Operation State Evaluation Method of Smart Distribution Network Based on Free Probability Theory

Jiaxin Zhang¹, Bo Wang¹*, Hongxia Wang¹, Hengrui Ma², Fuqi Ma¹, Yifan Li¹ and Yingchen Zhang¹

¹School of Electrical Engineering and Automation, Wuhan University, Wuhan, China, ²Tus-Institute for Renewable Energy, Qinghai University, Qinghai, China

In view of the current situation that the new generation of smart grids with “double high” characteristics is in urgent need of effective state evaluation methods due to the characteristics of strong volatility and diverse demands, a method of operation state evaluation of smart distribution networks based on free probability theory is proposed, which is combined with high-order moment indexes to describe the operation trajectory of distribution networks from a data-driven perspective. First, the state assessment problem of smart distribution networks is modeled as a binary hypothesis testing problem, and the asymptotic free equation is established based on free probability theory to provide a framework for state assessment of distribution networks. Then, a high-order moment evaluation index is proposed, combined with the sliding time window processing, and the high-order moment sequence was obtained based on the high-dimensional data of the distribution network, which is used to describe the state evolution of the distribution network. Finally, this method is applied to a certain 110-kV distribution network. The analysis of an example shows that the proposed evaluation framework and indicators can effectively reflect the data changes in the distribution network and support the state assessment and evolution analysis of the distribution network.

Keywords: free probability theory, asymptotic spectral distribution, free convolution operation, distribution network status assessment, moments

INTRODUCTION

As China proposes “to achieve carbon peak before 2030 and achieve carbon neutrality before 2060,” improving overall energy utilization efficiency and focusing on developing renewable energy has become an inevitable choice. The power system is closely related to the production, transportation, and consumption of renewable energy, which also plays a key role in promoting energy transformation (Zhou et al., 2018). Moreover, the new generation smart grid has the characteristics of “double high”: the high proportion of renewable energy and the high proportion of power electronic equipment. Distributed power supply, charging pile/station, and controllable load, along with other devices, are developing rapidly, and their large-scale access to the distribution network has brought strong uncertainties to the distribution network due to their characteristics of strong intermittent and diverse demands, making the operation mode of the distribution network increasingly complex and changeable.

As a pivotal link of energy, an intelligent distribution network is the key to the smooth operation of “the production-supply-marketing” of electric energy. The ultimate goal of the intelligent
The application of cloud computing, big data, Internet of Things, 5G information communication technology, and artificial intelligence in the power system provides a data basis for the realization of real-time state estimation and situation awareness of the smart distribution network (Yang et al., 2018; Shen et al., 2020a; Ma et al., 2020; Wang et al., 2020). However, the existing distribution network state estimation and situation awareness methods find it difficult to meet the requirements in many aspects of calculation accuracy, calculation speed, and visualization. Thus, it has become a hot topic for experts, scholars, and engineers at home and abroad to construct an effective state estimation and situation awareness system for the intelligent distribution network to support comprehensive, accurate, and real-time control of the operation situation of the distribution network.

In the field of power system state assessment, apart from the specific model method, classical research methods also include the analytic hierarchy process (AHP), the fuzzy comprehensive evaluation method, and principal component analysis (PCA). The literature (Cao et al., 2007) has proposed a comprehensive evaluation method for a new rural low-voltage distribution network based on the AHP and realized practical application. However, this method solved the problems of distribution network construction planning and transformation in the near future and could not evaluate the real-time status. The literature (Sun et al., 2017) combined PCA and system clustering analysis to establish a comprehensive evaluation system of county power grids, which could evaluate the power grid from five aspects of security, economy, reliability, adaptability, and quality, but also could not evaluate the real-time state.

In recent years, data-driven state assessment has also been applied in the power field. The literature (Xu et al., 2016) proposed a correlation analysis method based on random matrix theory (RMT). Combining real-time separation window technology with RMT, the mean spectrum radius (MSR) index was used to evaluate the correlation of distribution network data. In the literature (Xu et al., 2018), the evaluation indexes of the hit ratio and the false alarm rate were proposed, and the vulnerability of the distribution network was evaluated based on RMT. The literature (He et al., 2017a) used the basic breakthrough of high-dimensional statistics in recent years to put forward the research framework of space-time big data of the distribution network based on the random matrix. For power system fault identification, the literature (Xu et al., 2019) proposed a feature self-learning method based on deep learning for high-dimensional space-time fault samples, which had fast calculation speed and strong robustness. However, the method itself had high requirements on source data but low comprehensibility. The literature (Wei et al., 2016) proposed a high-dimensional power data fusion method based on correlation mining in order to solve the key problem of online stability analysis of large power grids. This method mainly solved the problem of data fusion and did not directly evaluate the status of the large power grid.

The key of power system state evaluation based on high-dimensional big data should be the construction of an evaluation framework and an evaluation index system and identification of a power grid or equipment evolution situation. In this study, the free probability theory (FPT) is introduced into the electric power field for the first time in China, providing a complete and clear evaluation framework for the operation status of the smart distribution network, combining with sliding time window processing to solve the high-dimensional source data to obtain the time series of the index. In this study, the high-order moment index is also proposed to analyze the distribution network from the perspective of the state assessment and evolution trend and is applied to the state assessment of the 110-kV distribution network to verify the effectiveness of the proposed method.

**FREE PROBABILITY THEORY AND BIG DATA PROCESSING METHODS**

**Introduction to Free Probability Theory**

A random matrix is a matrix whose elements are random variables. Through the high-dimensional statistical analysis, important information can be extracted from massive disordered data in a random matrix.

Free probability theory can provide an effective analysis framework for the asymptotic spectrum distribution of high-dimensional random matrices. In the 1980s, Voiculescu proposed FPT to deal with abstract “non-commutative space,” and the random matrix is a special case of “non-commutative space” (Dan, 1986; Voiculescu, 1987). The purpose of FPT is to introduce a concept similar to “independence” in classical probability theory, namely, “freedom,” and make it applicable to non-commutative random variables such as random matrices and extend it to the case of large dimensions, namely, “asymptotic freedom.”

Different from traditional mathematical theory, FPT defines some new operators, including additive-free convolution ★ and its inverse operation additive-free deconvolution †, multiplicative-free convolution ⊗ and its inverse operation multiplicative-free deconvolution †, similar to the addition, subtraction, multiplication, and division operations in classical mathematics. Based on the asymptotic spectrum theory of RMT, combining the concept of asymptotic freedom with the above new operators, some difficult problems in classical mathematics can be solved. In traditional mathematics, if and only if two matrices are commutative, the eigenvalues of their sum matrix or product matrix can be obtained from their respective eigenvalues. In FPT, if two random matrices are asymptotically free and their respective asymptotic spectral distributions are known, the asymptotic spectral distributions of their sum matrix or product matrix can be obtained and vice versa. It is worth mentioning that in FPT, the semicircular law is similar to the classical Gaussian distribution, that is, the normalization of the free random matrix (given spectral distribution) and the spectral
distribution of matrix converge to the semicircular law; the Marchenko–Pastur law (M-P law) is similar to the classical Poisson distribution, that is, the normalization of those single-rank free random matrices and the spectral distribution of matrices converge to the M–P law (Tulino and Verdu, 2004).

FPT has become a powerful tool to describe the characteristics of wireless communication systems. Spectrum sensing algorithms based on FPT have fast convergence, which are also suitable for the limited number of samples, and have high sensing performance in the case of low signal-to-noise ratio (Tulino and Verdu, 2004). Domestic and foreign scholars have made great academic achievements in this field. This study is the first attempt to apply the FPT to the electric power field, which selects the operation of the smart distribution network as the application scenario and evaluates the operation status of the distribution network based on FPT, providing real-time and efficient support for intelligent operation and maintenance.

Free Probability Theory

Definition. 1 (Dan, 1986) The empirical spectrum distribution of an \( N \times N \) random matrix \( B_N \) is defined as follows:

\[
\mu_{B_N}(x) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(\lambda_i(B_N) \leq x),
\]

where \( \lambda_i(B_N) \), \( i = 1, 2, \ldots, N \) are the eigenvalues of \( B_N \), \( \mathbb{I}(\cdot) \) is the indicator function.

In RMT, the asymptotic spectral distribution (ASD) is the empirical spectral distribution of \( B_N \) when \( N \to \infty \), which is represented by the symbol \( \mu_B \), and can be expressed uniquely by the following moment:

\[
m_k = \lim_{N \to \infty} \frac{1}{N} \mathbb{E}[\text{tr}(B_N^k)] = \int x^k d\mu_B(x),
\]

where \( \text{tr}(\cdot) \) represents the rank of the matrix and \( k \) represents the order of moments. In mathematics and statistics, moments can represent the distribution and morphological characteristics of variables. The specific moment algorithm of the method proposed in this study will be introduced in detail in Calculation of High-Order Moment Index section.

As mentioned above, Voiculescu proposed FPT in order to introduce the concept of “freedom” and summarize the law applicable to non-commutative variables such as random matrices, which is similar to the law in classical probability theory.

Random matrices are just one kind of non-commutative variable, and non-commutative variables are all elements of “non-commutative probability space.” The concept of a non-commutative probability space is as follows.

Definition. 2 (Couillet and Debbah, 2011) Let \( B \) be a non-commutative algebraic system with unit element \( I \); if \( \phi \) is a linear function on \( B \) and meet \( \phi(I) = 1 \), then the order pair \( (B, \phi) \) is called a non-commutative probability space.

For random matrices, the identity element \( I \) is the identity matrix \( I_N \), and \( \phi \) is defined as follows:

\[
\phi(C) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}[c_{ii}^{N}] = \frac{1}{N} \mathbb{E}[\text{tr}(C)],
\]

where \( C \in B \), \( c_{ii}^{N} \) represents the element of \( C \). It can be seen from Eq. 3 that \( \phi \) is the function of solving the moment. This kind of non-commutative probability space meets trace lemma, that is, \( \phi(ab) = \phi(ba) \).

Definition. 3 (Couillet and Debbah, 2011) Let \( (B, \phi) \) be a non-commutative probability space, and for all n-dimensional sequences \( (b_1, b_2, \ldots, b_n) \), if \( \phi(b_1 b_2 \ldots b_n) = 0 \) satisfies the following conditions:

1) \( b_j \in B_j \), where \( i_j \leq K \)
2) \( i_1 \neq i_2 \neq i_3 \neq \ldots \neq i_{n-1} \neq i_n \)
3) for all \( j \in \{1, \ldots, n\} \), \( \phi(b_j) = 0 \)

Then, a family of the subalgebra systems of \( B \{B_1, \ldots, B_K\} \) is free.

Obviously, if \( \{b_1, \ldots, b_n\} \) is free (each subalgebra system consists of only one of their elements), then random variables \( \{b_1, \ldots, b_n\} \) are free.

Furthermore, let us extend the concept of “freedom” to “asymptotic freedom.”

Definition. 4 (Couillet and Debbah, 2011) If the following two conditions are met

1) for all \( k \in \{1, \ldots, K\} \), \( X_{N,k} \) has an asymptotic spectral distribution;
2) for all \( \{i_1, \ldots, i_n\} \subset \{1, \ldots, K\} \), \( i_1 \neq i_2 \neq i_3 \neq \ldots \neq i_{n-1} \neq i_n \) and family of unary polynomials \( \{p_1, \ldots, p_n\} \),

\[
\lim_{N \to \infty} \phi'\{P(X_{N,i_j})\} = 0, j \in \{1, \ldots, n\}
\]

and

\[
\lim_{N \to \infty} \phi'\left(\prod_{j=1}^{n} P(X_{N,i_j})\right) = 0.
\]

Then the \( N \times N \) random matrix family \( \{X_{N,1}, \ldots, X_{N,K}\} \) of non-commutative probability spaces \( (B_N, \phi) \) is asymptotically free.

Based on the above definition of asymptotic freedom, combined with new operators such as additive-free convolution, the following illustration is made. If two random matrices \( A_N \in \mathbb{C}^{N \times N} \) and \( B_N \in \mathbb{C}^{N \times N} \) are asymptotically free, with their asymptotic spectral distributions denoted as \( \mu_A \) and \( \mu_B \), respectively, and \( A_N + B_N \) is known to have asymptotic spectral distribution \( \mu_{A+B} \), then

\[
\mu_{A+B} = \mu_A \boxplus \mu_B,
\]

where \( \boxplus \) is called additive-free convolution, namely, \( \mu_{A+B} \) is the additive-free convolution of \( \mu_A \) and \( \mu_B \).
Furthermore, $\bowtie$ is defined as additive-free deconvolution, that is, if $\mu_C = \mu_A \bowtie \mu_B$, then $\mu_A = \mu_C \bowtie \mu_B$, and $\mu_B = \mu_C \bowtie \mu_A$. So additive-free convolution and additive-free deconvolution are inverse operations of each other.

Similarly, if $A_NB_N$ has an asymptotic spectral distribution $\mu_{AB}$, then

$$\mu_{AB} = \mu_A \bowtie \mu_B, \quad (7)$$

where $\bowtie$ is called multiplicative-free convolution, namely, $\mu_{AB}$ is the multiplicative-free convolution of $\mu_A$ and $\mu_B$.

Furthermore, $\boxdot$ is defined as additive-free deconvolution, that is, if $\mu_C = \mu_A \boxdot \mu_B$, then $\mu_A = \mu_C \boxdot \mu_B$, and $\mu_B = \mu_C \boxdot \mu_A$. So multiplicative-free convolution and multiplicative-free deconvolution are inverse operations of each other. Both additive- and multiplicative-free convolution are commutative, namely, $\mu_A \boxdot \mu_B = \mu_B \boxdot \mu_A$, $\mu_A \bowtie \mu_B = \mu_B \bowtie \mu_A$. In this way, non-commutative variables (such as random matrices) can be exchanged in the operation.

**DISTRIBUTION NETWORK STATE ESTIMATION METHOD BASED ON FPT**

**Data Pre-Processing**

Supervisory Control and Data Acquisition (SCADA) is widely applied in the power system and collects the branch power, branch current amplitude, and node voltage amplitude in the system with high maturity, mainly through the remote terminal unit (RTU) and the feeder terminal unit (FTU) (Yang et al., 2020a; Shen et al., 2020b; Zhu et al., 2020; Li et al., 2021a; Shen et al., 2021a; Shen et al., 2021b; Qi et al., 2021; Xiang et al., 2021). The data collected by SCADA has the characteristics of mass and high dimension, so a high-dimension random matrix can be constructed according to the collected data. Combined with the sliding time window processing, the data characteristics of the distribution network can be analyzed based on FPT, and the distribution network state before and after can be compared in time to further realize the evolution of the distribution network operation state.

Assume that $i$ nodes in the distribution network are equipped with measuring devices, and the sampling interval is 0.01 s. At the sampling time $t_{ni}$, $i$ nodes each generate a state data (which can be voltage, current, and power angle), and the state data of all nodes at this time constitute a column vector $x(t_n)$, as shown in the following formula:

$$x(t_n) = [x_{1t_n}, x_{2t_n}, ..., x_{it_n}]^T. \quad (8)$$

When there are a total of $j$ sampling moments, the $j$ column vectors are arranged to form a high-dimensional random matrix $X_{ijp}$, as shown below:

$$X_{ij} = \begin{bmatrix}
x_{11} & x_{12} & \cdots & x_{1j} \\
x_{21} & x_{22} & \cdots & x_{2j} \\
\vdots & \vdots & \ddots & \vdots \\
x_{i1} & x_{i2} & \cdots & x_{ij}
\end{bmatrix}, \quad (9)$$

In the above formula, each row of $X_{ijp}$ is the state data of the same node at different times, and each column is the state data of different nodes at the same time.

The normalization process is carried out according to Eq. 10 below, and the normalized matrix $\hat{X}$ with mean value $E = 0$ and variance $\sigma^2 = 1$ is obtained as follows:

$$\hat{X}_i = (X_i - E(X_i))/\sigma^2(X_i), i = 1, 2, ..., i, \quad (10)$$

where $X_i$ is the $i$th row of $X$.

In statistical analysis of high-dimensional data, when the amount of data is large enough, the data as a whole will show certain random statistical characteristics after corresponding processing, such as the single ring theorem and the M–P law (Ling et al., 2018; Jain et al., 2019; Deepa et al., 2020; Xiong et al., 2020; Yang et al., 2020b; Li et al., 2021b; Li et al., 2021c; Yang et al., 2021c; Li et al., 2021d; Ye et al., 2021; Dong and Li, 2021; Liu et al., 2021; Mousavizadeh et al., 2021; Ouyang and Xu, 2021; Zhu et al., 2021). In the statistical analysis of high-dimensional power data, the corresponding linear eigenvalue statistics (LES) are constructed, such as MSR, high-order moment, etc., which can effectively represent the state of the distribution network. When there are only random fluctuations and measurement errors in the measured data, the data present a random statistical characteristic as a whole. If abnormal events occur in the power system, the original stable operation state of the system will be broken, and the measured data will change accordingly.

**State Assessment Model Construction**

The problem of distribution network operation state assessment is understood as a binary hypothesis testing problem, as follows:

$$y(n) = \begin{cases} v(n), & H_0 \\ x(n) + v(n), & H_1, \end{cases} \quad (11)$$

where $y(n)$ represents the received sampled signal, $x(n)$ represents the event signal component, and $v(n)$ represents the noise component.

The above binary hypothesis testing problem is further explained. $H_0$ means that no abnormal events occur, and the received sampled signal only has randomly distributed noise components. $H_1$ indicates that abnormal events occur, event signals exist in the sampled signals, and the original stable operation state of the system is broken.

In this study, the basic idea of the distribution network state estimation method based on FPT is to estimate the asymptotic spectrum distribution of the event signal component $x(n)$ by establishing and solving the asymptotic free equation and then calculate the high-order moment index $m_i (i = 1, 2, \ldots, n)$ of the event signal component $x(n)$. Thus, the high-order moment is the detection statistic of the algorithm.

Assuming that $N$ received sampled signals $y(1), y(2), \ldots, y(N)$ are used for distribution network state estimation and each sampled signal is composed of $M$ signal components, the sample covariance matrix of received sampled signals is as follows:

$$\begin{bmatrix}
X_{11} & X_{12} & \cdots & X_{1j} \\
X_{21} & X_{22} & \cdots & X_{2j} \\
\vdots & \vdots & \ddots & \vdots \\
X_{i1} & X_{i2} & \cdots & X_{ij}
\end{bmatrix}.$$
The asymptotic spectral distributions of the above two sample covariance matrices $\Sigma_y$ and $\Sigma_x$ satisfy the following asymptotic free equation (Ryan and Debbah, 2007):

$$\mu_{\Sigma_y} \boxplus \mu_c = \left( \mu_{\Sigma_x} \boxplus \mu_{\Sigma_x} \boxplus \mu_{\Sigma_x} \right) \boxplus \mu_{c^2}. \quad (14)$$

where $c = M/N$, and $\mu_{c^2}$ represents a probability distribution that has value only at point $c^2$.

Based on the established asymptotic free equation and the sliding time window method, the high-order moment index of the continuous time window matrix is obtained to observe the state of the distribution network.

The state assessment process of the distribution network is shown in Figure 1 below.

**Calculation of High-Order Moment Index**

From the above analysis, it can be seen that the calculation process of solving the high-order moment index is based on the asymptotic free equation, which involves new operators defined by FPT, namely, additive-free convolution and its inverse operation and multiplicative-free convolution and its inverse operation. It is a relatively simple calculation method to calculate additive-free convolution through the moment-cumulant formula (Ryan and Debbah, 2007).

1) Additive-free convolution

The moment-cumulant formula describes the relationship between the moments of a certain measure and the related $R$ transformation. The $R$ transformation of a probability distribution $\mu$ is defined as follows:

$$R_\mu = \sum_n \alpha_n^\mu z^n, \quad (18)$$

where $\alpha_n^\mu$ is the $n$th order cumulant of $\mu$. Based on $R$ transformation, additive-free convolution can be realized, as shown in the following formula:

$$R_{\mu_A \boxplus \mu_B} = R_{\mu_A}(z) + R_{\mu_B}(z), \quad (19)$$

which is equivalent to that the cumulative measure has additivity under additive-free convolution. That is,

$$a_n^{\mu_A \boxplus \mu_B} = a_n^{\mu_A} + a_n^{\mu_B}. \quad (20)$$

The moment-cumulant of the distribution $\mu$ is given as follows:

$$m_n^\mu = \sum_{n \leq k} a_n^\mu \text{cof}_{f_{k-n}} \left( (1 + m_n^\mu z + m_2^\mu z^2 + ...)^n \right), \quad (21)$$

where $\text{cof}_{f_{k-n}}(\cdot)$ is the coefficient of $z^n$. 

Substituted into Eq. 10, the asymptotic free equation becomes:

$$\mu_{\Sigma_y} \boxplus \mu_c = \left( \mu_{\Sigma_x} \boxplus \mu_{\Sigma_x} \boxplus \mu_{\Sigma_x} \right) \boxplus \mu_{c^2}. \quad (16)$$

The asymptotic spectrum distribution of event signal component $x(n)$ can be obtained by solving the asymptotic free equation as follows:

$$\mu_{\Sigma_x} = \left[ \left( \mu_{\Sigma_x} \boxplus \mu_{\Sigma_x} \right) \boxplus \mu_{\Sigma_x} \right] \boxplus \mu_{c^2}. \quad (17)$$
The bidirectional conversion between the cumulant sequence and the moment sequence can be completed by using the above formula, that is, the first n-order cumulants can be obtained from the first n-order moments and vice versa.

This article gives a brief description of the use of the moment-cumulant formula in free convolution calculation, and the specific process is described as follows:

1) Taking the sequence of moments as input, vector $f=(m_1, m_2, ..., m_n)$ with length $n+1$ is formed, where $m_1$ is the first-order moment, and $m_n$ is the nth-order moment. Then $n$ vectors are obtained by convolution calculation according to the following formula:

$$F_1 = f, F_2 = f * f, ..., F_n = *_nf,$$

where $*$ stands for the convolution operation, and $*_n$ stands for n-fold classical convolution with itself. With the accumulation of the convolution operation, the length of vector $F$ increases gradually. Since only the first $n+1$ elements of $M_1, ..., M_n$ are used in subsequent operations, the length of vector $F$ is uniformly trimmed to $n+1$ after the convolution operation in order to simplify calculation and reduce the storage space of the operation.

2) Calculate each cumulant iteratively. After the cumulants $\alpha_1, ..., \alpha_{n-1}$ are obtained by solving the moment-cumulant formula shown in Eq. 21 for n-1 times, $\alpha_n$ can be obtained by solving the equation for the nth time. It should be added that the relation between each $F$ vector in Step 1) and Eq. 21 can be expressed by the following:

$$\text{coef}_{n-k}(1 + \mu_1 z + \mu_2 z^2 + ... + \mu_k z^k) = F_k(n - k). \quad (22)$$

This equation can also be understood as writing the coefficients in the moment-cumulant formula as k-fold convolution. Based on this formula, it can be known that the $k$th cumulant is equivalent to the following expression:

$$\alpha_k = \frac{M_k(n+1) - \sum_{i=1}^{k-1} \alpha_i M_i(k-r)}{M_k(0)}. \quad (23)$$

Thus, the additive-free convolution and additive-free deconvolution can be easily calculated by means of the moment-cumulant formula.

2) Multiplicative-free convolution

Computation involving multiplicative-free convolution and its inverse operation requires the transformation of boxed convolution, denoted as $\boxplus$. Boxed convolution can be understood as a convolution operation acting on a power series polynomial, which involves the concept of non-cross partition not being repeated. Among the various forms of power series, the commonly used power series is Zeta-series, defined as $\text{Zeta}(z) = \sum z^i$. The sequence of moments under the deterministic measure is defined as $M(\mu)(z) = \sum_{k=1}^{\infty} m_k z^k$. The literature has proven that the above R transformation is equivalent to the following equation:

$$M(\mu) = R(\mu) \boxplus \text{Zeta}. \quad (24)$$

It can be proven that the boxed convolution acting on the power series polynomial is the combination of multiplicative-free convolution on each measure, where the boxed convolution of power series $c^{n-1} \text{Zeta}$ represents the convolution of measure $\mu$, as shown in the following formula:

$$M_{\text{prop}} = M_p \boxplus (c^{n-1} \text{Zeta}) \quad (25)$$

also written as follows:

$$c M_{\text{prop}} = (c M_p) \boxplus \text{Zeta}. \quad (26)$$

It can be found that, in fact, the above equation is the moment-cumulant formula, which is equivalent to Eq. 21. Thus, in the calculation process, the cumulant is replaced by the coefficient of $c M_p$, and the moment is replaced by $c M_{\text{prop}}$. It is concluded that the calculation process of additive-free convolution is also applicable to multiplicative-free convolution operation.

**CASE STUDIES**

**Data Sources**

Based on Matlab, this study uses the voltage data of 40 buses under the maximum operation mode of the 110-kV distribution network in a certain province in the summer of 2020 to conduct simulation verification. The total duration of voltage data is 5s, the number of each sampled signal components is $M = 40$, and the sampling interval is $\Delta t = 0.01$s, so the high-dimensional random matrix $X_{40 \times 500}$ of source data can be obtained. The specific case description is shown in Table 1. In the following two cases, the high-order moments of the event signal component are obtained based on FPT, and the operation status of the distribution network is analyzed and compared with the classical PCA method (Rong et al., 2019) and the MSR indicator (Zheng et al., 2020) in the commonly used random matrix theory to further verify the effectiveness of the indicators proposed in this study.

**Analysis of Cases**

In this study, the high-dimensional random matrix of source data is $X_{40 \times 500}$, that is, $M = 40$ and $N = 500$, and the sliding time window size is set as $40 \times 60$, that is, $p = 40, m = 60$, so $c = p/m \in (0, 1)$. The index selected in the method based on FPT is the third-order moment $m_3$ of the event signal component $x(n)$. The classical PCA assessment indexes for abnormal state detection are $T^2$ statistics and SPE statistics. The control limit of $T^2$ statistics is $T_{n}$, and $T^2 < T_n$ should be satisfied if the system runs normally; otherwise, it can be considered abnormal. The control limit of SPE statistics is $Q_n$. If the system is running properly, it should meet the SPE < $Q_n$ requirement; otherwise, it can be considered abnormal. In the evaluation of the MSR index based on RMT, and the calculated inner ring radius is 0.52. If MSR falls below the threshold of the inner ring radius, it indicates the occurrence of abnormal events. As the width of the sliding
time window is 60, all moments and MSR indicators in the following figures are 0 before the 60th sampling moment, as hereby stated.

In addition, in the power system, the voltage of each user must be kept at the rated value or within the allowable range of voltage offset. Currently, the percentage of voltage offset at the power supply end of 35 kV and above is defined as ±5% in China.

1) Case 1

By observing the voltage fluctuation of the 40 buses in Figure 2 below, it can be seen intuitively that the voltage unit values of the 40 buses are initially distributed between (0.955, 0.995), which are in a normal level, and obvious drops occur at about the 100th sampling time point, that is, the bus voltage drops at about 1s. Subsequently, the buses’ voltage stops falling around the 300th
sampling time point, that is, the bus voltage gradually stabilized after 3s and successfully reached a new stable state.

In Figure 3, the red dotted lines are control limits $T_\alpha$ and $Q_\alpha$, which is the same in Figure 7. Observing various indicators in the figures of case 1 (i.e. Figures 3, 4, Figure 5), the $T^2$ index in the PCA method changes dramatically at the 101th sampling point and exceeds $T_\alpha$, indicating abnormal status. It climbs to the 300th sampling point and then begins to fall but fails to return to the control limit level, indicating that the power grid tried to re-establish a new balance after the third second, but it does not return to the normal state (not in line with the actual situation). Here, the SPE indicator is similar to the $T^2$ indicator. The MSR index begins to fall significantly at the 100th sampling point and falls below the inner ring radius at the 150th sampling point, indicating that the power grid is in an abnormal state (also not in line with the actual situation) and then gradually recovers to

FIGURE 6 | Original voltage data of 40 buses in case 2.

FIGURE 7 | $T^2$ and SPE indices of case 2.

FIGURE 8 | MSR index of case 2.
above the inner ring radius at the 200th sampling point. In the moment index of each order, obviously, the first-order moment and the second-order moment are not sensitive to voltage changes, while the third-order moment is more sensitive to power grid fluctuations. Therefore, the third-order moment is selected as the final index in case 1, and the case of case 2 is the same as that of case 1. The third-order moment $m_3$ changes around the value of 1.5 under normal circumstances and changes dramatically at the 100th sampling point, almost climbing in a straight line. The value of $m_3$ reaches at 1900, which means that the event signal shows up during 1–1.5 s; so $H_1$ is true. After the 150th sampling point, $m_3$ begins to decline and shows a downward trend, indicating that the power grid tried to establish a new equilibrium state in this period. Then $m_3$ fluctuates at the 300th sampling point and recovers to the normal level after the 400th sampling point, representing that the power grid successfully establishes a new stable state, which means that $H_3$ is true. Based on the above analysis, in this case, the third-order moment index $m_3$ can reflect the original data more truly.

2) Case 2

By observing the voltage fluctuation of the 40 buses in Figure 6, it can be seen intuitively that the voltage unit values of the 40 buses are initially distributed between (0.95,1), which are in a normal level. After the 100th sampling point, as the impact load continues to increase, the voltage level drops sharply, and the minimum voltage unit value is as low as 0.82, which is at the abnormal operation level. After the 300th sampling point, the load level remains 3 times that of the original load, and the power grid tries to re-establish a stable state. It can be speculated that due to the limited capacity of the system, the new equilibrium state is not reached, and the voltage level is in continuous oscillation during 3–5 s.

By observing various indicators in the figures of case 2 (i.e. Figures 7, 8, 9), the $T^2$ index in PCA begins to climb gradually at the 101th sampling point, exceeding $T_a$ and indicating abnormal status. The violent fluctuation occurs at the 280th sampling point, showing that the power grid is in an extremely unstable state. Then the $T^2$ index begins to fall and enters an oscillation state without returning to the control limit level, representing that the power grid tried to re-establish a new balance at about 3 s but failed to restore the normal state. Here, the trend of the SPE index and $T^2$ index is slightly different after the 280th sampling point, but generally consistent, indicating that the power grid fails to return to the normal level and is in oscillation. The MSR index begins to decline at the 100th sampling point and then falls below the inner ring radius and remains below the threshold of the inner ring radius, failing to return to normal. The third-order moment $m_3$ changes dramatically at the 100th sampling point and remains at a high level, which means $H_0$ is true, and fluctuates near the 300th sampling point and then remains at a high level at about 4,000, indicating that the power grid tried to establish a new equilibrium in this period but failed, that is, $H_1$ is true.

Through the analysis of the above two cases, it can be found that the performance of the third-order moment index is better than that of the $T^2$ index, SPE index, and MSR index in the small disturbance monitoring of the power grid. When detecting abnormal events, all four indexes can reflect the operation state of the power grid effectively.

CONCLUSION

This study proposes a state evaluation method based on FPT, aiming to evaluate the operation of the distribution network according to high-order moment indexes. Through simulation cases, the following conclusions are obtained:

1) Based on the high-dimensional measurement data of the distribution network, the relevant asymptotic free equation is established, and the high-order moment index is proposed to evaluate the distribution network state, which verifies the feasibility of the proposed evaluation framework based on FPT, applied to the distribution network state analysis.

2) The free probability theory itself tends to be abstract. When calculating the high-order moment index, the moment-cumulant formula can effectively simplify the calculation process of the high-order moment. Theoretically, the index system can be extended to N-order moments, and the selection of specific indicators should be determined based on the actual application scenarios.

3) The proposed high-order moment index is compared with the classical $T^2$ index, SPE index, and the commonly used MSR index. The simulation results show that the above indexes can accurately detect the occurrence of abnormal events in the distribution network, and the high-order moment index performs better than other indexes mentioned in this article when only a small disturbance occurs in the distribution network.

In this study, the free probability theory is applied to the electric power field for the first time, and the proposed evaluation framework can be extended to high-dimensional electric power data processing, such as power dispatching, operation and maintenance control, new energy consumption, reliability
evaluation, and other scenarios to provide decision support. Based on the above analysis, the follow-up work of this study will focus on three aspects: the application of multiple power scenarios, internal performance comparison of high-order moment indicators, and further expansion as well as optimization of the evaluation index system (Xue and Lai, 2016, He et al., 2016, Liu et al., 2016, Wang et al., 2019, Chen et al., 2017, He et al., 2017b, Zhang et al., 2018, Xue, 2015).

**DATA AVAILABILITY STATEMENT**

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

**REFERENCES**

Cao, Y., Meng, H., Zhao, L., Wang, T., and Chen, S. (2007). Comprehensive Analysis of Power System Cognition. Power Grid Techn. 41 (04), 1165–1173. doi:10.13335/j.1000-3673.pst.2007.1956

He, X., Qiu, C., Ai, Q., Chu, L., and Ling, Z. (2017). Research on Spatial and Temporal Big Data of Distribution Network Based on Random Matrix Theory. Power supply and consumption 34 (06), 14–19+7. doi:10.19241/j.cnki.1006-6357.2017.06.003

He, X., Qiu, R. C., Ai, Q., Chu, L., and Ling, Z. (2016). Designing for Situation Awareness of Future Power Grids: An Indicator System Based on Linear Eigenvalue Statistics of Large Random Matrices. IEEE Access 4, 3557–3568. doi:10.1109/access.2016.2581838

Jain, T., Ghosh, D., and Mohanta, D. K. (2019). Augmentation of Situation Awareness by Fault Passage Indicators in Distribution Network Incorporating Network Reconfiguration. Prot. Control. Mod. Power Syst. 4 (4), 323–330. doi:10.1186/s41601-019-0140-6

Li, H., Deng, J., Feng, P., Pu, C., Arachchige, D. D., and Cheng, Q. (2021). Short-term Nacelle Orientation Forecasting Using Bilinear Transformation and IEC61850 Framework. Front. Energ. Res. 9, 697. doi:10.3389/fenrg.2021.769028

Li, H., Deng, J., Yuan, S., Feng, P., and Arachchige, D. D. Monitoring and Identifying Wind Turbine Generator Bearing Faults Using Deep Belief Network and EWMA Control Charts. Front. Energ. Res. 9, 770. doi:10.3389/fenrg.2021.799039

Liu, Z., Wang, J., Wang, P., Weng, H., and Li, Z. (2021). A Novel Fault Section Locating Method Based on Distance Matching Degree in Distribution Network. Prot. Control. Mod. Power Syst. 6 (2), 253–263. doi:10.1186/s41601-021-00194-y

**AUTHOR CONTRIBUTIONS**

JZ selected and studied the sources, designed the structure of the manuscript, and wrote the first draft of the manuscript. BW and HW contributed with supervision over the study of the literature and the writing of the manuscript. All authors contributed to manuscript revision and read and approved the submitted version.

**FUNDING**

This research is supported by a project from the National Natural Science Foundation of China (51907096).

Ling, Z., Qiu, R., and He, X. (2018). A New Approach of Exploiting Self-Adjoint Matrix Polynomials of Large Random Matrices for Anomaly Detection and Fault Location. arXiv:1802.03503v1.

Liu, W., Zhang, D., and Wang, X. (2016). Power System Transient Stability Analysis Based on Random Matrix Theory. Proc. CSEE 36 (18), 4854–4864. doi:10.13334/j.0258-8013.pcsce.160779

Zhang et al. State Evaluation of Distribution Network

Liu, Y., Chen, S., Cong, Z., Jiang, Q., Yan, Y., and Jiang, X. (2021). Key Technology and Application Prospect of Digital Twin in Power Equipment Industry. High Voltage Tech. 47 (05), 1539–1554. doi:10.13336/j.1003-6520.hve.20210194

Liu, Z. (2010). Modern Signal Processing. Beijing: China Electric Power Press.

Li, F., Wang, B., and Dong, X. (2020). Power Vision Edge Intelligence: Power Depth Vision Acceleration Technology Driven by Edge Computing. Power Syst. Tech. 44 (06), 2020–2029. doi:10.13335/j.1000-3673.pst.2019.2382

Mousavizadeh, S., Alahyari, A., Ghodsinya, S. R. M., and Haghifam, M. (2021). Incorporating Microgrids Coupling with Utilization of Flexible Switching to Enhance Self-Healing Ability of Electric Distribution Systems. Prot. Control. Mod. Power Syst. 6 (3), 300–310. doi:10.1186/s41601-021-00197-9

Ouyang, S., and Xu, W. (2021). Power Grid Fault Identification and Cause Identification Method Considering External Environmental Information. High Voltage Tech. 47 (03), 1075–1082. doi:10.13336/j.1003-6520.hve.20200106

Qi, B., Zhang, P., Zhang, S., Zhao, L., Wang, H., Huang, M., et al. (2021). Application Status and Development Prospect of Digital Twin Technology in State Evaluation of Power Transmission and Transformation Equipment. High Voltage Tech. 47 (05), 1522–1538. doi:10.13336/j.1003-6520.hve.20210093

Rong, Z., Qi, B., Zhang, P., Li, C., Yang, Y., Gu, C., et al. (2019). Fast Identification of Abnormal State of Oil Chromatography Online Monitoring Device Based on Nuclear Principal Component Analysis. High voltage technol. 45 (10), 3308–3316. doi:10.13336/j.1003-6520.hve.20190624006

Ryan, O., and Debabh, M. (2007). “Free Deconvolution for Signal Processing Applications[C],” in Proceeding of the Information Theory, 2007. ISIT 2007. IEEE International Symposium on, Nice, France, June 2007 (IEEE), 1846–1850.
