

## Research Article

# Research on a Location Method for Complex Voltage Sag Sources Based on Random Matrix Theory

Lizhen Wu , Yongnian Zhang, Xiaohong Hao, and Wei Chen

*College of Electrical and Information Engineering of Lanzhou University of Technology, Lanzhou, Gansu 730050, China*

Correspondence should be addressed to Lizhen Wu; [wulzlut@163.com](mailto:wulzlut@163.com)

Received 2 January 2020; Revised 18 March 2020; Accepted 17 April 2020; Published 28 May 2020

Academic Editor: Pietro Bia

Copyright © 2020 Lizhen Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The precise location of voltage sag sources plays an important role in formulating a voltage sag governance plan and clarifying the responsibility for the accident. The traditional location method of voltage sag sources is difficult to establish a precise mathematical model and locate the complex voltage sag sources accurately in the complex distribution network. In order to solve these problems, a location method for complex voltage sag sources based on random matrix theory is proposed in this paper. Firstly, the augmented matrix is constructed based on the influence factor data and the operation state data of the distribution network as the data source matrix, and the statistical characteristics of each data are analyzed by using the random matrix theory to determine the suspicious area of voltage sag source. Then, the disturbance signal of each node in the suspicious area of voltage sag source is analyzed by using the atomic algorithm and disturbance active power method, and it can determine the location of each disturbance source in the complex voltage sag event and the cause of voltage sag at each node. Compared with the existing model-based methods, the proposed method in this paper is a data-driven approach, which does not need the physical model and topology information. Furthermore, it can reduce the amount of data and improve the analysis efficiency on the basis of determining the suspicious area of the voltage sag source. Finally, the examples are given to show that the proposed method can accurately locate the disturbance sources in complex voltage sag events.

## 1. Introduction

In the distribution network, the power quality problems caused by voltage sag occupy the majority, which are difficult to avoid and do great harm. With a large number of renewable energies, electric vehicles, and other flexible loads connected to the distribution network, the structure of the distribution network becomes more complex, and the factors affecting voltage sag become more complex. As a result, the consequence of voltage sag is more and more serious, and the range of influence is also larger and larger. If the problem of voltage sag cannot be dealt with timely and accurately, it will develop into a serious accident, which will lead to the continuous deterioration of the distribution network and even large-scale failure [1]. So, one of the most concerned problems of power system operators is how to diagnose the distribution network fault quickly and accurately. Unfortunately, the complex voltage sag events are a whole of single voltage sag events with causal spatiotemporal

correlation at different locations and times, and these voltage sag processes overlap each other to form a complex voltage sag process. These reasons make it more difficult to locate the voltage sag sources [2]. Therefore, it is necessary to analyze and locate the origin event sources and the intermediate event sources when locating the complex voltage sag sources, which can provide theoretical support for the whole process of a comprehensive restoration of complex voltage sag events, and it has guiding value for the analysis of event consequences and event accountability.

At present, a lot of research works on the location of voltage sag sources have been carried out, which is mainly based on physical characteristics, including features' extraction and pattern recognition. The extraction of characteristics is to transform and reconstruct voltage sag signal through signal processing and mathematical statistics, including Fourier transform, Hilbert Huang transform, S transform, and wavelet transform [3]. Karimi et al. [3] suggested a method to analyze the voltage sag signal that is

based on the frequency-domain characteristics of wavelet transform and singularity detection theory, which can reduce the computational complexity. However, it has the worse ability to resist noise, and its accuracy is greatly affected by harmonic, noise, and other interferences. Zhao et al. [4] proposed a method that establishes the location state estimation model by estimating fault location and fault resistance. However, it is difficult to establish an accurate and general mathematical model for complex time series data in the complex power grid. In [5], two new methodologies are introduced: improved distance relay (IDR) and improved current based method (ICBM). These methodologies establish the evaluation index by using the threshold current, saturation current, and positive and negative power factor of the inverter of the distributed generation unit, which can well determine the position of voltage sag sources in the complex distribution network including the distributed generation, but this method still depends on the accuracy of the network structure model. At the same time, the process of characteristics' extraction will result in information loss, and the overcomplexity of the classification model also makes the shortcomings of existing methods increasingly prominent. The pattern recognition methods use the classification algorithm to design a classifier, which is used to determine the voltage sag source category of the disturbance signal, mainly including support vector machine, high-order spectral analysis [6], and neural network [7]. In [6], the classification of power quality disturbances is researched by using high-order spectral analysis. Reference [7] carried out a self-monitoring voltage sag source identification method based on the convolution neural network (CNN), and this method can not only identify the voltage sag source but also accurately determine the specific fault phase. With the development of artificial intelligence technology, the methods based on artificial intelligence and machine learning are gradually applied to the research of voltage sag source location. However, these methods have many problems, such as large demand data and a high-order model, and so on. In particular, these methods cannot deal with complex voltage sag events well. To solve the problem of complex voltage sag sources locating, [8] proposed a location method of voltage sag sources based on the division of disturbance regions, which can accurately locate the disturbance sources of complex voltage sag events and has a strong fault-tolerant performance. However, this method needs to analyze data of all of the detection points, and the process is complex.

In recent years, big data technology has been widely concerned. With the popularity of wide-area management system (WAMS) and phasor measurement unit (PMU) in the distribution network system, a large amount of information (including correct alarm information, wrong transmission information, repeated information, and irrelevant information) is obtained, which provides a reliable data for the application of big data technology in power system fault diagnosis. Also, some big data analysis methods are proposed for voltage sag source recognition. As a big data analysis method, random matrix theory (RMT) does not rely on the simplification and assumption of the physical model

of the system, because it can extract effective information from a large number of data. Recently, it has achieved some results in a power system fault diagnosis. Considering that the occurrence of voltage sag is a Poisson process, [9] proposed a mathematical model describing the voltage sag process as a random process, which proves the randomness of the voltage sag process. It shows that the RMT method can effectively analyze the voltage sag process. In [10], the mean spectral radius evaluation index is established by using the single ring law in random matrix theory, and the operation state of the system is analyzed. Reference [11] investigated an evaluation method to identify the weak points of power grid accurately from a data-driven perspective, which is combined with high-dimensional random matrix theory and entropy theory. These methods are used only to analyze whether there is a fault state in the power system, but how to locate the fault location is not involved. A fault time determination method based on the mean spectral radius index and a fault area localization based on data correlation analysis are proposed in [12]; they can effectively locate the fault location, but the cause of each fault source cannot be determined for complex fault cases.

Therefore, this paper proposes a new location method of complex voltage sag sources based on random matrix theory. Firstly, the fault state is identified by the random matrix theory, the augmented matrix is established by using the state data matrix and the influence factor matrix, and the potential area of voltage sag source is determined by using the mean spectral radius difference as the index. Then, according to the starting and ending time of basic events, the complex voltage sags are decomposed by the atomic algorithm, and the whole complex voltage sags are divided into segments in time sequence. According to the polarity of the disturbance energy in each segment, the disturbance event source in each segment is located to locate the voltage sag source accurately. Finally, the feasibility and correctness of the proposed method are verified by the simulation data of the IEEE 39-bus system through the PSCAD simulation platform. The results show that the random matrix theory can accurately determine the location of complex voltage sag sources. It should be pointed out that this method can help operators to determine the location of voltage sag source timely and accurately and can also explain the causes of voltage sag events of each node according to the location results of segments.

The paper is organized as follows. Section 2 discusses the random matrix theory and big data processing method. The voltage sag source location method based on PMU data correlation analysis is presented in Section 3, including the method of determining the suspicious areas of voltage sag source and the location method of complex voltage sag sources. Section 4 includes simulation results. Finally, the paper is concluded in Section 5.

## 2. Random Matrix Theory and Big Data Processing Method

*2.1. Random Matrix Theory.* Random matrix theory is a mathematical tool that can perform statistical analysis of

complex systems. The RMT can reflect the fluctuation characteristics of data by analyzing the correlation between data and can also map the physical system state through data characteristics. When the dimension of data is in tens or hundreds, some properties of random matrix are still convergent and have considerable accuracy, which provides a possibility for the random matrix theory to be used in practical engineering problems [13, 14]. When the number of rows and columns of the random matrix tends to infinity, and the ratio of the number of rows to the number of columns tends to constant, the empirical spectral distribution (ESD) function of the random matrix will have certain characteristics, such as semicircle law, full-circle law, single ring theorem, and Marchenko-Pastur (M-P) law.

The ESD function is often used to characterize the distribution of the eigenvalues of the random matrix in random matrix theory. For the  $p$  order matrix  $A$ , the ESD function is as follows [15]:

$$F^A(x) = \frac{1}{p} \sum_{i=1}^p I(\lambda_i \leq x), \quad (1)$$

where  $I(\cdot \leq \cdot)$  is an indicative function. If the inequality in parentheses holds, the function takes a value of 1; otherwise, it takes a value of 0.

**2.1.1. Single Ring Theorem.** Assume that  $X = \{x_{i,j}\}$  is a  $p \times n$  random matrix with non-Hermitian features, and it consists of independent random variables, where the mean  $E(x_{i,j}) = 0$  and the variance  $\delta^2(x_{i,j}) = 1$ . The matrix  $Z = \{z_{i,j}\}$  is obtained by quadrature of  $L$  the random matrix  $X$ , the matrix  $Z$  is subjected to standardization processing to obtain a standard matrix  $\bar{Z} = \{\bar{z}_{i,j}\}$ , and each element satisfies with the mean  $E(z_{i,j}) = 0$  and the variance  $\delta^2(z_{i,j}) = 1/N$ . When  $p, n \rightarrow \infty$  and the ratio  $y = p/n \in (0, 1]$ , the ESD function of matrix  $\bar{Z}$  satisfies the single ring theorem, and its probability density function is expressed as follows [16]:

$$f_{\text{ESD}}(\lambda) = \begin{cases} \frac{1}{\pi y L} |\lambda_i|^{(L/2)-2} & (1-y)^{L/2} \leq |\lambda_i| \leq 1, \\ 0, & \text{others,} \end{cases} \quad (2)$$

where  $\lambda_i$  is the eigenvalues of matrix  $\bar{Z}$ . According to the characteristics of the single ring theorem, when the system is in a stable state, the eigenvalues of the matrix  $\bar{Z}$  are roughly distributed in a ring with an outer ring radius of 1 and an inner ring radius of  $(1-y)^{L/2}$ .

When the system is stable, the effect diagram of the single ring theorem is shown in Figure 1.

**2.1.2. M-P Law.** The sample covariance matrix is a statistic that is more commonly used in mathematical statistics and other areas of analysis system status. The M-P law can describe the distribution of eigenvalues of the large-dimensional covariance matrix, and it has strong applicability. A  $p \times n$  random matrix  $X = \{x_{i,j}\}$ , where the mean  $E(x_{i,j}) = 0$  and the variance  $\delta^2(x_{i,j}) < \infty$ . When  $p, n \rightarrow \infty$ ,  $y = p/n \in (0, 1]$ , the ESD of the sample covariance matrix

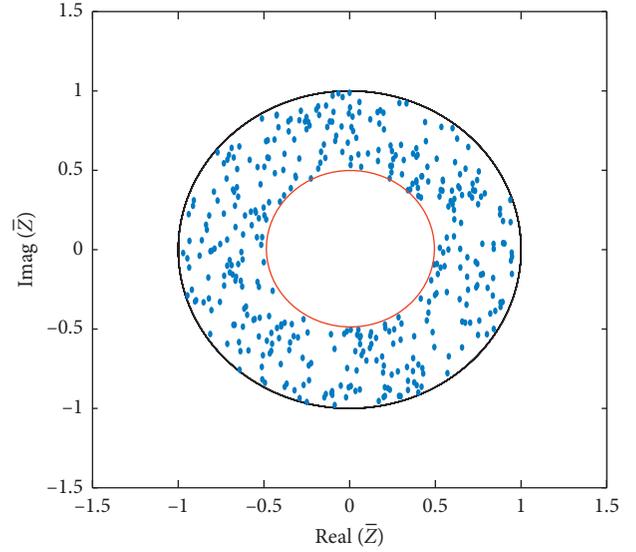


FIGURE 1: Effect diagram of the single ring theorem.

$S = (1/N)ZZ^H \in C^{p \times p}$  tends to the M-P law, and its probability density function is as follows [17]:

$$f_{\text{ESD}}(\lambda) = \begin{cases} \frac{1}{2\pi\lambda y \delta^2} \sqrt{(b-\lambda)(\lambda-a)}, & a < \lambda < b, \\ 0, & \text{others,} \end{cases} \quad (3)$$

where  $a = \delta^2(1 - \sqrt{y})^2$  and  $b = \delta^2(1 + \sqrt{y})^2$ .  $\lambda$  is the covariance matrix eigenvalue.

When the system is stable, the effect drawing of the M-P law is shown in Figure 2.

**2.1.3. Linear Eigenvalue Statistic.** Linear eigenvalue statistic (LES) can describe the distribution of eigenvalues of the random matrix. For the random matrix  $X$ , its linear eigenvalue statistic is defined as follows [18]:

$$N_p(f) = \sum_{i=1}^p f(\lambda_i), \quad (4)$$

where  $\lambda_i (i = 1, 2, \dots, p)$  is the eigenvalues of matrix  $X$ , and  $f(\cdot)$  is a test function. Different test functions result in different linear characteristic statistics.

Mean spectral radius (MSR) is a linear eigenvalue statistic commonly used in random matrix theory. The distance between the eigenvalue and the origin in the complex plane is used to reflect the distribution of the eigenvalues of the random matrix. It is defined as follows [19]:

$$L_{\text{MSR}} = \frac{1}{p} \sum_{i=1}^p |\lambda_i|, \quad (5)$$

where  $L_{\text{MSR}}$  is the value of MSR of the random matrix, and  $\lambda_i (i = 1, 2, \dots, p)$  is the eigenvalue root of the matrix. The operating state of the system is characterized by comparing the limits of the spectral radius between  $L_{\text{MSR}}$  and the normal case.

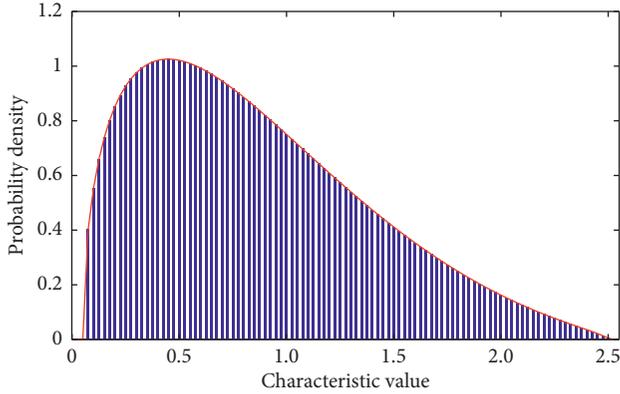


FIGURE 2: Effect diagram of M-P law.

## 2.2. Data Processing and the Random Matrix Establishing.

The PMU can acquire massive real-time operation data with spatiotemporal characteristics. Assume that there are  $n$  nodes in the distribution network system, all nodes are equipped with PMUs, and each PMU can obtain  $m$  state variables. At the sampling time  $t_i$ , the measurement data of PMU can form a column vector which is formed as

$$x(t_i) = [x_{1,1}, \dots, x_{1,m}, x_{2,1}, \dots, x_{n,1}, \dots, x_{n,m}]^T. \quad (6)$$

Arrange the measurement data of each sampling time into a matrix according to the time series, as shown in formula (7).

$$X = [x(t_1), x(t_2), \dots, x(t_i), \dots]. \quad (7)$$

In order to realize the real-time analysis of the data, the PMU data is collected by the real-time separation window method. The raw data matrix  $\hat{X}$  is obtained from the data matrix  $X$  by the real-time separation window, and the window size is  $N \times T$ . After each sampling, the separation window is moved backward by a one time point, so as to realize the dynamic process of monitoring state variables. Since the different state variables of the monitoring data have different measurement units, and the data does not meet the requirement of the single ring theorem and the M-P law, it is necessary to standardize the raw data matrix  $\hat{X}$ . The  $N \times T$  non-Hermitian standard matrix  $\bar{X}$  is obtained by equation (8) [20].

$$\bar{x}_{i,j} = \frac{\delta(\bar{x}_i)}{\delta(\hat{x}_i)} [\hat{x}_{i,j} - E(\hat{x}_i)] + E(\bar{x}_i), \quad (8)$$

where  $i = 1, 2, \dots, N, j = 1, 2, \dots, T$ , and  $\hat{x}_i = (\hat{x}_{i,1}, \hat{x}_{i,2}, \dots, \hat{x}_{i,t})$ .  $\delta(\bar{x}_i)$  and  $E(\bar{x}_i)$  are the standard deviation and the mean of the standard matrix  $\bar{X}$ , respectively, and  $\delta(\bar{x}_i) = 1$ , and  $E(\bar{x}_i) = 0$ .  $\delta(\hat{x}_i)$  and  $E(\hat{x}_i)$  are the standard deviation and the mean of the raw data matrix  $\hat{X}$ , respectively.

The singular value equivalent matrix of the standard matrix  $\bar{X}$  is

$$\bar{X}_u = U \sqrt{\bar{X} \bar{X}^T}, \quad (9)$$

where  $U$  is a unitary matrix satisfying Haar distribution, and  $\bar{X}_u \bar{X}_u^T = \bar{X} \bar{X}^T$ . After performing the above processing,  $L$ -

independent standard non-Hermitian matrices  $\bar{X}_{u,i}$  ( $i = 1, 2, \dots, L$ ) can be obtained, and the matrix product  $Z$  is obtained as follows:

$$Z = \prod_{i=1}^L \bar{X}_{u,i} \in C^{N \times N}. \quad (10)$$

The matrix  $\bar{Z}$  can be obtained according to

$$\bar{z}_i = \frac{z_i}{\delta(z_i) \sqrt{N}}, \quad (11)$$

where  $z_i = (z_{i,1}, z_{i,2}, \dots, z_{i,N})$ ,  $\bar{z}_i = (\bar{z}_{i,1}, \bar{z}_{i,2}, \dots, \bar{z}_{i,N})$ .  $\delta(z_i)$  is the standard deviation of the matrix  $Z$ .

In addition, if the intercepted sample data has nonstationary conditions, the data should be differentiated and meet the requirements of stability.

## 3. Voltage Sag Source Location Method Based on PMU Data Correlation Analysis

**3.1. Establishing Augmented Matrix.** In order to accurately locate the voltage sag sources, it is necessary to analyze the correlation between the distribution network state data and the influencing factors of each node. Utilizing the distribution network state data as the basic matrix and taking the influencing factors of each node as the augmentation part, the state augmentation matrix is constructed together with the basic matrix and the augmentation part. The internal relationship between the distribution network operation state and the influencing factors of each node is explored by analyzing the augmented matrix. The operation state data can be composed of a variety of distribution network state data, such as node voltage amplitude and phase angle, injected active power and reactive power, and branch current. Data of influencing factors include the output of distributed power supply, node load, temperature, humidity, weather, and other data. Because the proposed method needs to determine the position of voltage sag source by data correlation, the data of influencing factors should have only the main measurements associated with voltage sags, such as the residual voltage, the time the voltage sags occurred, and their duration, and so on [21].

Assume that there are  $m$  nodes in the monitoring data of the distribution network.  $m_c$  was selected as the state variable and  $m_f$  as the influencing factors of each node. After  $t$  times of real-time separation window acquisition, the state variable data constitutes a basic state matrix  $B_c \in C^{(m \times m_f) \times T}$ , and the influence factor data constitutes an influencing factor matrix  $B_f \in C^{(m \times m_f) \times T}$ . In practice, the number of state variables is far greater than the number of influencing factors; that is  $m_c \gg m_f$ .

For the purpose of reducing the influences of repeated data in the influencing factors matrix, it is necessary to process the matrix of influencing factors as shown in equation (12) [19]:

$$N_f = B_f + g \times N, \quad (12)$$

where  $N_f$  is the influencing factors matrix after processing.  $N$  is the random noise matrix, and  $N \sim N(0, 1)$ .  $g$  is the magnitude of the random noise.

It should be noted that the value of  $g$  should not be too large or too small, and if its value is too large, this will affect the accuracy of the factor matrix; if the value is too small, this will reduce the correlation of repeated data. Therefore, it is essential to determine a suitable signal-to-noise ratio to get an appropriate  $g$ , and so as to ensure the consistency of the results, the signal-to-noise ratio should take a constant value. The definition of SNR can be expressed as follows [11]:

$$SNR = \frac{\min(|N_{f(i,j)}|)}{g}, \quad (13)$$

where SNR represents the signal-to-noise ratio.  $N_{f(i,j)}$  ( $i = 1, 2, \dots (m \times m_c)$ .  $j = 1, 2, \dots T$ ) is an element of the factor matrix.

The state data matrix and influencing factors matrix are constructed as state augmented matrix, which can be represented as follows:

$$A = \begin{bmatrix} B_c \\ N_f \end{bmatrix}. \quad (14)$$

In addition, the reference state augmentation matrix is constructed by the state data matrix and the random noise matrix, and the influence of the correlation between the data can be eliminated by comparing it with the state augmentation matrix and the state data matrix. The reference state augmentation matrix is formed as

$$\tilde{A} = \begin{bmatrix} B_c \\ N \end{bmatrix}. \quad (15)$$

**3.2. The Method to Determine the Suspicious Areas of Voltage Sag Source.** For the sake of locating the voltage sag sources, it is necessary to analyze the correlation between the influencing factors of each node and the system state matrix. The MSR is used as the index of correlation analysis for matrix data. In order to eliminate the interference caused by the repeated data in the state data and the influencing factors, the difference of the MSR of the augmented matrix is defined as an index, and the formula is shown in

$$d_{MSR}(t) = L_{MSR,A}(t) - L_{MSR,\tilde{A}}(t), \quad (16)$$

where  $k = m(m_c + m_f)$ .  $L_{MSR,A}(t)$  and  $L_{MSR,\tilde{A}}(t)$  are the value of MSR of the state augmentation matrix  $A$  and the reference augmentation matrix  $\tilde{A}$ , respectively.  $d_{MSR}(t)$  can determine whether the corresponding influencing factors are related to the system voltage disturbance at this time.

The integral of the  $d_{MSR}$  in the time scale of real-time separation window is defined as

$$Q_{MSR}(t) = \int_{t_0}^t d_{MSR}(t)dt, \quad (17)$$

where  $t_0$  and  $t$  are the starting and ending time of the real-time separation window, respectively. The  $Q_{MSR}(t)$  can be used to characterize the degree of correlation between different influencing factors and the state data.

When the distribution network operates stably, the critical value of the difference of mean spectral radius is  $d_{MSR,0}(t)$  ( $d_{MSR,0}(t)$  is a reference value, which is determined by realities). If  $d_{MSR}(t) > d_{MSR,0}(t)$ , the influence factor of the node is related to the system disturbance; then, all the values of  $Q_{MSR}(t)$  are sorted from big to small. If the value of  $Q_{MSR}(t)$  is larger, the corresponding buses or lines are more likely to be the area where voltage sag sources exist.

### 3.3. The Location Method of Complex Voltage Sag Sources.

In order to accurately locate the position of the complex voltage sag sources, the disturbance data in the fault area is analyzed by the atomic algorithm and the method of disturbance active power (DP) and the disturbance energy (DE). The analysis process mainly includes two parts: the basic voltage sag event segmentation method and location method of disturbance event source in each segment.

**3.3.1. Basic Voltage Sag Event Segmentation Method.** The complex voltage sag disturbance signal contains multiple basic events, and there are chain-type relationships, development-type relationships, parallel-type relationships, or multiple mixed-type relationships between basic events. The atomic decomposition algorithm has a good application in dealing with this composite, overlapping, segmented, and mixed data signal [22]. In the atomic decomposition algorithm, a complete atomic database is constructed. According to the characteristics of the signal, a set of best-matching atoms are adaptively selected in the highly redundant atomic database by search algorithm, and the signal is represented by a set of best-matching atoms to achieve sparse signal decomposition. The matching pursuit algorithm is commonly used in signal sparse decomposition [23].

Assuming that  $f$  is the analyzed signal,  $D = \{g_r\}_\Gamma$  is a complete atomic database,  $\Gamma$  is the set of values of  $r$ , and all atoms are normalized; that is  $\|g_r\| = 1$ . After the  $j$ th decomposition, the search algorithm finds the best matching atom  $g_{r,(j-1)}$  from the complete atomic database with the original signal or the current residual signal, and  $g_{r,(j-1)}$  should satisfy

$$|\langle f, g_{r,(j-1)} \rangle| = \max_{r \in \Gamma} |\langle f, g_r \rangle|, \quad (18)$$

where  $\langle f, g_r \rangle$  is the inner product of  $f$  and  $g_r$ .

The best matching atom  $g_{r,(j-1)}$  is extracted to get the residual signal  $R_f^{j-1}$  after each iteration process, and the residual signal is decomposed into a new residual component  $R_f^j$ . The iteration relationship is shown in

$$R_f^{j-1} = \langle R_f^{j-1}, g_{r,(j-1)} \rangle g_{r,(j-1)} + R_f^j. \quad (19)$$

After  $n$  iterations, the residual signal  $R_f^n$  is ignored, and the original signal  $f$  can be represented as a linearly expanded form of  $n$  atoms.

$$f = \sum_{j=0}^{n-1} \langle R_f^j, g_{r,j} \rangle g_{r,j}. \quad (20)$$

In each iteration process of the matching pursuit algorithm, a matching atom closest to the signal which is to be decomposed can be obtained by searching all the atoms in the complete atomic database. Usually, the main components of the signal can be represented by a few matching atoms. The flow chart of the basic event detection for complex voltage sag events using the atomic decomposition algorithm is shown in Figure 3.

Through signal decomposition, the correlation and causal relationship of each basic event in the data record of the voltage sag event can be obtained. According to the starting and ending time of each basic event, the basic voltage sag event segmentation method is divided into the following steps:

*Step 1.* The basic events detected are represented in the phase-time diagram.

*Step 2.* According to the distribution of the occurrence time of each of the basic events in the three-phase signal data, search for the starting and ending time of each disturbance event, and establish a three-phase starting and ending time set.

*Step 3.* The starting time and ending time of each phase are combined to establish a set of critical moments. Because there will be some error in the process of basic event detection, if the difference between the values of multiple starting and ending times is too small, it can be regarded as the same time.

*Step 4.* The three-phase voltage waveform records are divided into segments by using the critical values in the three-phase critical time set.

For a simple voltage sag event, the relationship between the events caused by the voltage sag source is relatively simple, so only one basic event is included in the segment. However, there are multiple basic events in the segment for complex voltage sag events.

**3.3.2. Location Method of Disturbance Event Source in Each Segment.** After dividing the basic segments of the complex voltage sag event, firstly, the direction and location of the basic event sources are determined based on the method of DP and DE within each segment. Then, the circuit in which voltage sag events occur is regarded as a result of the interaction of power sources and each disturbance event sources in the system. And the superposition principle is used to analyze each segment and determine the location of each basic disturbance event source relative to the monitoring point. According to the analysis of each section, the location of the disturbance source of the most serious basic events affecting the voltage and current of the monitoring point is determined. And the complete process of complex voltage sag events can be recorded according to the result of the location of the voltage sag source in the segment [24].

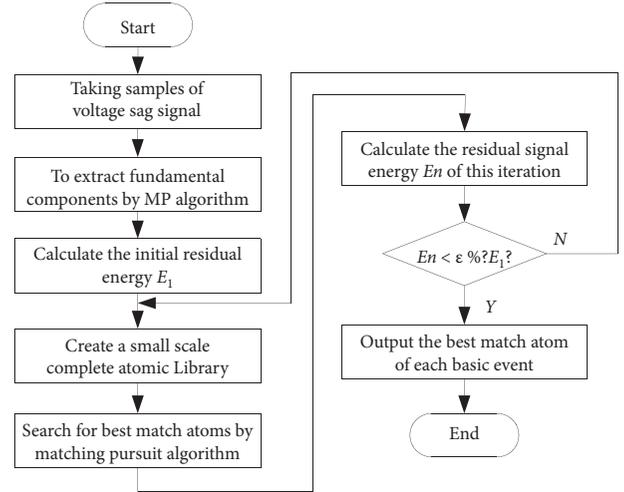


FIGURE 3: The flow chart of basic event detection.

When the voltage sag event occurs, the phase voltage or current of each monitoring point in the suspicious area of the voltage sag source can be divided into the sum of the steady-state component and the disturbance component. The voltage and current components generated by the voltage sag source in the  $i$ th segment can be obtained as shown in formula (21), respectively:

$$\begin{aligned} u_{d,i} &= u_i - u_{p,i} \\ i_{d,i} &= i_i - i_{p,i} \end{aligned} \quad (21)$$

where  $u_{p,i}$  and  $i_{p,i}$  represent the voltage and current steady-state components produced by all system power supplies in segment  $i$ .  $u_i$  and  $i_i$  represent the total voltage and total current of the segment  $i$ , that is, the instantaneous voltage and current values at the corresponding time.

The three-phase instantaneous power in this segment can be expressed:

$$\begin{aligned} p_i &= (u_{p,i} + u_{d,i})(i_{p,i} + i_{d,i}) \\ &= u_{p,i}i_{p,i} + u_{p,i}i_{d,i} + u_{d,i}i_{p,i} + u_{d,i}i_{d,i}. \end{aligned} \quad (22)$$

Then, the disturbance power can be obtained as shown in (23):

$$DP_i = u_{d,i}i_{d,i} = p_i - (u_{p,i}i_{p,i} + u_{p,i}i_{d,i} + u_{d,i}i_{p,i}). \quad (23)$$

By integrating the disturbance power  $DP_i$ , the disturbance energy  $DE_i$  of the segment  $i$  can be obtained; namely,

$$DE_i = \int DP_i dt. \quad (24)$$

When the  $DE_i$  symbol is positive, this basic disturbance event source is located upstream of the monitoring point; when the  $DE_i$  symbol is negative, it is located downstream of the monitoring point.

There may be an action of multiple basic event disturbance sources in a segment. If the location results of the segments in the monitoring point are consistent, this indicates that the monitoring point is affected by the basic

event disturbance source on the same side; when the location results of the segments are inconsistent, this means that there are multiple basic event disturbance sources have impacts on the monitoring point. And according to the value and symbol of the energy value of the section, which basic event disturbance source has a greater impact on the monitoring point can be determined. Furthermore, the location of the originating disturbance event sources and the intermediate disturbance event sources can be determined based on the analysis of the starting and ending time of the segments.

#### 4. Simulation Experiment Analysis

Based on the PSCAD simulation platform, the IEEE39-bus system is built as a simulation example. The corresponding short-circuit fault is set to simulate the voltage sag event, voltage sag data is obtained by simulation, and the validity of the proposed method is verified by analyzing voltage sag data. The analysis of the examples mainly includes three parts: the system operation state verification, the method for determining suspicious areas where the voltage sag source is located, and the verification of the location method of complex voltage sag sources. The structure diagram of the IEEE39-bus system built by the PSCAD simulation platform is shown in Figure 4.

*4.1. The System Operation State Verification.* A three-phase short-circuit fault occurs on line 3–4 close to bus 3 in the IEEE39-bus system, and the simulated bus voltage data is taken as the system state data. The program of single ring theorem and M-P law is written with Python language. The changes in system state data before and after the fault under the single ring theorem and M-P law expression are shown in Figures 5 and 6.

It can be seen from Figure 5(a) that the characteristic roots of the system state matrix are distributed between the outer ring and the inner ring when the system is before the fault, which conforms to the single ring theorem. Some of these points are distributed outside the outer ring, which may be caused by bad data, and have no influence on the result of the single ring theorem. In Figure 5(b), when the system is at fault time, the characteristic roots are concentrated in the inner ring, and most of the characteristic roots are concentrated at the center of the circle. And the single ring theorem of the system state matrix is no longer satisfied. It can be seen from Figure 6(a) that before the fault, the characteristic roots of the covariance matrix of the system state data matrix satisfy the M-P law, the spectral distribution histogram is evenly distributed, and the characteristic roots are distributed between the two limits. In Figure 6(b), when the system is at fault time, the characteristic roots of the covariance matrix do not satisfy the M-P law, and the distribution of the histogram of the spectral distribution is not uniform. The eigenvalues appear abnormally large eigenvalues, and the probability density is abnormally large when the eigenvalues are close to zero. The

single ring theorem and the M-P law can clearly reflect the operating state of the system at a certain moment.

*4.2. The Method of Determining Suspicious Areas.* In the simulation system,  $d_{MSR,0}(t) = 0.1$ . A three-phase short-circuit fault occurs on line 3–4 close to bus 3 at  $t = 5$  s, and the fault is cleared at  $t = 5.5$  s. The voltage data of all bus is used to construct the state data matrix, and the voltage value of each bus is used to construct the influencing factor matrix. The size of the real-time separation window was taken as  $39 \times 50$ . According to the proposed voltage sag source location method, the state augmentation matrix and the reference state augmentation matrix can be established, and the values of  $d_{MSR}(t)$  and  $Q_{MSR}(t)$  of each matrix can be obtained. The changes in  $d_{MSR}(t)$  and  $Q_{MSR}(t)$  are shown in Figures 7 and 8.

It can be seen from Figures 7 and 8 that at  $t = 5$  s, the  $d_{MSR}(t)$  corresponding to the buses 3, 4, 5, 14, and 18 is obviously greater than the critical value  $d_{MSR,0}(t)$ , and the value of  $Q_{MSR}(t)$  corresponding to these buses is larger than that of the other buses, which indicates that the voltage sag sources have a great influence on these buses. The area of those buses is a more likely suspicious area of the voltage sag source.

In order to further explain that the voltage sag source is located in the area of buses 3, 4, 5, 14, and 18, the voltage variation of the system is analyzed. The unit value change diagram of system voltage is shown in Figure 9.

In Figure 9, the voltage sag  $o$  occurs at  $t = 5$  s. The voltage variation results of buses 3, 4, 5, and 18 are more serious than those of the other buses. Therefore, this voltage sag source has a higher impact on these buses, and bus 3 and bus 4 should be closer to the location of the voltage sag source than the other buses. The results of the example show that the indicators using  $d_{MSR}(t)$  and  $Q_{MSR}(t)$  can well determine the suspicious areas where the voltage sag sources are located.

*4.3. Verification of the Location Method of the Complex Voltage Sag Sources.* A three-phase short-circuit fault occurs on line 1–2 close to bus 2 at  $t = 5$  s, and the fault is cleared at  $t = 5.5$  s. Moreover, a three-phase short-circuit fault occurs on line 3–4 close to bus 3 at  $t = 5.2$  s, and the fault is cleared at  $t = 5.7$  s. The simulation data is analyzed according to the method of determining the suspicious area where the voltage sag sources are located, and the curve of voltage variation per unit, the variation curve of  $d_{MSR}(t)$ , and the distribution of  $Q_{MSR}(t)$  of each bus are got as shown in Figures 10–12.

As can be seen from Figure 10, when multiple voltage sag events occur in the system, the system voltage changes will become more complicated. In Figure 11, since the two voltage sag sources appear simultaneously in the system, the variation curve of  $d_{MSR}(t)$  of complex voltage sag events is more complicated than that of the single voltage sag source, and the influence of the two voltage sag sources on each bus at different times is very different. In Figure 12, it can be seen that in the complex voltage sag event, the  $Q_{MSR}$  of buses 1, 2, 3, 4, 5, 14, 18, and 25 are larger than those of the other buses,

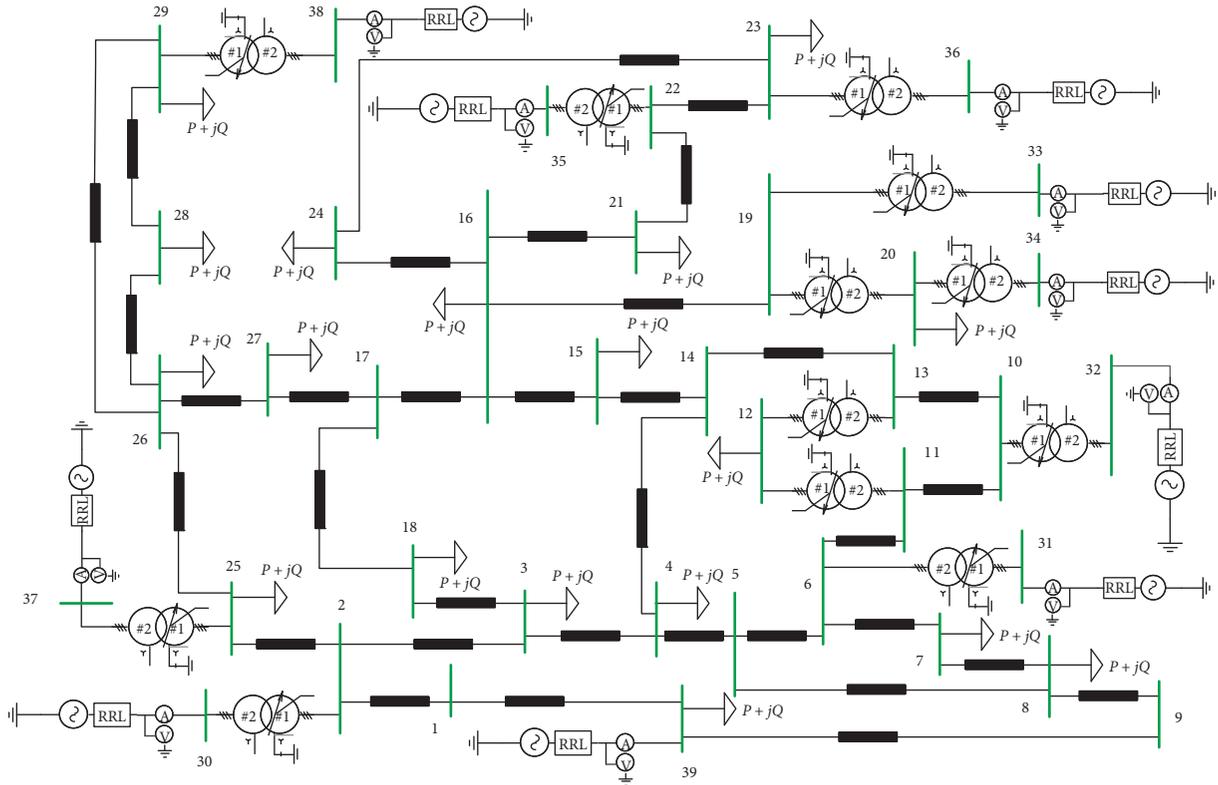


FIGURE 4: IEEE39-bus system.

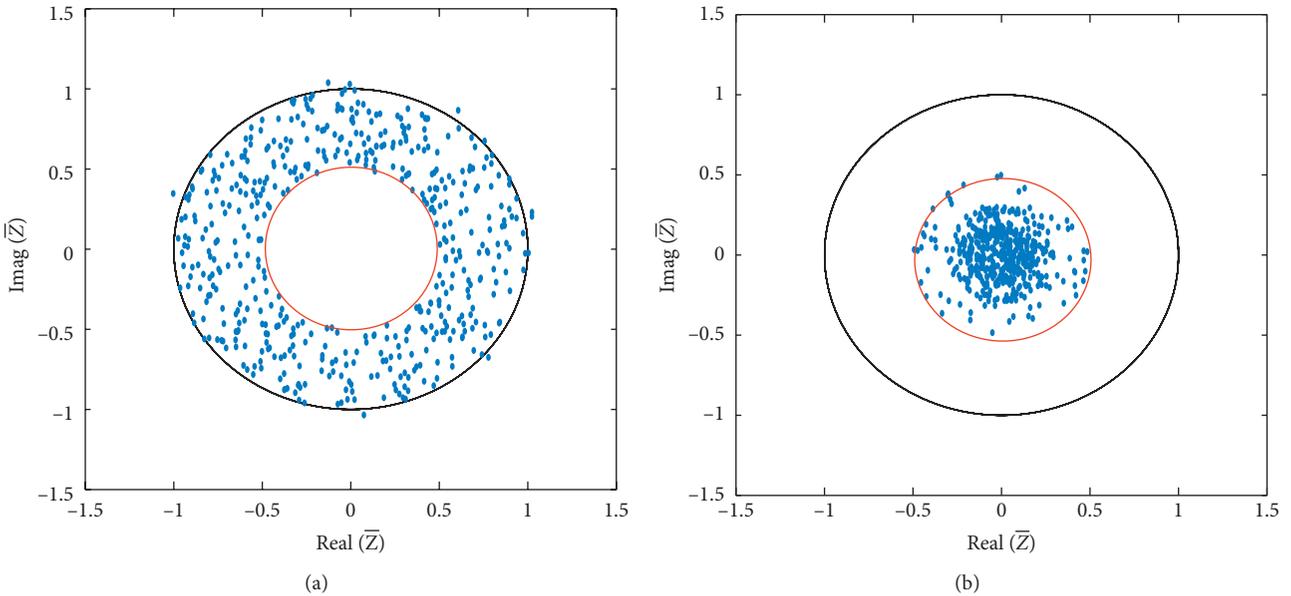


FIGURE 5: Effect diagram of single ring theorem before and after the fault. (a) Single ring theorem of system state before the fault. (b) Single ring theorem of system state after the fault.

indicating that voltage sag sources are more likely to exist in the area of these buses, but it is difficult to determine the exact location of the two voltage sag sources, and it is impossible to determine the impact of the two voltage sag sources on each bus.

Based on the suspicious areas where the voltage sag sources located, the location method of the complex voltage sag sources proposed in this paper is used to analyze the bus voltage and bus current data of buses 1, 2, 3, 4, 5, 14, 18, and 25. According to the three-phase voltage disturbance data of

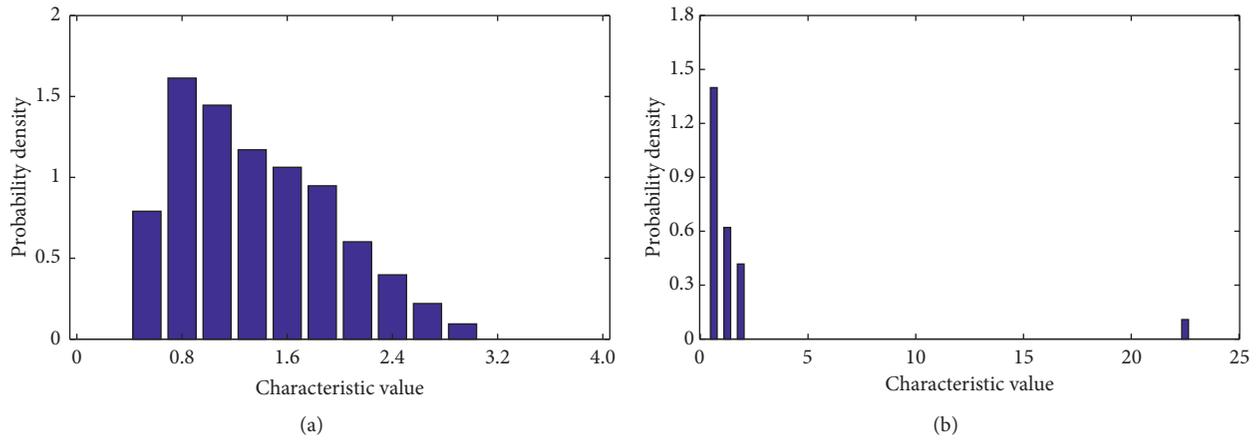


FIGURE 6: Effect diagram of M-P law before and after the fault. (a) M-P law of system state before the fault. (b) M-P law of system state after the fault.

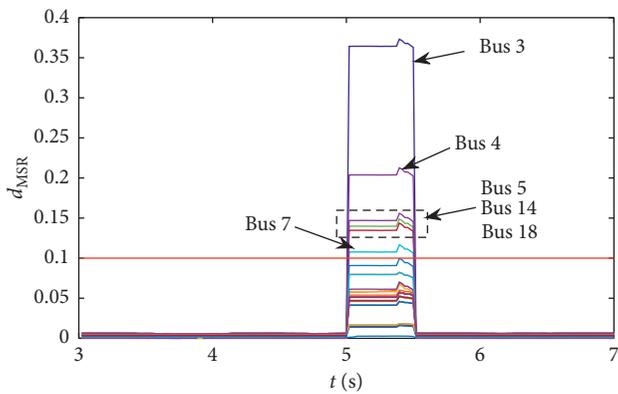


FIGURE 7: Difference curve of the mean spectral radius.

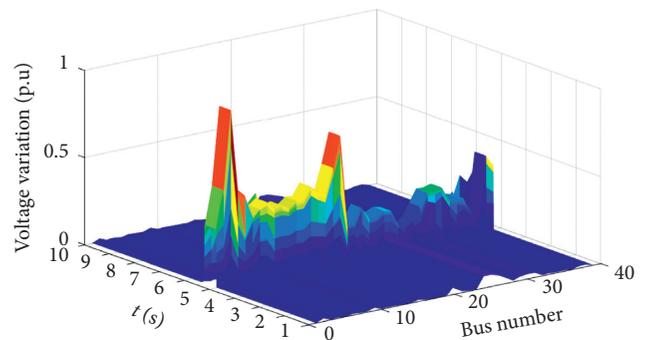


FIGURE 9: Unit value variation of system voltage.

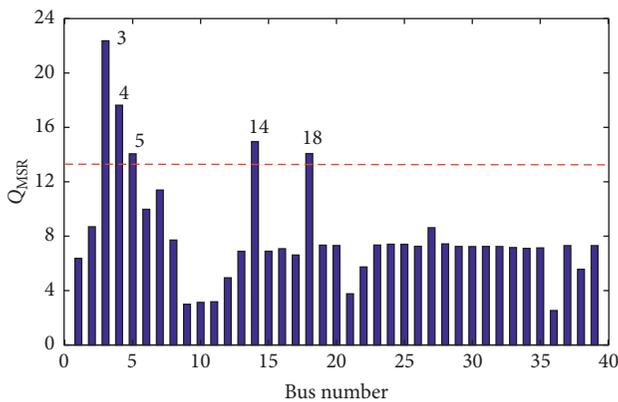


FIGURE 8: Distribution of  $Q_{MSR}$ .

each bus in the voltage sag period, the time distribution diagram of the basic event is established. Since all fault types are three-phase short-circuit fault, which are symmetric faults, the three-phase segments of each bus are basically consistent; therefore, one of the phases is used to obtain the diagram of event segment division for each bus as shown in Figure 13.

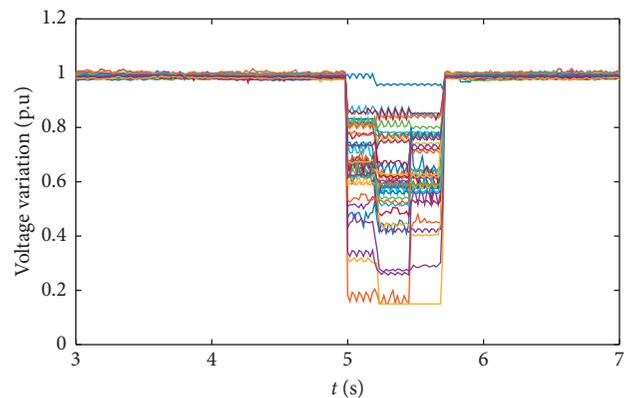


FIGURE 10: Voltage variation curve of each bus in complex voltage sag events.

As shown in Figure 13, bus 2 has a small segment at 5.7 s, which may be a voltage swell segment caused by the transformer connected to bus 2 after all faults are restored, and buses 3, 4, and 5 are the same as bus 2, which have a small nonfaulty voltage swell segment at 5.7 s caused by the transformer. Bus 14 and bus 18 have only two segments, and the long segment occurs at 5.2 s, indicating that a voltage sag source has a much higher impact on bus 14 and bus 18 than the other voltage sag sources.

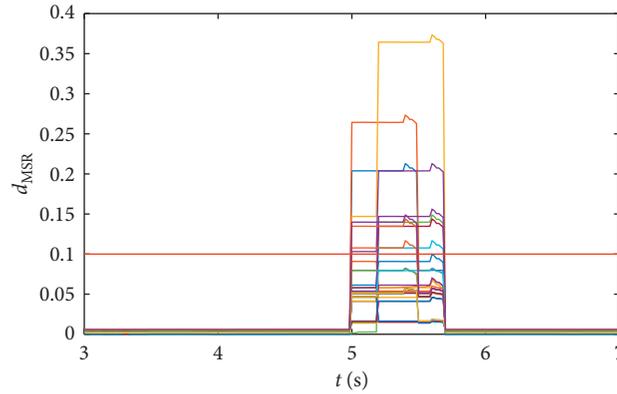


FIGURE 11: Variation curve of  $d_{MSR}$  of complex voltage sag events.

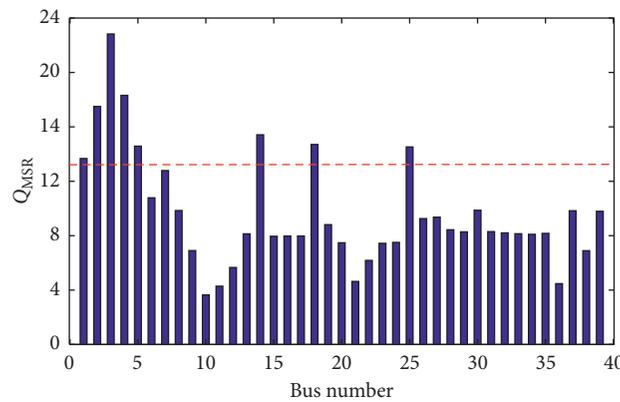


FIGURE 12: Distribution of  $Q_{MSR}$  of complex voltage sag events.

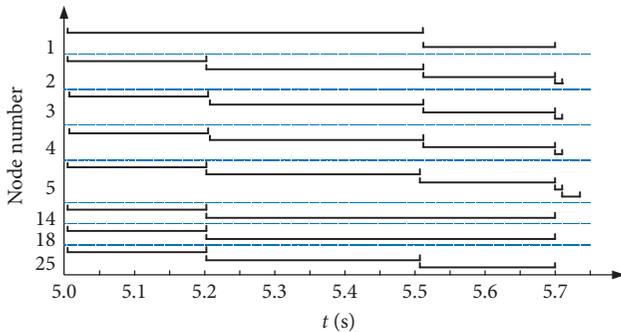


FIGURE 13: The event segment division for each bus.

Table 1 gives the starting and ending times of the voltage segments of each bus and the results of the disturbance location of each segment, where the symbol “↑” indicates that the disturbance source is located in the upstream direction of the bus, the symbol “↓” indicates that the disturbance source is located in the downstream direction of the bus, and the symbol “—” indicates that the disturbance segment does not exist.

As shown in Table 1, the results of the location of all disturbance event segments are the same for bus 1, so all voltage sag sources are located upstream of bus 1. For bus 2, the location results of all disturbance event segments show

that voltage sags are located downstream of bus 2. Similarly, the results of the location of all disturbance event segments for buses 4, 5, 14, and 18 indicate that all voltage sag sources are located upstream of these nodes. The voltage sag sources are located downstream of bus 25. For bus 3, the results of the location of disturbance segments are different, which means that bus 3 is affected by multiple voltage sag sources in different directions. In segment 1, it can be judged that there is a voltage sag source in this period according to the location results, and this voltage sag source should be located on line 2–3 or line 2–1. In segment 2, because the result of the location for bus 3 is changed, it can be concluded that another voltage sag source begins to affect the system at 5.2 s, and the voltage sag source should be located on line 3–4. This is similar to the result of the set simulation example. Therefore, the location of voltage sag sources can be determined more accurately by dividing the complex voltage sag events into the basic disturbance event segments and analyzing the result of the location of all disturbance segments.

In addition, as can be seen from Table 1, buses 1, 14, and 18 have only two segments, the starting and ending time of segment 1 of bus 1 is 5.004 s–5.505 s, but the starting and ending time of bus 14 and bus 18 is 5.004 s–5.200 s, which means that the first voltage sag source has a greater influence on bus 1 than the second

TABLE 1: The results of the disturbance location of each segment.

Bus number	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
1	(5.004–5.200) ↑	(5.200–5.505) ↑	(5.505–5.701) ↑	—	—
2	(5.004–5.200) ↓	(5.200–5.504) ↓	(5.504–5.701) ↓	(5.701–5.712) ↓	—
3	(5.004–5.201) ↑	(5.200–5.504) ↓	(5.504–5.701) ↓	(5.701–5.712) ↓	—
4	(5.004–5.201) ↑	(5.200–5.504) ↑	(5.504–5.701) ↑	(5.701–5.712) ↑	—
5	(5.004–5.200) ↑	(5.200–5.505) ↑	(5.505–5.701) ↑	(5.701–5.712) ↑	(5.712–5.736) ↑
14	(5.004–5.200) ↑	(5.200–5.701) ↑	—	—	—
18	(5.004–5.200) ↑	(5.200–5.701) ↑	—	—	—
25	(5.004–5.200) ↓	(5.200–5.504) ↓	(5.504–5.701) ↓	—	—

voltage sag source, and the influence of the first voltage sag source on bus 14 and bus 18 is less than the second voltage sag source. For bus 3, the first voltage sag source at 5.504 s–5.201 s is affected; the second voltage sag source begins to affect the system, and this voltage sag source has a greater impact on bus 3 after 5.2 s. Therefore, the second voltage sag source is closer to the position of bus 3. Analyzing the starting and ending time of each bus's disturbance segments, we can obtain the main cause of the voltage sag of the bus in a certain period of time. This method can be used as a support for analyzing the consequences of the voltage sag event and the accountability of the voltage sag event.

## 5. Conclusion

In this paper, a new method is proposed to locate the complex voltage sag source, which is based on random matrix theory and traditional disturbance power and disturbance energy method. Utilizing the single ring theorem and M-P law of random matrix theory, the grid operation state can be judged obviously at a certain time, and the correlation between the data of the state augmentation matrix is analyzed by the mathematical statistics method to determine the possible area of voltage sag source. On the basis of judging the suspicious area where the voltage sag sources are located, the sequence division method and the disturbance power and the disturbance energy method are combined to analyze the data of each bus in the suspicious area, thereby accurately determining the location of each voltage sag source in the complex voltage sag source event. The proposed method is not limited to the model structure and the operation mechanism of the distribution network and can improve the accuracy of sag source location with fewer data.

For the location problem of complex voltage sag sources, further research is needed. For example, in the timing sequence division method, the matching pursuit algorithm will have a large computational problem in the analysis process, which is not conducive to the online analysis of the voltage sag disturbance signal, and can be optimized by the combination of the optimization algorithm and the matching pursuit algorithm which improves the calculation speed while ensuring accuracy. In addition, for the complex distribution network system, we can consider the reasonable division of the large power grid and then use the method in this paper to analyze each subnetwork.

## Data Availability

The data used to support the results of this study are obtained from PSCAD simulation. The authors can provide relevant data to support the results of this study upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 51467009), the Basic Research Innovation Group Project of Gansu Province (No. 18JR3RA133), and the Science and Technology Foundation of State Grid Corporation of China.

## References

- [1] H. Shareef, A. A. A. Mohamed, and A. A. Ibrahim, "Identification of voltage sag source location using S and TT transformed disturbance power," *Journal of Central South University*, vol. 20, no. 1, pp. 83–97, 2013.
- [2] H. Dong, Q. Jia, N. Wang et al., "Extension mining method of complex power quality correlation disturbance mode," *Chinese Journal of Electrical Engineering*, vol. 36, no. 5, pp. 1269–1278, 2016.
- [3] M. Karimi, M. R. H. Mokhtari, and M. R. Iravani, "Wavelet based on-line disturbance detection for power quality applications," *IEEE Transactions on Power Delivery*, vol. 15, no. 4, pp. 1212–1220, 2000.
- [4] C. Zhao, S. Tao, and X. Xiao, "fault location estimate-on based on voltage sag information of PQMS," *Power System Technology*, vol. 40, no. 02, pp. 642–648, 2016.
- [5] Y. Mohammadi and R. C. Leborgne, "Improved DR and CBM methods for finding relative location of voltage sag source at the PCC of distributed energy resources[J]," *International Journal of Electrical Power and Energy Systems*, vol. 117, 2020.
- [6] O. N. Gerek and D. G. Ece, "Power-quality event analysis using higher order cumulants and quadratic classifiers," *IEEE Transactions on Power Delivery*, vol. 21, no. 2, pp. 883–889, 2006.
- [7] D. Li, F. Mei, C. Zhang, H. Sha, and J. Zheng, "Self-supervised voltage sag source identification method based on CNN," *Energies*, vol. 12, no. 6, p. 1059, 2019.
- [8] H. Dong, Q. Jia, N. Wang, L. Shi, and H. Yu, "Study on location method of complex power quality disturbance events," *Chinese Journal of Electrical Engineering*, vol. 37, no. 17, pp. 4931–4940, 2017.

- [9] A. d. Santos, T. Rosa, and M. T. Correia De Barros, "Stochastic characterization of voltage sag occurrence based on field data," *IEEE Transactions on Power Delivery*, vol. 34, no. 2, pp. 496–504, Article ID 8516317, 2019.
- [10] W. Liu, D. Zhang, and Y. Ding, "Power grid vulnerability identification methods based on random matrix theory and entropy theory," *Proceedings of the CSEE*, vol. 37, no. 20, pp. 5893–5901, 2017, in Chinese.
- [11] Bo Wang, J. Wang, D. Liu et al., "Research on evaluating vulnerability of power network based on high-dimensional random matrix theory," *Proceedings of the CSEE*, vol. 39, no. 6, pp. 1682–1691, 2019, in Chinese.
- [12] Y. Xiong, W. Yao, W. Chen, J. Fang, X. Ai, and J. Wen, "A data-driven approach for fault time determination and fault area location using random matrix theory," *International Journal of Electrical Power and Energy Systems*, vol. 116, 2020.
- [13] B. Han, G. L. Luo, and G. LiJiang, "Framework of random matrix theory for power system data mining in a non-Gaussian environment," *IEEE Access*, vol. 4, pp. 9969–9977, 2016.
- [14] Z. Bai and J. W. Silverstein, "Spectral analysis of large dimensional random matrices," *Journal of the Royal Statistical Society*, vol. 175, no. 3, pp. 822–823, 2012.
- [15] Y. Yan, G. Sheng, R. C. Qiu, and X. Jiang, "Big data modeling and analysis for power transmission equipment: a novel random matrix theoretical approach," *IEEE Access*, vol. 6, pp. 7148–7156, 2018.
- [16] Z. Ling, R. C. Qiu, X. He et al., "A new approach of exploiting self-adjoint matrix polynomials of large random matrices for anomaly detection and fault location," *Journal of Latex Class Files*, vol. 14, no. 8, pp. 1–11, 2015.
- [17] J. Bun, J.-P. Bouchaud, and M. Potters, "Cleaning large correlation matrices: tools from random matrix theory," *Physics Reports*, vol. 666, pp. 1–109, 2017.
- [18] Y. Wang, J. Gu, and Z. Jin, "Electric larceny recognition method based on high dimensional random matrix analysis," *Modern Electric Power*, vol. 34, no. 6, pp. 71–78, 2017.
- [19] X. He, R. C. Qiu, Q. Ai, L. Chu, X. Xu, and Z. Ling, "Designing for situation awareness of future power grids: an indicator system based on linear eigenvalue statistics of large random matrices," *IEEE Access*, vol. 4, pp. 3557–3568, 2016.
- [20] X. Xu, X. He, and Ai Qian, "A correlation analysis method for operation status of distribution network based on random matrix theory," *Power System Technology*, vol. 40, no. 3, pp. 781–790, 2016, in Chinese.
- [21] M. De Santis, C. Noce, P. Varilone, and P. Verde, "Analysis of the origin of measured voltage sags in interconnected networks," *Electric Power Systems Research*, vol. 154, pp. 391–400, 2018.
- [22] Z. Qu, H. Wanru, and N. Wang, "Application of atomic rapid decomposition algorithm in power quality disturbance analysis," *Electric Power Automation Equipment*, vol. 35, no. 10, pp. 145–150, 2015.
- [23] X. Hong, J. Zhou, and Y. He, "Damage detection of anchored region on the messenger cable based on matching pursuit algorithm," *Mechanical Systems and Signal Processing*, vol. 130, pp. 221–247, 2019.
- [24] Y. Tan, L. Wen, Y. Qi et al., "A direction finder for voltage sag source based on the disturbance power direction," *Proceedings of the CSEE*, vol. 35, no. 9, pp. 2202–2208, 2015, in Chinese.