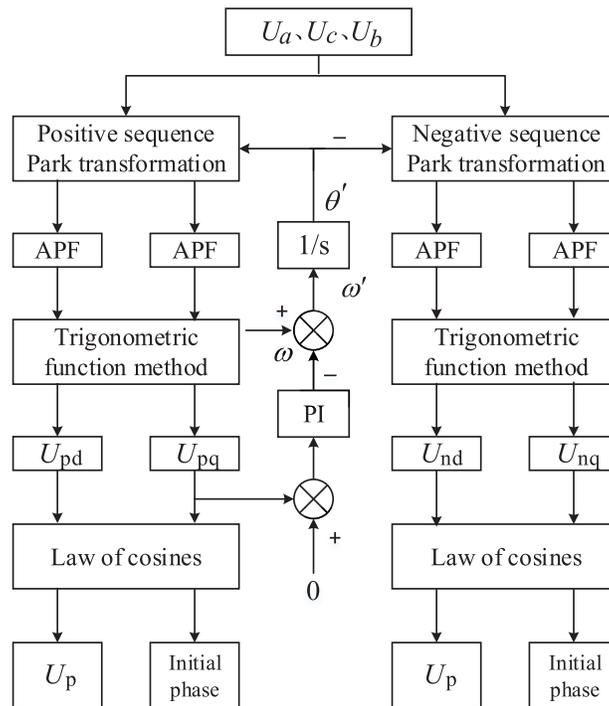


Figure 7: Phase locked loop trigonometric function method

From Eqs. (6) and (7):

$$\begin{cases} \Delta U_m^{(2)} = [(1 - \lambda) \cdot Z_{mm} + \lambda \cdot Z_{nn}] \cdot (I_m^{(2)} + I_n^{(2)}) \\ \Delta U_n^{(2)} = [(1 - \lambda) \cdot Z_{nn} + \lambda \cdot Z_{mm}] \cdot (I_m^{(2)} + I_n^{(2)}) \end{cases} \quad (14)$$

Thus, by Eq. (14), Eq. (7) becomes:

$$\begin{bmatrix} \Delta U_1^{(2)} \\ \Delta U_2^{(2)} \\ \vdots \\ \Delta U_N^{(2)} \end{bmatrix}_{N \times 1} = \begin{bmatrix} (1 - \lambda) Z_{11} + \lambda Z_{12} & (1 - \lambda) Z_{12} + \lambda Z_{13} & \cdots & (1 - \lambda) Z_{1,N-1} + \lambda Z_{1N} \\ (1 - \lambda) Z_{21} + \lambda Z_{22} & (1 - \lambda) Z_{22} + \lambda Z_{23} & \cdots & (1 - \lambda) Z_{2,N-1} + \lambda Z_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ (1 - \lambda) Z_{N1} + \lambda Z_{N2} & (1 - \lambda) Z_{N2} + \lambda Z_{N3} & \cdots & (1 - \lambda) Z_{N,N-1} + \lambda Z_{NN} \end{bmatrix}_{N \times 1} \begin{bmatrix} 0 \\ \vdots \\ I_p^{(2)} \\ 0 \end{bmatrix}_{N \times 1} \quad (15)$$

However, because the fault location coefficient has a value range of (0, 1), when all of them are substituted into Eq. (15), the number of columns in the λ -NIM dictionary will reach hundreds of columns, resulting in a long convergence time for the compressed sensing algorithm. This paper estimates $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_N\}$, where λ_1 is the estimated value of the fault location coefficient between the first and second wind turbines; $U_i^{(0)}$ is the zero-sequence voltage at node i ; R is the resistance of the grounding transformer; $U_0^{(0)}$ is the zero-sequence voltage value at the collection bus; and Z_{ij} is the impedance value between nodes i and j .

$$\lambda_1 = \frac{U_2^{(0)} R - U_0^{(0)} R - U_0^{(0)} Z_{01}}{U_0^{(0)} Z_{12}} \quad (16)$$

$$\lambda_2 = \frac{U_3^{(0)} R - U_0^{(0)} R - U_0^{(0)} (Z_{01} + Z_{12})}{U_0^{(0)} Z_{23}} \quad (17)$$

$$\lambda_3 = \frac{U_4^{(0)} R - U_0^{(0)} R - U_0^{(0)} (Z_{01} + Z_{12} + Z_{23})}{U_0^{(0)} Z_{34}} \quad (18)$$

$$\begin{cases} \frac{U_2^{(0)}}{Z_{12} + Z_{01} + R} = \frac{U_0^{(0)}}{R} \\ \frac{U_3^{(0)}}{Z_{12} + Z_{01} + Z_{23} + R} = \frac{U_0^{(0)}}{R} \\ \frac{U_4^{(0)}}{Z_{12} + Z_{01} + Z_{23} + Z_{34} + R} = \frac{U_0^{(0)}}{R} \end{cases} \quad (19)$$

The combined Eqs. (16)–(19) can be obtained:

$$\begin{cases} \lambda_2 = (\lambda_1 - 1) \cdot \frac{Z_{12}}{Z_{23}} + 1 \\ \lambda_3 = (\lambda_2 - 1) \cdot \frac{Z_{23}}{Z_{34}} + 1 \end{cases} \quad (20)$$

By the Eq. (20) can predict the value of $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_N\}$, then the Eq. (15) becomes:

$$\begin{bmatrix} \Delta U_1^{(2)} \\ \Delta U_2^{(2)} \\ \vdots \\ \Delta U_N^{(2)} \end{bmatrix}_{N \times 1} = \begin{bmatrix} (1 - \lambda_1) Z_{11} + \lambda_1 Z_{12} & (1 - \lambda_2) Z_{12} + \lambda_2 Z_{13} & \cdots & (1 - \lambda_{N-1}) Z_{1,N-1} + \lambda_{N-1} Z_{1N} \\ (1 - \lambda_1) Z_{21} + \lambda_1 Z_{22} & (1 - \lambda_2) Z_{22} + \lambda_2 Z_{23} & \cdots & (1 - \lambda_{N-1}) Z_{2,N-1} + \lambda_{N-1} Z_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ (1 - \lambda_1) Z_{N1} + \lambda_1 Z_{N2} & (1 - \lambda_2) Z_{N2} + \lambda_2 Z_{N3} & \cdots & (1 - \lambda_{N-1}) Z_{N,N-1} + \lambda_{N-1} Z_{NN} \end{bmatrix}_{N \times N} \begin{bmatrix} 0 \\ \vdots \\ I_p^{(2)} \\ 0 \end{bmatrix}_{N \times 1} \quad (21)$$

3.3 J-CoSaMP Algorithm

To compute the residual modulus, the Jaccard-CoSaMP algorithm proposed in this paper employs the generalized Jaccard coefficient matching criterion [24] rather than the inner product criterion. The results are shown in Fig. 8.

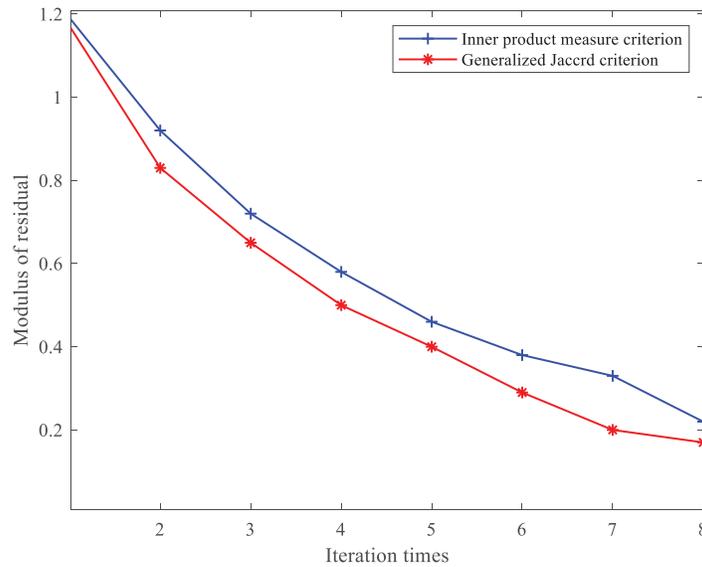


Figure 8: Comparison of residual calculation results

As a result, the difference between the selected atom and the residual signal is smaller when compared to the traditional algorithm. The data is then normalized to improve the rate of atomic convergence. The algorithm's main steps are as follows:

Initialization: residual $r_0 = y$, index set $\Lambda_0 = \emptyset$, support set $A_0 = \emptyset$, $t = 1$;

1. The correlation coefficient u is calculated by the generalized Jaccard coefficient matching criterion: $u = |\text{Jaccard}(A^T, r_{k-1})|$; and normalized and arranged in descending order;

2. Select the maximum $2K$ values for the result of Step 1 to form the set a_j and its corresponding subscripts to form J_0 ; updating index set $\Lambda_t = \Lambda_{t-1} \cup J_0$ and support set $A_t = A_{t-1} \cup a_j$;

3. Using the least square method to find the least square solution of $y = A_t x_t$: $\hat{x}_t = \arg \min \|y - A_t x_t\| = (A_t^T A_t)^{-1} A_t^T y$;
4. For the result of Step 3, K values with the largest absolute value are selected to form set \hat{x}_{tK} , K columns in the corresponding support set are marked as A_{tK} , the sequence number of the corresponding support set is marked as Λ_{tK} , and index set $\Lambda_t = \Lambda_{tK}$ and residual $r_t = y - A_{tK} \hat{x}_{tK}$ are updated;
5. Get the reconstructed signal \hat{x} .

The specific steps of the algorithm are shown in Fig. 9.

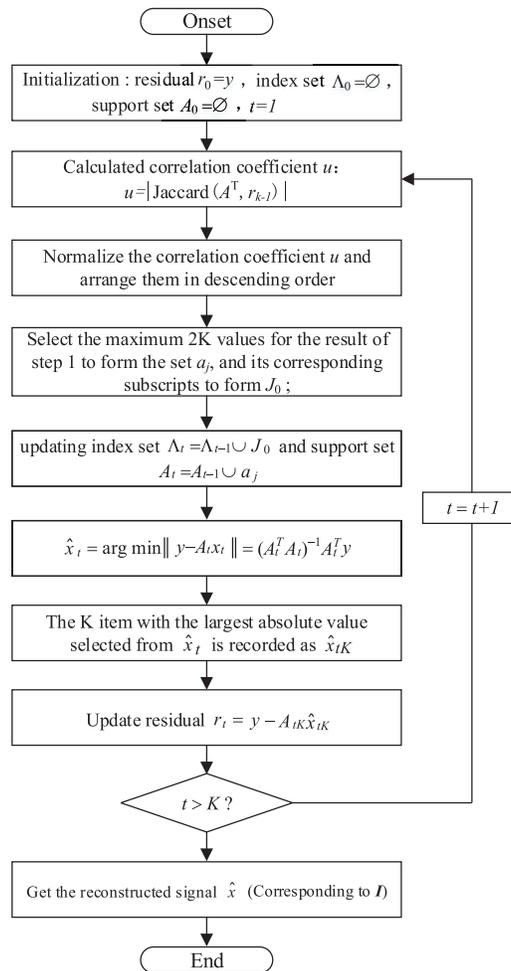


Figure 9: Flow chart of J-CoSaMP algorithm

4 Collection System Fault Location Equation and Steps

4.1 Fault Location Equation of Collecting System

When the negative sequence component I of the fault current is reconstructed using compressed sensing, it is fed into Eq. (21) to solve, and the negative sequence voltage amplitude of the adjacent

nodes at both ends of the fault point is obtained. Following that, ranging is done in conjunction with impedance characteristics.

Reference [25] developed the ranging equation to accurately locate the fault by examining the negative sequence voltage components and impedance characteristics at both ends of the fault section. In the event of a system fault, the negative sequence network diagram is depicted in Fig. 10, and the distance between the fault and the negative sequence voltage is as follows:

$$\frac{\Delta U_m^{(2)}}{\Delta U_n^{(2)}} = \frac{(1 - \lambda) Z_{mn} + Z_n}{\lambda Z_{mn} + Z_m} \cdot \frac{Z_m}{Z_n} \quad (22)$$

where Z_m is the equivalent impedance of all lines before node m ; Z_n is the equivalent impedance of all lines after node n ; Z_{mn} is the mutual impedance between node m and node n .

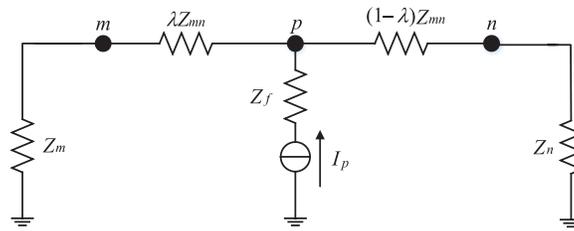


Figure 10: Fault zero sequence network diagram

The negative sequence voltage amplitude of two adjacent nodes is used in this paper to locate the fault using a Eq. (23):

$$\left| \frac{\Delta U_m^{(2)}}{\Delta U_n^{(2)}} \right| = \left| \frac{(1 - \lambda) Z_{mn} + Z_n}{\lambda Z_{mn} + Z_m} \right| \cdot \left| \frac{Z_m}{Z_n} \right| \quad (23)$$

The number of unknowns in Eq. (23) for fault location is three, which are the fault location coefficient, Z_m and Z_n , respectively. As a result, the equivalent impedance values Z_m and Z_n must be calculated ahead of time. However, due to the complexity of the lines in the wind farm collection system, solving the equivalent impedance Z_m and Z_n directly is difficult. As a result, based on the wind farm collection system's unique structural distribution characteristics, this paper proposes a method to solve the equivalent impedance.

During the design of the λ -NIM dictionary, the nodes m and n are viewed as a whole, and the injected fault negative sequence current is I_p . The flow path is shown in Fig. 11.

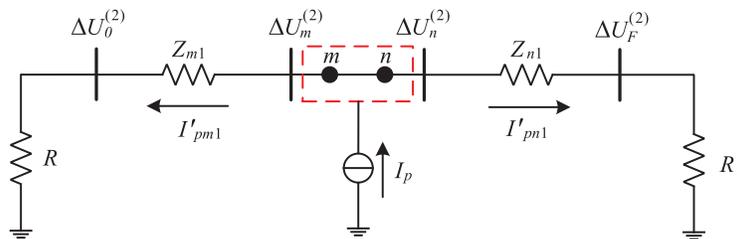


Figure 11: Fault current flow path diagram

As can be seen from Fig. 11:

$$\left\{ \begin{array}{l} \frac{\Delta U_m^{(2)} - \Delta U_0^{(2)}}{I'_{pm1}} = Z_{m1} \\ \frac{\Delta U_0^{(2)}}{R} = I'_{pm1} \\ \frac{\Delta U_n^{(2)} - \Delta U_F^{(2)}}{I'_{pn1}} = Z_{n1} \\ \frac{\Delta U_F^{(2)}}{R_S} = I'_{pn1} \end{array} \right. \quad (24)$$

where $\Delta U_0^{(2)}$ is the fault negative sequence voltage amplitude at the collection bus; $\Delta U_F^{(2)}$ is the fault negative sequence voltage amplitude at the last fan of the collection line; R is the resistance value of the grounding transformer; R_S is the resistance value of the three-phase load; The equivalent impedance is $Z_m = Z_{m1} + R$, $Z_n = Z_{n1} + R_S$.

The values $\Delta U_0^{(2)}$ and $\Delta U_F^{(2)}$ can be obtained after determining the fault interval by measuring the fault negative sequence voltage amplitude at the bus and the last fan of the collection line. As a result, using Eq. (24), the values Z_{m1} and Z_{n1} can be calculated, and the values of equivalent impedance Z_m and Z_n can be obtained. The only unknown variable in the Eq. (23) is the fault location coefficient λ . To create a quadratic equation, square both sides of Eq. (23):

$$\left(\left| \frac{\Delta U_m^{(2)}}{\Delta U_n^{(2)}} \right|^2 \left| \frac{Z_n}{Z_m} \right|^2 - 1 \right) |Z_m|^2 \lambda^2 + \left[\left| \frac{\Delta U_m^{(2)}}{\Delta U_n^{(2)}} \right|^2 \left| \frac{Z_n}{Z_m} \right|^2 |Z_{m1}| |Z_m| + \left| \frac{\Delta U_m^{(2)}}{\Delta U_n^{(2)}} \right|^2 \left| \frac{Z_n}{Z_m} \right|^2 |Z_m|^2 \right] 2\lambda - (|Z_{m1}| + |Z_n|^2) = 0 \quad (25)$$

When Eq. (25) is solved, two solutions can be obtained, with the solution satisfying $0 < \lambda < 1$ being the fault distance. Because only one solution satisfies the condition in general, this method can meet the majority of the positioning requirements in practice.

4.2 Converging System Fault Location Steps

Combined with J-CoSaMP algorithm, the proposed fault location steps of wind farm collection system are as follows:

1. The branch addition method is used to calculate the node impedance matrix and introduce the fault location coefficient based on the topological structure and line parameters of the wind farm collection system. The distribution of the measuring points yields the corresponding λ -NIM dictionary.
2. For the purpose of obtaining the appropriate fault voltage negative sequence component $\Delta U^{(2)}$, read the voltage data of the measuring point before and after the fault at a specific time.
3. The λ -NIM dictionary and $\Delta U^{(2)}$ obtained in Steps 1 and 2 are used as parameters in the J-CoSaMP algorithm to compute the fault negative sequence current component I .
4. If you read $\Delta U^{(2)}$ times in a row, go to Step 5, otherwise go back to Step 3.
5. To obtain the fault negative sequence voltage amplitude of the corresponding node in the fault interval, the reconstruction result I is substituted into Eq. (21).

6. The obtained negative sequence voltage amplitude and system impedance Z_{mn} , Z_m , and Z_n are entered into Eq. (25) to solve, and the solution satisfying $0 < \lambda < 1$ is chosen to locate the fault.

The flow chart is shown in Fig. 12.

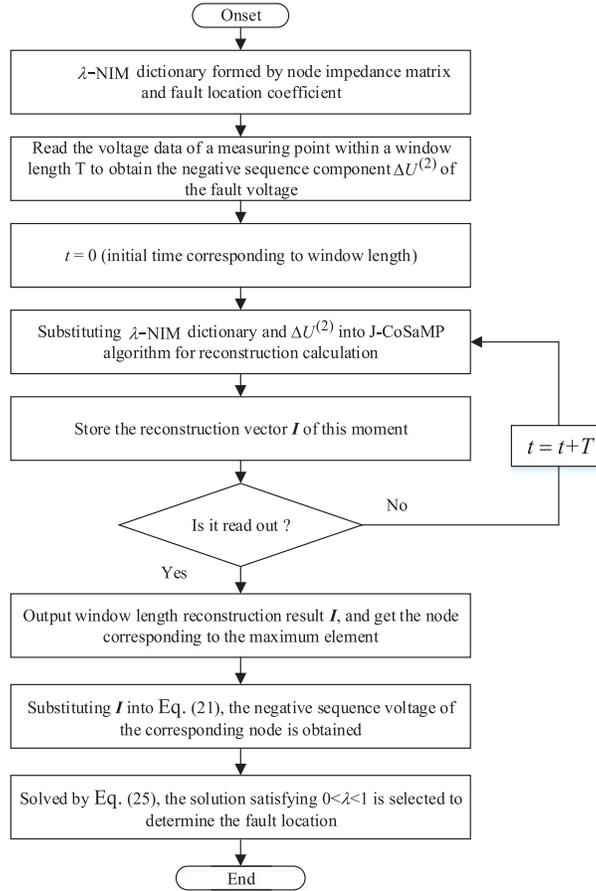


Figure 12: Fault location flow chart

In theory, the location of non-zero elements in the compressed sensing reconstruction result is the location of the fault node. However, it is difficult to ensure that the results of a reconstruction are accurate during the reconstruction process. As a result, in order to reduce the presence of false fault points, this paper reconstructs the same fault multiple times and selects the results with the most times to determine.

5 Experimental and Simulation Verification

5.1 Analysis of Short Circuit Experiment

The 35 kV collection system in a large wind farm with a rated capacity of 504 MW is used as the research object in this paper, and the Simulink simulation platform is built, as illustrated in Fig. 13. The system includes 56 wind turbines. The voltage of each wind turbine is raised to 35 kV by the box transformer. Following cable transmission to the collection bus, it is raised to 220 kV via the main transformer via the transmission line and sent to the power grid.

The short-circuit experiment is performed on the collection system’s 35 kV collection line. As indicated by the k point in Fig. 13, the fault point is located on the collection line between the No. 12 and No. 13 fans. At 5 s, a phase-to-ground fault occurs, and at 5.15 s, the fault is removed. Fig. 14 depicts the three-phase current waveform on the collection line during the fault.

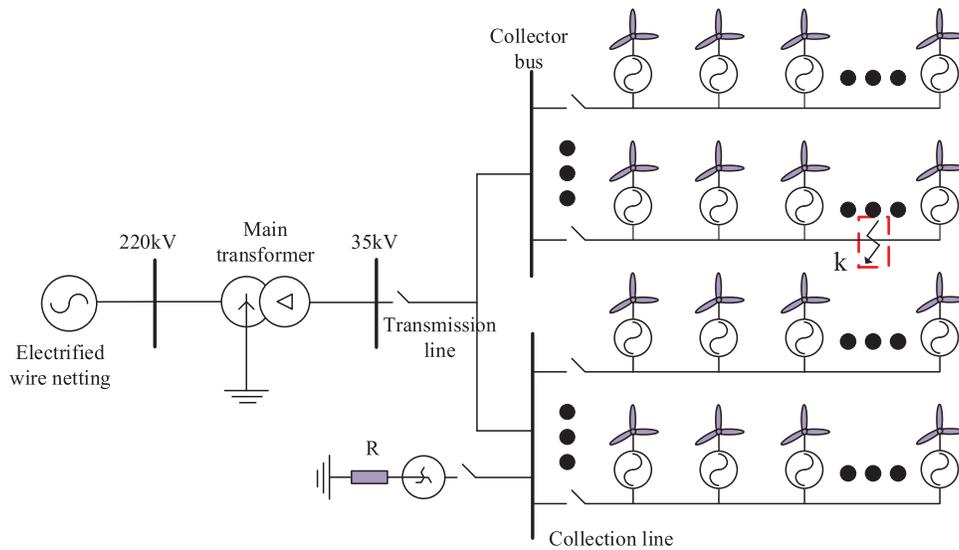


Figure 13: Topology of large wind farm collection system

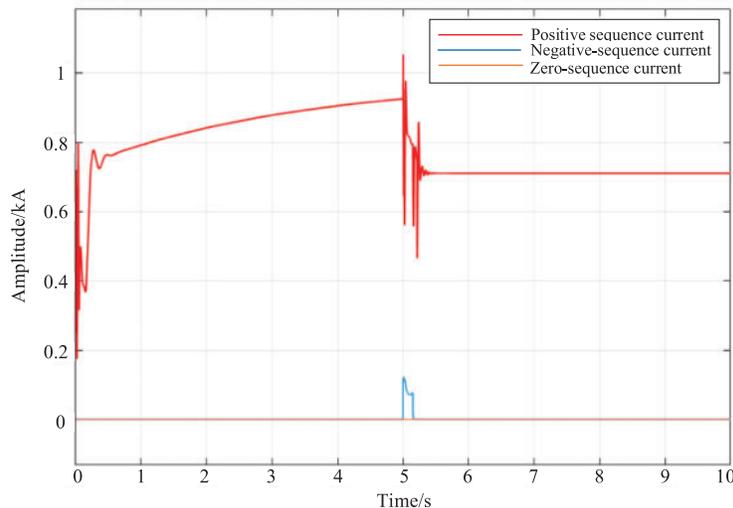


Figure 14: Three phase current waveform of fault line

The examination of the fault current sparse features of the wind farm collection system is validated by comparing Figs. 14 and 15. The fan’s negative sequence current output is 10^{-16} kA, or almost nil, when a single-phase grounding fault on the collecting line happens. In the wind farm collection system’s negative sequence network, the fault point is the only place where the fault negative sequence current is generated.

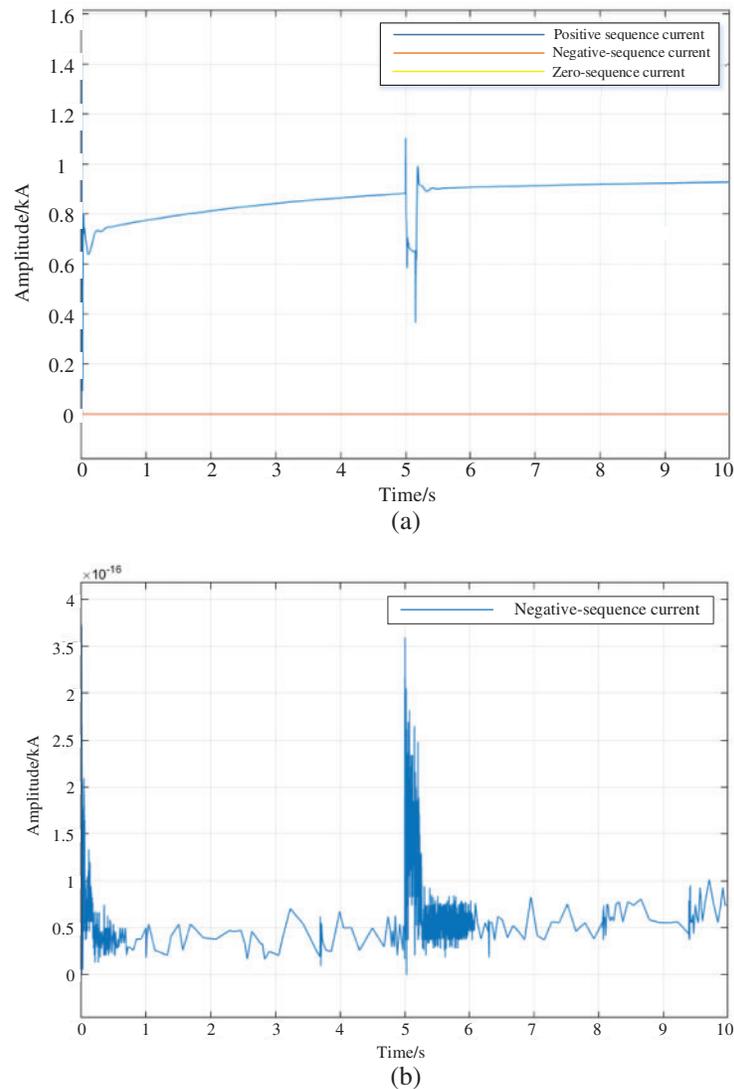


Figure 15: Current waveform of fan when fault occurs. (a) Three phase current waveform sent by the fan in case of fault, (b) Negative sequence current waveform sent by the fan in case of fault

5.2 Simulation Example of New Fault Location Algorithm

5.2.1 Fault Type and Location Impact Analysis

The distributed parameter model is primarily used in this paper, and a series of π -type equivalent circuits are used to simplify the modeling of a collection line in the wind farm collection system. Fig. 16 depicts the equivalent circuit. This circuit consists of series resistance R_1 , inductance L_1 , and grounding capacitance C_1 .

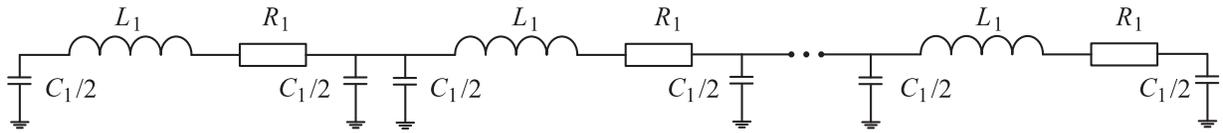


Figure 16: Simplified π unit based modeling

The parameter calculation of π type equivalent circuit is shown in Eq. (26):

$$\begin{cases} R_1 = l_s R_0 \\ L_1 = l_s L_0 \\ C_1 = l_s C_0 \end{cases} \quad (26)$$

where R_0 , L_0 , and C_0 are the resistance, inductance, and capacitance per unit length of the cable line, respectively, and l_s is the length of a single-type equivalent circuit.

When a single-phase ground fault occurs in the wind farm collection system, the negative sequence network structure is depicted in Fig. 17 (the line segments between the nodes represent line impedance and transformer impedance, etc.).

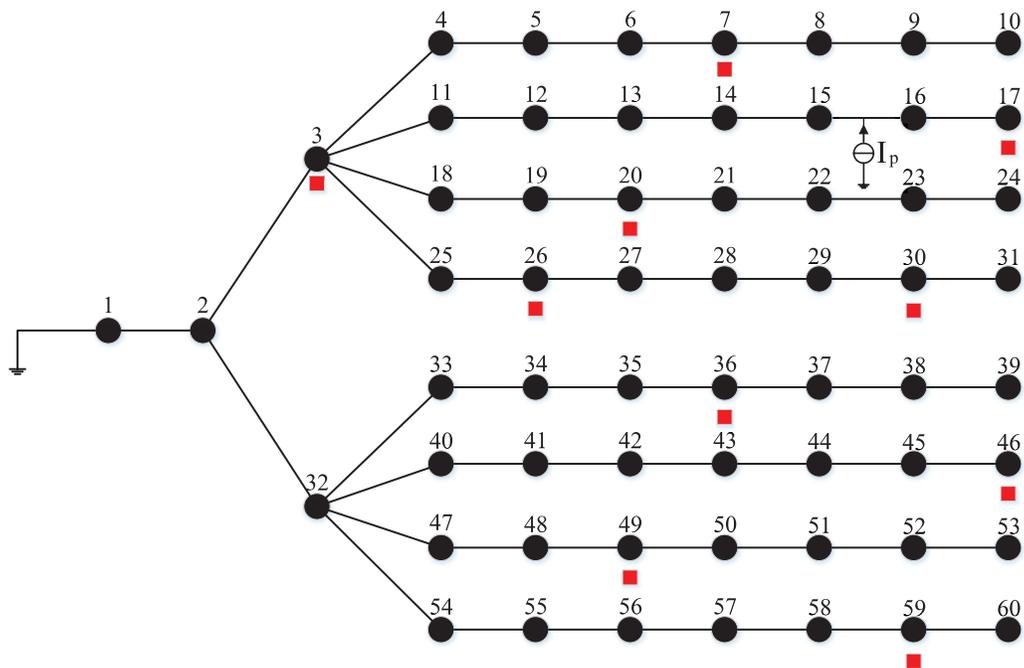


Figure 17: Negative sequence network diagram of collection system

(1) Failure mode

To make the following calculation easier, assume nodes 15 and 16 are 1 km apart. If a single-phase ground fault occurs at a distance of 0.4 km from node 15, the fault location coefficient is 0.4. At the same time, in order to ensure the accuracy of compressed sensing fault location, it is required that there is at least one measurement node in each collection line. Therefore, this paper selects 10 nodes of 3, 7, 17, 20, 26, 30, 36, 46, 49, and 59 in the collection system (60 nodes) as the measurement points to install the voltage measurement device, so that the 10 measurement nodes are evenly distributed in

the wind farm collection system. The voltage data at the ten measuring points following the fault are recorded, and the amplitude of the fault negative sequence voltage is extracted using the phase-locked loop trigonometric function method, as illustrated in [Table 1](#).

Table 1: Fault negative sequence voltage value

Measuring point position	Voltage/kV	Measuring point position	Voltage/kV
3	0.05746	30	0.05129
7	0.05311	36	0.05314
17	0.3898	46	0.05087
20	0.0524	49	0.0525
26	0.0522	59	0.05121

As shown in [Fig. 18](#), the J-CoSaMP algorithm is used to reconstruct the fault negative sequence current component I . The node number is represented by the x -axis, and the amplitude of the fault negative sequence current is represented by the y -axis.

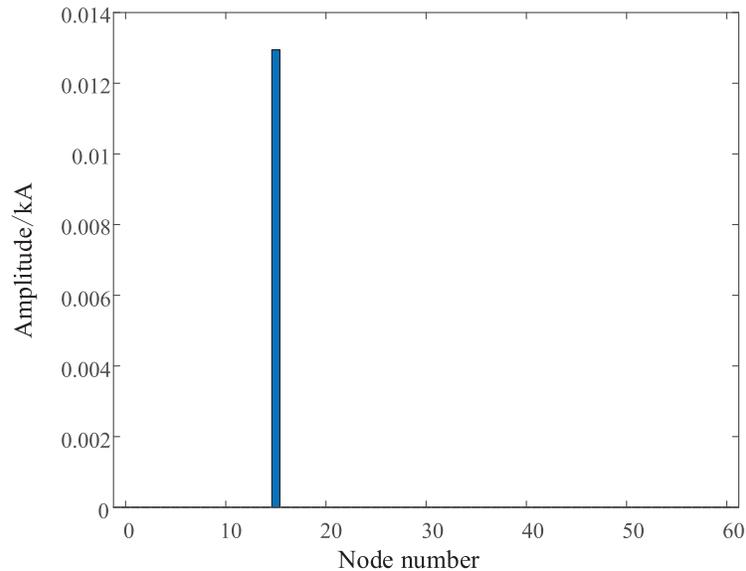


Figure 18: Vector I when nodes 15 and 16 fail A-0 Ω

The voltage amplitude is measured over a time range of 0.25–15.25 ms. 30 results can be obtained when the sampling frequency is set to 2 kHz. [Fig. 19](#) depicts the fault negative sequence current component I at each instantaneous time. The x axis in the figure represents the node number, the y axis time, and the z axis the amplitude of the fault negative sequence current.

[Section 4.1](#) states that after determining the fault occurrence interval, the obtained fault negative sequence current component I is substituted into [Eq. \(21\)](#) to obtain the voltage amplitudes of buses 15 and 16 of 0.3222 kV and 0.3875 kV, respectively. The negative sequence voltage amplitudes Z_m and Z_n are then entered into [Eq. \(25\)](#). The value of λ is 0.3991 according to the conditions for selecting roots in [Section 3.1](#). By analogy, the fault location coefficient is 0.3987 until all 30 results are calculated and

averaged. The fault distance is 0.3987 km, the absolute error is 0.0013 km, and the relative error is 0.325% because the distance between nodes 15 and 16 is 1 km.

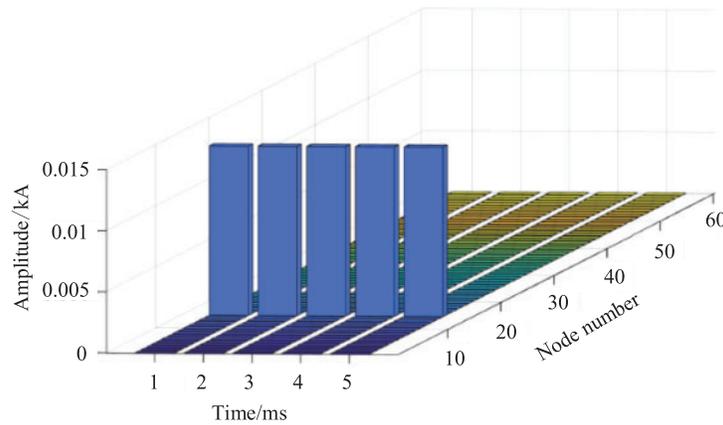


Figure 19: Reconstructed current vector I in each period

(2) Effect of transition resistance

This paper sets five different transition resistances $R_f = 10, 20, 40, 80, 100$ for fault analysis when the fault occurs in order to study the influence of the transition resistance on the fault location of the wind farm collection system. Table 2 summarizes the simulation results. Because the element values in the node impedance matrix are constants, the transition resistance has no effect on them. Both the fault negative sequence current and the fault negative sequence voltage are variables. When the transition resistance varies, the value changes dramatically. As shown in Table 2, the transition resistance has no effect on the fault interval positioning results. Despite having an effect on the ranging results, the relative error is still less than 1%. As a result, the proposed method is less sensitive to transition resistance.

Table 2: Fault location results with transition resistance

Failure mode	R_f/Ω	Abort situation	Location	Positioning accuracy/%	Ranging relative error/%
AG	0	15–16	15–16	100	0.4
	20	15–16	15–16	100	0.65
	40	15–16	15–16	100	0.82
	80	15–16	15–16	100	0.75
	100	15–16	15–16	100	0.5

(3) Noise effect

This paper adds Gaussian white noise after reading the negative sequence voltage value to locate in order to investigate whether noise affects ranging error. Fig. 20 depicts the ranging's relative error.

As shown in Fig. 20, the presence of noise increases the ranging relative error, but most of the errors are less than 1%, and in the actual situation of a wind farm collection system, the two fans are

not far apart. As a result, a small amount of noise in the ranging process has little effect on the ranging results, and the proposed method's anti-noise ability meets the practical application requirements.

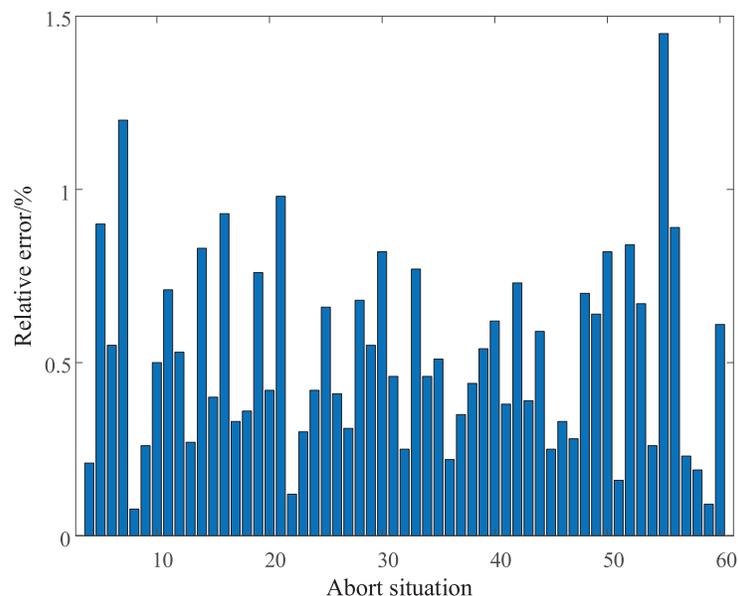


Figure 20: Relative error of each position during fault A-0 Ω

5.2.2 Figure Labels and Captions

(1) Comparative analysis of fault voltage extraction schemes

Because the fault voltage reading value has a large influence on the fault interval positioning accuracy when using compressed sensing to reconstruct the fault negative sequence current. A single-phase ground fault is set between nodes 15 and 16 to see if the phase-locked loop trigonometric function method used in this paper can improve fault location accuracy. The fault voltage is extracted using the instantaneous symmetrical component method and the phase-locked loop trigonometric function method, and the fault interval is located using the J-CoSaMP algorithm. Table 3 displays the simulation results. In this paper, the reconstruction error is defined as the ratio of reconstruction failure times to reconstruction times.

Table 3: Comparison of errors of different extraction schemes

Scheme	Measuring point	Read values/kV	Error ratio/%	Measuring point	Read values/kV	Error ratio/%	Reconstruction error/%
Instantaneous symmetrical component method	3	0.06215	9.1	30	0.05283	7	16.7
	6	0.05669	8.8	35	0.05536	5.8	
	13	0.219	7.4	46	0.05272	7.3	
	17	0.4052	7.6	50	0.0528	5	
	21	0.0541	7.5	59	0.05194	5.3	

(Continued)

Table 3 (continued)

Scheme	Measuring point	Read values/kV	Error ratio/%	Measuring point	Read values/kV	Error ratio/%	Reconstruction error/%
Phase-locked	3	0.05746	0.84	30	0.05129	3.9	6.7
loop trigono-	6	0.05312	1.9	35	0.05336	2	
metric	13	0.2113	3.6	46	0.05087	3.6	
function	17	0.3898	3.5	50	0.0534	6.2	
method	21	0.0521	3.6	59	0.05121	3.8	

Table 3 shows that when using the instantaneous symmetrical component method to extract the fault voltage value, the error with the actual fault voltage value exceeds 5%, and the reconstruction error exceeds 16.7%. When using the phase-locked loop trigonometric function method to extract the fault voltage value, the error is typically less than 4%, and the reconstruction error is reduced to 6.7%. When calculating the fault voltage value, it can be seen that the phase-locked loop trigonometric function method is more accurate than the traditional instantaneous symmetrical component method. Meanwhile, during the fault location process, the phase-locked loop trigonometric function method can reduce reconstruction error and improve positioning accuracy.

(2) Comparative analysis of overcomplete dictionary selection

This paper uses a node impedance matrix and a λ -NIM dictionary to locate fault sections in order to investigate the effect of selecting an over-complete dictionary on fault section location accuracy. In the wind farm collection system, 23 nodes between nodes 4–26 are chosen to set up A-0 Ω grounding faults sequentially, and the reconstruction error is summarized in Fig. 21. Fig. 21 shows that when the node impedance matrix is used as the over-complete dictionary, the reconstruction errors of different fault points are mostly greater than 10%, with the maximum error being 33.3%. When using the λ -NIM dictionary for fault location, the reconstruction errors of various fault points are mostly reduced to less than 10%, demonstrating that using the λ -NIM dictionary for fault location can effectively improve fault location accuracy.

(3) Comparative analysis of reconstruction algorithms

The compressed sensing reconstruction algorithm is enhanced in this paper. To investigate its improvement effect, the OMP algorithm, CoSaMP algorithm, and J-CoSaMP algorithm are used to locate the fault. Fig. 22 depicts the comparison results.

The figure shows that the sparsity K has a significant impact on the algorithm's reconstruction success probability. The reconstruction success probabilities of the three algorithms used in this paper, namely the OMP algorithm, the CoSaMP algorithm, and the J-CoSaMP algorithm, are 97.3%, 90.1%, and 100%, respectively, when the sparsity is 1. When the sparsity is set to 2, the J-CoSaMP algorithm still has a 99.4% chance of successful reconstruction, whereas the other two algorithms have a 90.9% and a 66.3% chance, respectively. The comparison of the three algorithms demonstrates that the proposed algorithm can be applied to various sparsity conditions in the wind farm collection system and can effectively improve fault location accuracy.

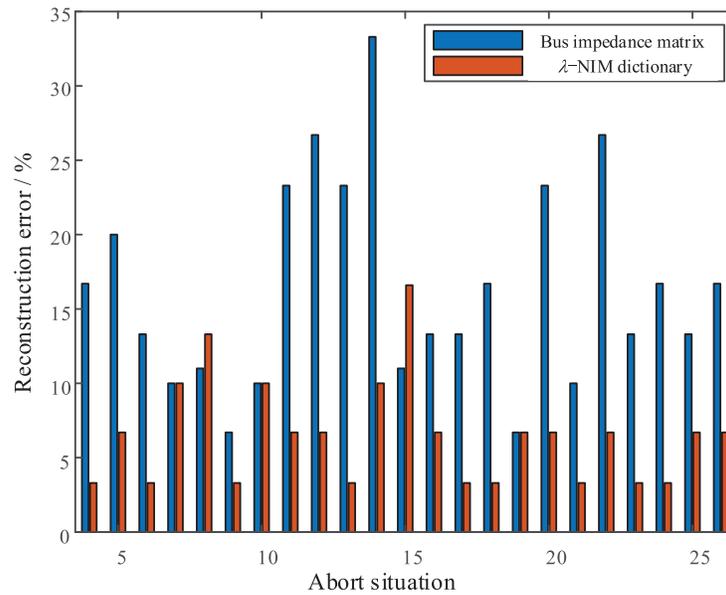


Figure 21: Reconstruction error of each position during fault A-0 Ω

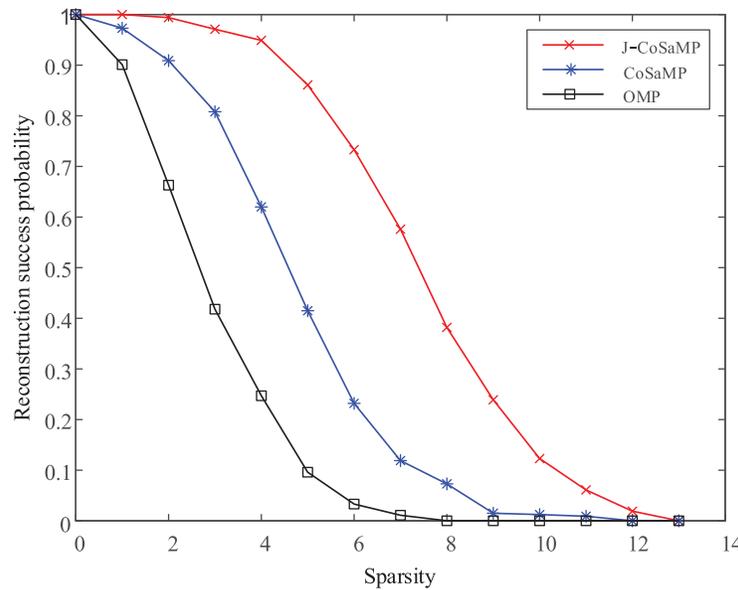


Figure 22: Algorithm reconstruction probability comparison chart

(4) Ablation study

Ablation experiments are included to test the effectiveness of the phase-locked loop trigonometric function method, λ -NIM dictionary, and J-CoSaMP algorithm proposed in this paper. As an example:

a. Phase-locked loop trigonometric function method (PLTFM): The phase-locked loop trigonometric function method is replaced by the instantaneous symmetrical component method in this experiment, and the fault is located by combining the λ -NIM dictionary and the J-CoSaMP algorithm.

b. λ -NIM dictionary: The λ -NIM dictionary is replaced by a node impedance matrix in this experiment, and the fault is located by combining the phase-locked loop trigonometric function method and the J-CoSaMP algorithm.

c. J-CoSaMP algorithm: In this experiment, the J-CoSaMP algorithm is replaced by the CoSaMP algorithm, which is used in conjunction with the phase-locked loop trigonometric function method and the λ -NIM dictionary to locate the fault.

d. Integral experiment (IE): In this experiment, there is no variant, and the fault is found by combining the phase-locked loop trigonometric function method, the λ -NIM dictionary, and the J-CoSaMP algorithm.

To make a fair comparison, using the entire experiment as the reference group, a single-phase grounding fault with the same distance is set between nodes 15–16, 19–20, 34–35, 41–42, and 57–58, and 30 simulation experiments are performed. When performing fault location, all variants use the same parameter settings. Fig. 23 depicts the fault location results.

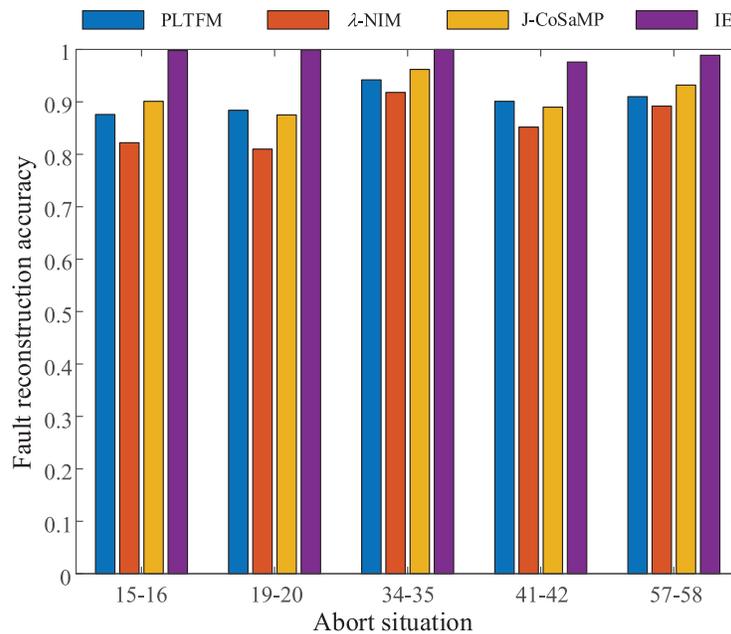


Figure 23: Comparison of performance on different variants

The λ -NIM dictionary has a significant influence on the accuracy of the method in this paper, as shown in Fig. 23 by comparing the three groups of experiments a, b, and c. That is, the λ -NIM dictionary makes the atoms in the observation matrix more matched by combining the fault location coefficient, which can significantly improve the fault location accuracy. Its improvement effect is clearly superior to the other two methods of improvement. Second, by comparing the two groups of a and c experiments, it can be seen that improving the extraction method and improving the algorithm have little difference in improving the proportion of fault location, but in general, the algorithm improvement is slightly larger than the extraction method improvement. As a result of the ablation experiment, the phase-locked loop trigonometric function method, λ -NIM dictionary, and J-CoSaMP algorithm proposed in this paper can effectively improve fault location accuracy, but the λ -NIM dictionary designed in this paper has a clear effect on fault location accuracy.

(5) Comparative analysis of fault location methods

The proposed method, impedance method [5], and traveling wave method [9] are used to locate a single-phase ground fault that is the same distance between nodes 15–16, 34–35, and 41–42. Table 4 displays the location results.

Table 4: Comparison of fault interval location results

Positioning method	Abort situation	Number of accurate positioning	Location deviation	Positioning accuracy/%
Method of this article	15–16	29	1	96.7
	34–35	28	2	93.3
	41–42	30	0	100
Impedance method	15–16	25	5	83.3
	34–35	26	4	86.7
	41–42	27	3	90
Traveling wave method	15–16	23	7	76.7
	34–35	22	8	73.3
	41–42	24	6	80

Table 4 shows that when the impedance method and the traveling wave method are used to locate the fault interval multiple times in the process of interval positioning of wind farm collection system faults, the positioning accuracy is generally below 90%, whereas the accuracy of the positioning method proposed in this paper is mostly above 90%. As a result, the proposed method outperforms the impedance method and the traveling wave method for interval positioning.

To investigate the overall performance of the fault location method proposed in this paper, single-phase ground faults of varying distances are placed between nodes 15–16, and the three methods listed in Table 4 are used to accurately locate faults. Table 5 shows the location results, and Fig. 24 shows the relative error of location.

Table 5: Different fault distance location results

Fault point/kM	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Method of this article	0.0907	0.1965	0.2976	0.4094	0.4911	0.5906	0.6912	0.814	0.91
Impedance method	0.1029	0.1912	0.3088	0.3971	0.5043	0.6129	0.7074	0.8088	0.9831
Traveling wave method	0.0994	0.1984	0.2963	0.4016	0.499	0.5958	0.7035	0.7976	0.8937

According to Table 5 and Fig. 24, the relative error of the fault location method proposed in this paper is mostly less than 1% in the fault location process, whereas the relative error of the impedance method and the traveling wave method is mostly around 2%, with the highest being 9.3%. As a result, the method proposed in this paper is more precise than the other two.

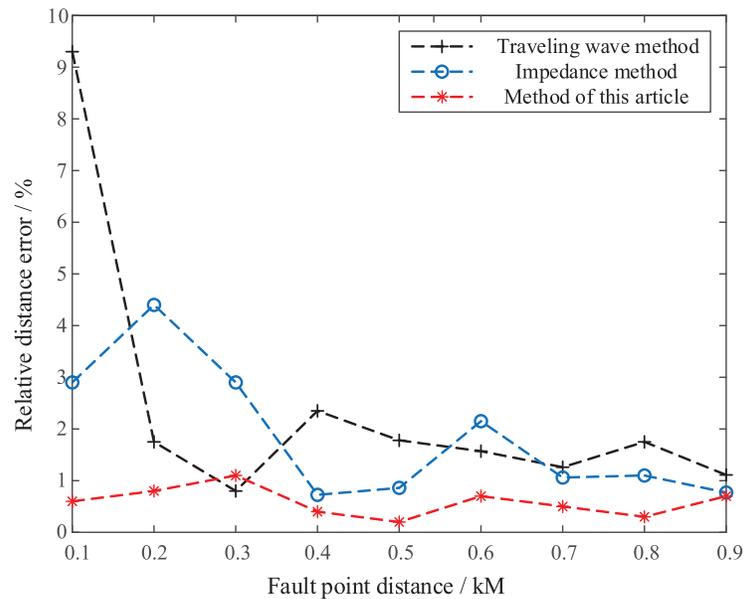


Figure 24: Relative error of different ranging methods

6 Conclusion

To address the difficulty of fault inspection in offshore wind farm collection systems, this paper proposes a fault location method for wind farm collection systems based on the λ -NIM dictionary and the J-CoSaMP algorithm, and simulates, verifies, and analyzes a 60-node wind farm collection system. This paper's main contribution is as follows:

(1) The phase-locked loop trigonometric function method is used to extract the fault negative sequence voltage to form the node negative sequence voltage equation, suppressing the peak phenomenon and narrowing the gap between the read and actual voltage values.

(2) The compressed sensing theory is introduced, and the node impedance matrix is deduced further in the node voltage equation. The fault location coefficient creates the λ -NIM dictionary.

(3) The J-CoSaMP algorithm reconstructs the negative sequence current, and the fault location is realized by combining the voltage and impedance characteristics of the measuring point.

Funding Statement: This work was partly supported by the National Natural Science Foundation of China (52177074).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Ben, K. (2020). Wind energy beijing declaration developing 3 billion of wind power, leading green development and implementing the “30–60” target. *Wind Energy*, 11(5), 34–35.
2. Wang, B., Ren, X. (2021). Single-line-to-ground fault location in wind farm collection line with neutral point grounding with resistor. *Proceedings of the CSEE*, 41(6), 2136–2144. <https://doi.org/10.13334/j.0258-8013.pcsee.200340>

3. Zhai, Y. J., Zhang, K., Zhu, Y. L. (2020). Single-line-to-ground fault location method for wind farm collection line based on segmented impedance matching. *Smart Power*, 48(12), 26–32.
4. Zhang, K., Ma, C. X., Cao, Y. Q. (2020). Fault location based on parameter estimation in hybrid transmission lines of wind farm. *Acta Energiæ Solaris Sinica*, 41(8), 323–329. <https://doi.org/10.19912/j.0254-0096.2020.08.043j.0254-0096.2020.08.043>
5. Pen, H., Zhu, Y. L., Yuan, S. (2020). Combined single-phase ground fault ranging for wind farm collector lines. *Chinese Journal of Scientific Instrument*, 41(9), 88–97. <https://doi.org/10.19650/j.cnki.cjsi.J2005976>
6. Hamidi, R. J., Livani, H. (2017). Traveling-wave-based fault-location algorithm for hybrid multiterminal circuits. *IEEE Transactions on Power Delivery*, 32(1), 135–144. <https://doi.org/10.1109/TPWRD.2016.2589265>
7. Wang, X. D., Wang, Y. C., Liu, Y. C. (2022). Fault location method for collector lines by combining decision coefficient and ESMD-TEO. *Proceedings of the CSU-EPSA*, 34(3), 51–58. <https://doi.org/10.19635/j.cnki.csu-epsa.000845>
8. Zhang, K., Zhu, Y. L., Ma, C. X. (2018). A new method for traveling wave ranging of transmission lines based on optimized traveling wave velocity. *Journal of North China Electric Power University (Natural Science Edition)*, 45(4), 34–40.
9. Peng, H., Wang, W. H., Zhu, Y. L. (2021). Single-phase grounding intelligent ranging of wind farm collector line based on LSTM neural network. *Power System Protection and Control*, 49(16), 60–66. <https://doi.org/10.19783/j.cnki.pspc.201457>
10. He, Q., Pang, Y., Jiang, G., Xie, P. (2021). A spatio-temporal multiscale neural network approach for wind turbine fault diagnosis with imbalanced SCADA data. *IEEE Transactions on Industrial Informatics*, 17(10), 6875–6884. <https://doi.org/10.1109/TII.2020.3041114>
11. Wu, X., Jiang, G., Wang, X., Xie, P., Li, X. (2019). A multi-level-denoising autoencoder approach for wind turbine fault detection. *IEEE Access*, 7, 59376–59387. <https://doi.org/10.1109/ACCESS.2019.2914731>
12. Rezamand, M., Kordestani, M., Carriveau, R., Ting, D. S. (2020). A new hybrid fault detection method for wind turbine blades using recursive PCA and wavelet-based PDF. *IEEE Sensors Journal*, 20(4), 2023–2033. <https://doi.org/10.1109/JSEN.2019.2948997>
13. Yang, L., Zhang, Z. (2021). A conditional convolutional autoencoder-based method for monitoring wind turbine blade breakages. *IEEE Transactions on Industrial Informatics*, 17(9), 6390–6398. <https://doi.org/10.1109/TII.2020.3011441>
14. Dhiman, H. S., Deb, D., Muyeen, S. M., Kamwa, I. (2021). Wind turbine gearbox anomaly detection based on adaptive threshold and twin support vector machines. *IEEE Transactions on Energy Conversion*, 36(4), 3462–3469. <https://doi.org/10.1109/TEC.2021.3075897>
15. Lu, S., Gao, Z., Xu, Q. (2022). Class-imbalance privacy-preserving federated learning for decentralized fault diagnosis with biometric authentication. *IEEE Transactions on Industrial Informatics*, 18(12), 9101–9111.
16. Xu, Q., Lu, S., Jia, W. (2020). Imbalanced fault diagnosis of rotating machinery via multi-domain feature extraction and cost-sensitive learning. *Journal of Intelligent Manufacturing*, 31(6), 1467–1481.
17. Ge, Y. Z. (2007). *The principle and technology of advanced relay protection and fault location*. China: Xi'an Jiaotong University Press.
18. Jia, K., Feng, T., Zhao, Q. J. (2020). Bipolar short-circuit fault location of DC distribution network based on sparse measurement of fault high frequency voltage. *Automation of Electric Power Systems*, 144(1), 142–151.
19. Xu, G., Gong, Q. W., Li, X. (2014). Single-ended fault location method based on atomic decomposition and natural frequency of traveling wave. *Automation of Electric Power Systems*, 34(5), 133–138.
20. Yu, H. N., Ma, Q. Q., Wang, H. (2020). Fault location of VSC-HVDC transmission line based on compressed sensing. *Acta Energiæ Solaris Sinica*, 41(3), 158–166.

21. Baraniuk, R. (2007). A lecture on compressive sensing. *IEEE Signal Processing Magazine*, 24(4), 118–121.
22. Lu, C., Liu, Y. H. (2015). RIP criterion in compressive sensing theory. *Automation and Instruments*, 190(8), 1001–9227. <https://doi.org/10.14016/j.cnki.1001-9227.2015.08.211>
23. Zhang, S. N., Luo, H. Y., Cheng, X. X. (2020). Realization of positive & negative sequence component picking up and phase locked loop using FPGA based on dual-dq transform. *High Voltage Apparatus*, 56(3), 182–189. <https://doi.org/10.13296/j.1001-1609.hva.2020.03.027>
24. Wang, K., Wu, Y. M., Zhuge, J. C. (2019). MsGOMP infrared image denoising algorithm based on generalized jaccard coefficient. *Infrared Technology*, 41(6), 577–584.
25. Li, Z. X., Tian, B., Li, Z. H. (2016). Two-terminal nonsynchronized fault location algorithms for single/double transmission lines. *Automation of Electric Power Systems*, 40(22), 105–110.