A joint configuration method for multiple flexible resources in low carbon distribution networks based on massive scene dimensionality reduction

Zhen Zheng¹,*,†, Guanhua Zhang¹, Duhong Wang¹, Kai Mou¹, Zhizhuo He¹ and Haijun Xing²

¹State Grid Shanghai Qingpu Electric Power Supply Company, No. 6155, Wai-Qingsong Highway, Qingpu District, Shanghai, China, 201700; ²Shanghai University of Electric Power, Shanghai, 200090

Abstract

Due to the extensive access of renewable sources in distribution networks, this research proposes a joint configuration method for multiple flexible resources in low carbon distribution networks based on massive scene dimensionality reduction to maximize the access capacity of renewable sources. The dimensionality reduction clustering is carried out on the wind-light-load mass high-dimensional scenes by the principal component Gaussian mixture clustering algorithm, and the typical scene set of wind-power-loads is obtained. Hereafter, a joint configuration model for multiple flexible resources of distribution network for massive scenarios is constructed, and using the second-order cone relaxation technique to convert the non-convex constraints in the model to convex processing. Last, a Portugal 54 distribution system is employed to verify the proposed method.

Keywords: low carbon distribution network; principal component analysis method; Gaussian mixed clustering; multiple flexible resources; joint configuration

1 INTRODUCTION

In order to reduce dependence on fossil fuels, China has begun to vigorously develop renewable sources [1]. However, the output of renewable sources is stochastic and difficult to accurately predict, which has a significant impact on the results of flexible resource configuration. Joint configuration of multiple flexible resources while considering the mutual influence between various flexible resources can achieve maximum resource utilization. Therefore, studying the joint configuration method of multiple flexible resources in distribution networks based on massive scene dimensionality reduction is of great significance.

Owing to the development of distribution networks and the popularization of new flexible loads, such as electric vehicle (EV), microgrid group (MMG), 5G base stations, adjustable loads, and data centers, the awareness of exploring these demand side flexible resources has increased. Many studies have proposed a methodological framework for optimizing the configuration of electric vehicle charging stations and related operational plans [2–5]. When renewable resources are deployed, the microgrid can operate in an autonomous or nonautonomous mode. The autonomous operation of microgrids requires careful planning, which has attracted the attention of many researchers [6–8]. In order to balance the power supply and consumption during peak demand periods, people are committed to reducing peak demand by considering adjustable load regulation. In addition, adjusting interruptible or transferable loads has enormous potential for various system costs [9–11]. With the rapid growth of cloud computing, the deployment of data centers and 5G base stations continues to increase, resulting in significant power consumption and higher operating costs. In order to reduce the long-term energy consumption associated with these new flexible loads,
many configuration methods have been proposed. However, the above research did not consider the joint configuration of multiple flexible resources in distribution networks, and a joint configuration method for multiple flexible resources should be developed.

In existing research, k-means algorithm is often used to generate typical scenarios. Peng et al. [12] proposed a multi-scenario analysis method for parallel iterative binary k-means clustering, which has a smaller sum of squared Euclidean distances between cluster centers and a shorter clustering time. Luo and Shi [13] utilized the k-means clustering algorithm to reduce typical scenarios in different regions while fully preserving load space and temporal information, effectively reducing the complexity of clustering. Shi et al. [14] used improved fuzzy C-means to cluster the temporal scenes of wind-light-load to solve its uncertainty, and synthesized the temporal full scene based on the probability of the temporal scenes. Lu et al. [15] proposed an improved feature selection and adaptive PSO-k-means clustering algorithm, which can quickly capture the time series of new energy power generation, and get the scene containing most of the key information. In the above research, k-means algorithm or its improved algorithm is mostly used for clustering. The time complexity of the k-means algorithm increases with the increase in the number of samples, which makes it difficult to apply to clustering analysis of massive high-dimensional scenes.

Based on this, our study proposes a joint configuration method for multiple flexible resources in distribution networks based on massive scene dimensionality reduction. First, the principal component-Gaussian hybrid clustering method is used to realize the dimensionality reduction clustering of massive scenes, and on the basis of fully retaining the original information, the typical operation scenes of wind and light loads are obtained. Then, on the typical scenarios generated by the dimensionality reduction clustering, the cooperative planning model of distribution network source, network, load and storage for massive scenarios is constructed with the objective of minimizing the total cost of distribution network planning. The presence of 0–1 variables and square terms in the model leads to a nonconvex nonlinearity, which is transformed into a second-order cone planning model based on cone theory. Finally, the typical daily curve obtained by using principal component-Gaussian hybrid clustering on the Portugal 54-node arithmetic example is substituted into the operation layer in the source-grid-load-storage planning model to verify the effectiveness of the method proposed in this paper. The main contributions of this study are summarized as follows:

1. A principal component Gaussian mixture clustering algorithm was proposed to achieve effective dimensionality reduction clustering in massive high-dimensional wind-light-load scenes.
2. A joint configuration method for multiple flexible resources in distribution networks has been proposed to maximize the access capacity of renewable energy.

The remaining parts of this study are organized as follows: section II introduces the dimensionality reduction clustering method for massive scenes, section III introduces a detailed configuration model, section IV analyzes the results, and section V summarizes this study.

2 MASSIVE SCENE DIMENSIONALITY REDUCTION CLUSTERING METHOD

The massive operation scenario in this article refers to a large number of typical days composed of wind power, photovoltaic, and load operation data in the distribution network. Conventional clustering algorithms are difficult to cluster high-dimensional data. Therefore, this article first uses principal component analysis to reduce the dimensionality of massive operating scenarios of wind power, photovoltaic, and loads, extract important information from them, and then cluster them to improve efficiency.

2.1 Principal components

The main idea of the principal component analysis method (PCAM) is to use a k-dimensional feature orthogonal basis to replace the original n-dimensional data, in order to reduce the data from n-dimensional to k-dimensional. Taking the high-dimensional output data of photovoltaics as an example, the specific steps of PCAM are introduced [16]. Assuming that the original photovoltaic output data contains $N_1$ $T_1$-dimension vectors $S$, the specific steps of PCAM are as follows:

1) Normalization of raw photovoltaic output data:

$$S_{nt}^* = \frac{S_{nt} - \bar{S}_n}{\sqrt{M_n}}, \quad n = 1, 2, ..., N_1, \quad t = 1, 2, ..., T_1$$

$$\bar{S}_n = \frac{1}{T_1} \sum_{t=1}^{T_1} S_{nt}$$

$$M_n = \text{Var}\left(\frac{1}{T_1 - 1} \sum_{t=1}^{T_1} (S_{nt} - \bar{S}_n)^2\right), \quad n = 1, 2, ..., N_1$$

where $S_{nt}^*$ represents the photovoltaic output data after vector S normalization; $\bar{S}_n$ and $M_n$ represents the first and second moments of the vector $S$, respectively.

2) Calculate the normalized covariance coefficient of data and form a covariance matrix:

$$R = \begin{bmatrix}
    r_{11} & r_{12} & \cdots & r_{1N_1} \\
    r_{21} & r_{22} & \cdots & r_{2N_1} \\
    \vdots & \vdots & \ddots & \vdots \\
    r_{N_11} & r_{N_12} & \cdots & r_{N_1N_1}
\end{bmatrix}$$

where $r_{mn}$ represents the covariance between the $m$-th and $n$-th elements in two $T_1$-dimensional vectors.

3) Calculate the eigenvalues and eigenvectors of $R$:

$$\varphi = U^T R U$$

where $\varphi = \text{diag}(\varphi_1, \varphi_2, ..., \varphi_L)$ is the eigenvalue of the covariance matrix $R$; $U$ is the orthogonal matrix.
4) By performing a linear transformation on matrix $U$, the principal component matrix $Z$ can be obtained:

$$Z = U^T S.$$  

(6)

### 2.2 Gaussian mixture clustering algorithm

Existing research commonly uses k-means clustering algorithms to generate typical scenes, but the k-means algorithm has the advantages of being fast and simple, but it also has limitations such as limited accuracy and inflexible class shapes. Based on this, our article proposes a principal component-Gaussian hybrid clustering algorithm. The biggest difference between this algorithm and the Gaussian clustering algorithm is that the massive data are downscaled and covariance matrix

$$
\text{Objective function to make its derivative $\phi$ as an example. Take the logarithm of the GMM general model:}
$$

$$
\ln \Pr (x | \mu_w, \Sigma_w) = \sum_{n=1}^{N} \ln \sum_{w=1}^{W} \phi_w N_w (x | \mu_w, \Sigma_w).$$

(13)

Step 4: Check whether the parameter or logarithmic likelihood function converges. If not, return to step 1 until it does converge.

In order to evaluate the quality of clustering results, based on the dispersion degree $B$ [13] and intra class concentration degree $B_r$ [16], this paper proposes the comprehensive clustering index $O(R)$, which is calculated using the following equation:

$$
O(R) = \frac{(B)^2}{R} \sum_{r=1}^{R} (B_r)^2 / [n_r (N - R)].
$$

(14)

The larger the comprehensive clustering index, the more compact the scenes in each category, the more obvious the boundaries between classes, and the better the clustering effect.

### 3 MODEL OF JOINT CONFIGURATION FOR MULTIPLE FLEXIBLE RESOURCES

In this section, a joint configuration model for multiple flexible resources based on massive scene dimensionality reduction is developed. This configuration approach determines what actions should be taken to investigate the interactivity of multiple flexible resources in the configuration horizon. The objective function of the proposed model is presented with network constraints. The linearization of the proposed configuration model is also analyzed.

#### 3.1 Objective function

Decision variables in the model include: allocated capacity of multiple types of flex resources; real-time output of multiple types of flex resources; and real-time output of WT/PV. The configuration
model takes the maximum access capacity of renewable sources (wind power, photovoltaic) as the objective function which can be denoted as:

\[
\max \left( \sum_{i \in N_{PV}} S_{i}^{PV} + \sum_{i \in N_{WT}} S_{i}^{WT} \right) \tag{15}
\]

where \( N_{PV} \) is the set of nodes for newly configured photovoltaic power; \( N_{WT} \) is the set of nodes for newly configured wind power; \( S_{i}^{PV} \) is the newly configured photovoltaic capacity for node \( i \); and \( S_{i}^{WT} \) is the newly configured wind power capacity for node \( i \).

### 3.2 Constraints

#### 3.2.1 Power balance constraints:

\[
\begin{align*}
\sum_{j} P_{ij,t} - \sum_{k} (P_{ji,t} - R_{ik}I_{ik,t}) &= I_{P,t} \\
\sum_{j} Q_{ij,t} - \sum_{k} (Q_{ji,t} - X_{ik}I_{ik,t}) &= I_{Q,t}
\end{align*}
\]

where \( P_{ij,t} \) and \( Q_{ij,t} \) is the active and reactive power flow of branch \( ij \); \( R_{ij,t} \) and \( X_{ij,t} \) denotes the resistance and reactance of branch \( ij \); \( I_{P,t} \) and \( I_{Q,t} \) refers to the active and reactive power injected from node \( i \); \( I_{ij,t} \) describes the squared value of current amplitude for branch \( ij \).

#### 3.2.2 Node voltage and branch current constraints:

\[
\begin{align*}
V_{i,t}^{2} &\leq V_{i,t}^{2} \leq V_{i,\text{Max}}^{2} \\
X_{ij,\text{Max},ij}^{2} &\geq I_{ij,t}^{2}
\end{align*}
\]

where \( V_{i,t}^{2} \) represents the square value of voltage at node \( i \) with upper and lower bounds; \( V_{i,\text{Max}}^{2}, V_{i,\text{Min}}^{2} \) refers to the square value of maximum and minimum voltage of the node \( j \); \( I_{ij,t}^{2} \) refers to the square value of current of branch \( ij \).

#### 3.2.3 Constraints of PV and WT:

\[
\begin{align*}
0 &\leq P_{i,t}^{PV} \leq P_{i,t}^{PV,\text{max}} \\
0 &\leq P_{i,t}^{WT} \leq P_{i,t}^{WT,\text{max}} \\
Q_{i,t}^{PV} &= P_{i,t}^{PV} \tan \varphi_{PV} \\
Q_{i,t}^{WT} &= P_{i,t}^{WT} \tan \varphi_{WT}
\end{align*}
\]

where \( P_{i,t}^{PV} \) and \( P_{i,t}^{WT} \) are the active output of wind power and photovoltaic power; \( P_{i,t}^{PV,\text{max}} \) and \( P_{i,t}^{WT,\text{max}} \) are the maximum output power of wind power and photovoltaic power; \( Q_{i,t}^{PV} \) and \( Q_{i,t}^{WT} \) are the reactive output of wind power and photovoltaic power, \( \varphi_{PV} \) and \( \varphi_{WT} \) are the power factors of PVG and WTG.

#### 3.2.4 EES constraints:

\[
\begin{align*}
0 &\leq P_{i,t}^{\text{ESS,dis}} \leq \gamma_{i,t} \mu_{\text{ESS}}^{\text{Max}} \\
0 &\leq P_{i,t}^{\text{ESS,cha}} \leq (1 - \gamma_{i,t}) \mu_{\text{ESS}}^{\text{Max}} \\
E_{\text{ESS},i,t+1} &= E_{\text{ESS},i,t} - \mu_{\text{ESS}}^{\text{Max}} P_{i,t}^{\text{ESS,cha}} + P_{i,t}^{\text{ESS,dis}} / \mu_{\text{dis}} \\
E_{\text{ESS},i,t+1} &= E_{\text{ESS},i,t} + \epsilon_{\text{ESS}} \eta_{\text{ESS}}^{\text{Max}} \\
\epsilon_{\text{ESS}} &\leq E_{\text{ESS}} \leq \gamma_{i,t} \mu_{\text{ESS}}^{\text{Max}}
\end{align*}
\]

where \( \mu_{\text{ESS}}, \mu_{\text{dis}} \) refers to the charging and discharging efficiency of ESS; \( P_{i,t}^{\text{ESS,cha}} \) and \( P_{i,t}^{\text{ESS,dis}} \) describes the charging and discharging power of ESS located at node \( i \); \( \epsilon_{\text{ESS}} \) determines ESS’s quantity of charge at time \( t \).

#### 3.2.5 EV constraints:

\[
\begin{align*}
E_{\text{EV},i,t,ch}^{\text{EV}} + E_{\text{EV},i,t,dis}^{\text{EV}} &\leq 1 \\
E_{\text{EV},i,t,ch}^{\text{EV}} &\in \{0, 1\}, E_{\text{EV},i,t,dis}^{\text{EV}} &\in \{0, 1\} \\
0 &\leq P_{i,t}^{\text{EV}} \leq \eta_{\text{EV}}^{\text{Max}} E_{\text{EV},i,t,ch}^{\text{EV}} S_{i,t}^{\text{EV}} \\
0 &\leq P_{i,t}^{\text{EV}} \leq \eta_{\text{EV}}^{\text{Max}} E_{\text{EV},i,t,dis}^{\text{EV}} S_{i,t}^{\text{EV}}
\end{align*}
\]

where \( N_{i} \) is the number of EVs in the charging station at time \( t \); \( E_{\text{EV},i,t,ch}^{\text{EV}}, E_{\text{EV},i,t,dis}^{\text{EV}} \) is the charging and discharging status of the \( i \)th EV at time \( t \); \( S_{i,t}^{\text{EV}} \) is the charging and discharging status of the \( i \)th EV; \( \eta_{\text{EV}}^{\text{Max}}, \eta_{\text{EV}}^{\text{Max}} \) is the charging and discharging efficiency of EVs.

#### 3.2.6 Micro grid constraints:

\[
\begin{align*}
-P_{G-PCC,\text{max}} &\leq P_{\text{grid},t} \leq P_{G-PCC,\text{max}} \quad \forall m \in M, t \in T \\
-P_{\text{PCC,\text{min}}} &\leq P_{m,t} \leq P_{\text{PCC,\text{max}}} \quad \forall m \in M, t \in T
\end{align*}
\]

where \( P_{\text{grid},t} \) denotes the active power exchanged between microgrids and the distribution network during time period \( t \); \( P_{G-PCC,\text{max}} \) denotes the capacity constraint that can be transmitted by the microgrid cluster; \( P_{m,t} \) denotes the active power exchanged between the microgrid and other microgrids during the operation time; \( P_{\text{PCC,\text{min}}} \) denotes the capacity that can be transmitted between microgrid \( m \) and \( n \). \( P_{m,t}^{\text{min}}, P_{m,t}^{\text{max}} \) is the upper and lower limit of the microgrid load demand during time period \( t \), respectively.

#### 3.2.7 5G constraints:

\[
p_{t,g}^{\text{BS}} = \sum_{m \in \Omega_{\text{BS}}} p_{m,t,g}^{\text{BS}} \quad \forall g \in G, t \in T
\]

\[
p_{m,t,g}^{\text{BS}} \geq 0 \quad \forall g \in G, t \in T
\]
where $P_{m}^{BS}$ denotes the total power consumption of the 5G base station group during time $t$; $\Omega_{BS}$ is the set of installed nodes of 5G base stations; $P_{m,t,g}^{BS}$ denotes the active power absorbed by base station $m$ during time $t$; $P_{m,t}^{air}$ denotes the charging/discharging power of the energy storage battery at base station $m$ during time $t$; $P_{m,t}^{com}$ denotes the power consumption of communication equipment at base station $m$ during time $t$; $P_{AAU}^{m,t}$ denotes the power consumption of air conditioners at the base station $m$ during time $t$; $P_{m,t}^{com}$ denotes the total power consumption of the communication equipment in the base station during time $t$; $P_{AAU}^{m,t}$ denotes the power consumption of BBU/AAU at base station $m$ during time $t$, respectively; $P_{m,t}^{BBU,Ac}/P_{m,t}^{BBU,Sl}$ denotes the power consumption of BBU/AAU when the base station is in active/standing mode, respectively; $P_{m,t}^{AAU,Ac}/P_{m,t}^{AAU,Sl}$ denotes the AAU power consumption when the base station is in active/standing mode, respectively; $P_{m,t}^{ir}$ denotes the total transmitting power of base station $m$; $\delta_{m}$ is the efficiency of the AAU power amplifier; $U_{m,t}^{bs}$ denotes the switching state of the base station, in which 1 indicates an active state, 0 indicates a sleep state; $P_{m,t}^{ir}$ indicates the transmitting power of base station $m$ during time $t$; $P_{max,m}$ indicates the maximum transmitting power of base station $m$; $c_{m}^{m,u}$ refers to the connections between the customer and base stations. A value of 0 indicates that no communication relationship has been established between the base station and the customer, while a value of 1 represents the opposite; $N_{u}$ is the set of neighboring base stations of the customers.

### 3.2.8 Adjustable load constraints:

$$P_{IL,min} \leq P_{i,t}^{IL} \leq P_{IL,max}$$

where $P_{i,t}^{IL}$ is the load value at time $t$; $P_{IL,min}/P_{IL,max}$ is the minimum/maximum interruptible load; $\delta(t)$ represents the control command at time $t$; $\Omega_{IL}$ is the set; $\Delta T$ is the operation period; $t_{cut,max}$ is the maximum interruption time; $t_{off, min}$ is the minimum time between two interruptions; $t_{onIL}$ is the maximum interruption time.

### 3.2.9 Data center constraints:

$$p_{i,t}^{DC} = p_{i,t}^{TT} + p_{i,t}^{IT} + p_{i,t}^{L}$$

$$p_{i,t}^{UPS} = \frac{1}{\eta_{rec,gs}} \left( E_{on,t} + \frac{p_{i,t}^{DC}}{\eta_{inv,ls}} \right)$$

$$\begin{align*}
\left\{ \begin{array}{l}
p_{i,t}^{TT} = \sum_{k=1}^{N_{k}} M_{i,k} p_{k,t}^{server} \\
p_{i,t}^{server} = p_{fix,server} + p_{CPU,i,t}
\end{array} \right.
\end{align*}$$

$$\begin{align*}
\left\{ \begin{array}{l}
p_{i,t}^{CPU} = C_{1} \sum_{k=1}^{N_{s}} b_{i,k,t,s} \left( \frac{p_{CPU,i,t}}{p_{CPU,i,t}} \right)^{2} \\
0 < b_{i,k,t,s} \leq a_{i,k,t,s} M
\end{array} \right.
\end{align*}$$

$$\begin{align*}
\left\{ \begin{array}{l}
d_{i,t}^{s} - (1 - a_{i,k,t,s}) M \leq b_{i,k,t,s} \\
d_{i,t}^{s} + (1 - a_{i,k,t,s}) M \leq b_{i,k,t,s}
\end{array} \right.
\end{align*}$$

$$\begin{align*}
\left\{ \begin{array}{l}
p_{i,t}^{DC} = p_{i,t}^{PUE} p_{i,t}^{TT} \\
p_{i,t}^{PT} + p_{i,t}^{L} = (\eta_{PU} - 1) P_{i,t}^{TT} \\
\end{array} \right.
\end{align*}$$

$$\left\{ \begin{array}{l}
d_{i,t} = \sum_{\delta=1}^{N_{k}} \lambda_{\delta,i,t} \\
\sum_{k=1}^{N_{i}} \left( \mu_{i,k,t} - d_{i,k,t} \right) \geq \frac{1}{\Delta_{\rho}} \forall i, \forall \rho, \forall t
\end{array} \right.
\end{align*}$$

$$\Delta_{\lambda_{i,t}} = \sum_{\delta=1}^{N_{k}} \lambda_{\delta,i,t} - d_{i,t}$$

$$\begin{align*}
\left\{ \begin{array}{l}
E_{i,t} = E_{i,t} - 1 + \Delta_{\lambda_{i,t}} \Delta t \\
E_{i,t} = E_{i,t} + 0 \\
E_{i,t} \geq 0 \\
0 \leq \sum_{\rho} E_{i,t,\rho} \leq E_{i,t} \max \forall i, \forall t
\end{array} \right.
\end{align*}$$

where $P_{i,t}^{DC}$ is the power consumption of the data center $i$ at time $t$; $P_{i,t}^{TT}$ is the power of IT equipment inside the data center; $P_{i,t}^{CO}$ is
the power consumption of air conditioning equipment; $P_{i}^{\text{dis}}$ is the internal power consumption of the data center; $P_{i}^{\text{UPSe}}$ is the active power supplied by the UPS at time $t$, which represents the load demand of the whole data center; $E_{\text{on},i,t}$ is the active power of the battery charging in the UPS at time $t$; $\eta_{\text{rec},i}$ is the efficiency of the rectifier; $\eta_{\text{inv},i}$ is the efficiency of the inverter; $P_{k}^{\text{server}}$ is the active power consumed by each type $k$ server; $M_{i,k}$ is the number of type $k$ servers; $N_{i}$ is the number of server types; $P_{k}^{\text{fixed}}$ is the fixed power consumption of type $k$ servers; $P_{k}^{\text{CPU}}$ is the CPU power consumption of type $k$ servers; $C_{i}$ is the CPU power consumption coefficient; $b_{i,k,t}$ is the variable associated with the data load processed by type $k$ servers; $N_{s}$ is the number of server frequency blocks; $f_{i,k,t}$ is the s-block operating frequency of type $k$ server CPUs in the data center; $d_{i,k,t}$ denotes the data load processed by each type of server; $a_{i,k,t}$ denotes the block operating frequency flag of CPUs; $M$ is a constant; PUE is a parameter that is defined as the ratio of total data center energy consumption to server energy consumption; $d_{i,k,t}$ refers to the $\rho$ type data load processed by the data center; $\lambda_{i,b,t}$ is the data load transmitted from the front-end server to the data center; $N_{s}$ is the total number of front-end servers; $d_{i,k,t}$ is the $\rho$ type data load processed by each server of type $k$; $D_{i}$ is the latency tolerance time of latency-sensitive data load; $\Delta \lambda_{i,b,t}$ is the variation of data load owing to time transfer; $E_{i,k,t}$ is the data load of type $\rho$ stored in the data center; $E_{i,max}$ is the maximum data load that is allowed to be stored.

### 3.3 The MISOCP transformation

Owing to the nonconvex nonlinear characteristics of the proposed planning model, a second order cone relaxation is applied. Subsequently, the power flow constraint would be reformulated as,

$$u_{i,t}u_{j,t} \geq R_{ij,t}^2 + T_{ij,t}^2. \quad (45)$$

This equation is then deformed to obtain the second order cone form,

$$\frac{u_{i,t} + u_{j,t}}{2} \geq \sqrt{R_{ij,t}^2 + T_{ij,t}^2 + \left(\frac{u_{i,t} - u_{j,t}}{2}\right)^2}. \quad (46)$$

The power balance can be transformed as,

$$P_{i,t} = \sum_{j \in C(i)} \alpha_{ij,t} \left[ G_{ij}u_{i,t} - R_{ij}G_{ij} - T_{ij}B_{ij} \right]$$

$$Q_{i,t} = \sum_{j \in C(i)} \alpha_{ij,t} \left[ -B_{ij}u_{i,t} + R_{ij}B_{ij} - T_{ij}G_{ij} \right]. \quad (47)$$

The node voltage and branch current constraint equations are transformed into,

$$\begin{cases} (V_{\min})^2 \leq u_{i,t} \leq (V_{\max})^2 \quad \text{and} \quad I_{ij,t} \leq (I_{ij})^2. \quad (48) \end{cases}$$

**DESS constraints can be linearized as,**

$$\begin{align*}
\gamma_{\text{ch}}^{\text{e},t} + \gamma_{\text{dis}}^{\text{e},t} & \leq 1, & \gamma_{\text{ch}}^{\text{e},t}, \gamma_{\text{dis}}^{\text{e},t} & \in \{0, 1\} \\
0 & \leq p_{\text{ch}}^{\text{e},t} \leq N_{i}^{\text{ESS},\text{max}}, & 0 & \leq p_{\text{dis}}^{\text{e},t} \leq N_{i}^{\text{ESS},\text{max}} \\
0 & \leq p_{\text{dis}}^{\text{e},t} \leq \gamma_{\text{ch}}^{\text{e},t}M, & 0 & \leq p_{\text{dis}}^{\text{e},t} \leq \gamma_{\text{dis}}^{\text{e},t}M \\
E_{\text{ESS}}^{\text{e},t+\Delta t} & = E_{\text{ESS}}^{\text{e},t} + \eta_{\text{ch}}\Delta t_{\text{ch}}^{\text{e},t} - \frac{1}{\eta_{\text{dis}}}\Delta t_{\text{dis}}^{\text{e},t} \\
E_{\text{min}}^{\text{e},t} & \leq E_{\text{ESS}}^{\text{e},t} \leq E_{\text{max}}^{\text{e},t}. \quad (49)
\end{align*}$$

Similarly, the EV charging stations constraints are linearized as,

$$\begin{align*}
\varepsilon_{\text{CS}}^{\text{tch},t} + \varepsilon_{\text{tch},t}^{\text{dis}} & \leq 1, & \varepsilon_{\text{CS}}^{\text{tch},t}, \varepsilon_{\text{tch},t}^{\text{dis}} & \in \{0, 1\} \\
p_{\text{CS}}^{\text{tch},t} & \leq E_{\text{EV}}^{\text{tch},t}, & 0 & \leq p_{\text{CS}}^{\text{tch},t} \leq E_{\text{EV}}^{\text{tch},t} \\
0 & \leq p_{\text{CS}}^{\text{tch},t} \leq E_{\text{EV}}^{\text{tch},t}M, & 0 & \leq p_{\text{CS}}^{\text{tch},t} \leq E_{\text{EV}}^{\text{tch},t}M \\
0 & \leq p_{\text{CS}}^{\text{tch},t} \leq \eta_{\text{ch}}E_{\text{EV}}^{\text{tch},t} \eta_{\text{ch}}, & 0 & \leq p_{\text{CS}}^{\text{tch},t} \leq \frac{1}{\eta_{\text{dis}}}E_{\text{EV}}^{\text{tch},t} \eta_{\text{dis}} \\
0 & \leq p_{\text{CS}}^{\text{tch},t} \leq \eta_{\text{ch}}E_{\text{EV}}^{\text{tch},t} \eta_{\text{dis}} \\
0 & \leq p_{\text{CS}}^{\text{tch},t} \leq \frac{1}{\eta_{\text{dis}}}E_{\text{EV}}^{\text{tch},t} \eta_{\text{dis}}. \quad (50)
\end{align*}$$

**Figure 1. The initial network.**
Figure 2. The relationship between the degree of feature retention and the dimension of retained features.

Figure 3. The comprehensive clustering index for different classification numbers.

4 CASE STUDY AND ANALYSIS

4.1 Case setting

For the case and result analysis of this study, a Portugal 54 bus distribution system is employed to validate the effectiveness of the proposed configuration approach. The presented method in this study is implemented based on the YALMIP optimization toolbox with MATLAB R2021a, and solved by GUROBI. The overall experimental process is computed on a PC with a processor of 12th Gen Intel (R) Core i5–12500 of 3.00 GHz and 16 GB RAM.

Herein, the employed flexible resources are DESS, EV charging stations, microgrid groups, adjustable load, data centers, and 5G base stations. The inputs of this case study are load and line parameters, substation capacity, flexible resource parameters, and candidate node of renewable generation. Subsequently, the result of this case study is the optimal configuration scheme of multiple flexible resources. The initial network to be configured is shown in Figure 1, 71.4 MW of total load is deployed in this test system.

Table 1. Comparison of computing time of different clustering methods.

<table>
<thead>
<tr>
<th>Feature dimension</th>
<th>Number of cluster classifications</th>
<th>Calculate times/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>k-means</td>
<td>Method proposed in this research</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>1.36</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>7.07</td>
</tr>
<tr>
<td>10</td>
<td>72</td>
<td>25.16</td>
</tr>
</tbody>
</table>

Figure 4. Typical day for clustering results.

4.2 Massive scene dimensionality reduction analysis

The historical data of wind speed, light intensity, and conventional load in the distribution network are based on the data in reference [16]. Due to the high dimensionality of wind speed, light intensity,
Table 2. Maximum access capacity of renewable sources of three configuration schemes.

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Maximum access capacity of WT</th>
<th>Maximum access capacity of PV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme 1</td>
<td>27.2 MW</td>
<td>14.6 MW</td>
</tr>
<tr>
<td>Scheme 2</td>
<td>29 MW</td>
<td>16.8 MW</td>
</tr>
<tr>
<td>Scheme 3</td>
<td>30.3 MW</td>
<td>18 MW</td>
</tr>
</tbody>
</table>

Table 3. The installation of multiple flexible resources.

<table>
<thead>
<tr>
<th>Device</th>
<th>Installation node</th>
<th>Scheme 1 Capacity (MW)</th>
<th>Scheme 2 Capacity (MW)</th>
<th>Scheme 3 Capacity (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EV</td>
<td>7</td>
<td>1.8</td>
<td>1.7</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>1.5</td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>49</td>
<td>0.9</td>
<td>1.0</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>1.8</td>
<td>1.8</td>
<td>1.7</td>
</tr>
<tr>
<td>MMG</td>
<td>8</td>
<td>1.0</td>
<td>1.0</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>0.8</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.5</td>
<td>0.6</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>0.9</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>Adjustable load</td>
<td>3</td>
<td>1.4</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>1.2</td>
<td>1.2</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>48</td>
<td>0.7</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>1.4</td>
<td>1.4</td>
<td>1.3</td>
</tr>
<tr>
<td>Data center</td>
<td>4</td>
<td>1.8</td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.5</td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>0.9</td>
<td>1.0</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>43</td>
<td>1.9</td>
<td>1.9</td>
<td>1.8</td>
</tr>
<tr>
<td>SG station</td>
<td>5</td>
<td>0.7</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>0.6</td>
<td>0.6</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>0.7</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>EES</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1.625 MW/5 MWh</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0.975 MW/3 MWh</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0.5 MW/1.8 MWh</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0.75 MW/2.3 MWh</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>0.48 MW/1.6 MWh</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0.325 MW/1 MWh</td>
</tr>
</tbody>
</table>

and conventional loads, the calculation time for direct clustering is longer and the clustering results are not typical enough. Therefore, this research first uses principal component analysis to reduce the dimension of the historical data. The relationship between the feature retention degree of the original data and the feature dimension is shown in Figure 2.

It can be seen from Figure 2 that when the feature dimension is 3, 95% of the data information can be effectively retained, and increasing the feature dimension will not significantly increase the degree of feature retention [18]. Therefore, this research takes the feature dimension as 3 pairs of original data to reduce the dimension.

4.3 Results and analysis of multiple types of flexible resource configuration

In order to analyze and consider the impact of flexible resource cluster interaction on configuration results, three scenarios were set as:

a) Scheme 1: joint configuration not considering interactions of five flexible loads without DESS.
b) Scheme 2: joint configuration considering interactions of five flexible loads without DESS.
c) Scheme 3: joint configuration considering interactions of five flexible loads and DESS.

Table 2 shows various maximum renewable sources access capacity of three configuration schemes. The maximum access...
capacity of renewable sources in scheme 2 is higher than that in scheme 1, indicating that considering interactions of five flexible loads can enhance the access capacity of renewable sources. The maximum access capacity of renewable sources in scheme 3 is higher than that in scheme 2, indicating that DESS can further enhance the access capacity of renewable sources.

Table 3 shows the multiple flexible resources configuration results for scheme 1, scheme 2 and scheme 3. It can be seen that the installation capacity of EV, MMG, adjustable load, data center and 5G station is 6 MW, 3.2 MW, 4.7 MW, 6.1 MW, 2.3 MW, respectively, in scheme 1 and scheme 2, because of the upper investment limit constraints. As for scheme 3, the installation capacity of EV, MMG, adjustable load, data center, 5G station and EES is 5.5 MW, 3.1 MW, 4.4 MW, 5.8 MW, 14.7 MW, respectively.

In order to compare the three configuration schemes more intuitively, the renewable sources penetration rate, network loss, abandonment rate of wind and light, mean voltage deviation, and maximum voltage deviation are displayed in the form of radar charts, as shown in Figure 5. It can be seen from Figure 5 that the configuration results of scheme 3 are superior to scheme 1 and scheme 2 in terms of renewable sources penetration rate, network loss, abandonment rate of wind and light, mean voltage deviation, and maximum voltage deviation. Based on the analysis results, the proposed joint configuration method for multiple flexible resources in the distribution network considering interactions of five flexible loads and DESS can improve the optimality of various indicators.

Configuration Scenario 3 energy storage SOC and output power are shown in Figure 6. From the figure, it can be seen that the energy storage stores energy when the PV output is high and releases energy to provide some power to the surrounding loads when the PV output becomes low.

5 CONCLUSION

This paper proposes a joint configuration method for multiple flexible resources in low carbon distribution networks based on massive scene dimensionality reduction. The dimensionality reduction clustering is carried out on the wind-light-load mass high-dimensional scenes by the principal component Gaussian mixture clustering algorithm. Different types of flexible resources are coordinated so that the access capacity of maximum renewable sources can be mitigated. The subsequent case study analyzes the rationality of the proposed configuration method to maximize the access capacity of renewable sources. On the premise of the analysis above, the key findings are:

• The proposed principal component Gaussian mixture clustering algorithm has higher efficiency compared with the
k-means clustering algorithm and can effectively cluster mass high-dimensional scenes.

• The proposed joint configuration method for multiple flexible resources in the distribution network considering interactions of five flexible loads and DESS cannot not only enhance the access capacity of renewable sources, but also improve the optimality of various indicators.

Future work based on this study may be to investigate the coordination between uniform and non-uniform flexible resources, in order to further maximize the renewable sources access capacity of the prescribed configuration methods.

Author contributions
Zhen Zheng (Conceptualization [Equal], Data curation [Equal], Formal analysis [Equal], Funding acquisition [Equal], Investigation [Equal], Methodology [Equal], Software [Equal], Supervision [Equal], Writing—original draft [Equal]), Duhong Wang (Formal analysis [Equal], Funding acquisition [Equal], Investigation [Equal], Methodology [Equal], Resources [Equal], Software [Equal], Supervision [Equal], Writing—original draft [Equal]), Huajun Xing (Investigation [Equal], Methodology [Equal], Visualization [Equal], Writing—original draft [Equal]), Zhihuo He (Data curation [Equal], Funding acquisition [Equal], Validation [Equal], Visualization [Equal], Writing—original draft [Equal], Kai Mou (Investigation [Equal], Methodology [Equal], Visualization [Equal], Writing—original draft [Equal]), Zhizhuo He (Data curation [Equal], Funding acquisition [Equal], Visualization [Equal], Writing—original draft [Equal], Validation [Equal], Visualization [Equal], Writing—original draft [Equal], Haijun Xing (Investigation [Equal], Methodology [Equal], Project administration [Equal], Resources [Equal], Software [Equal], Supervision [Equal], Validation [Equal], Visualization [Equal]).

Acknowledgements
This study is supported by the State Grid Shanghai Electric Power Company science and technology project Construction and application of ‘green electricity’ index in Yangtze River Delta integrated development Demonstration Zone based on big data of power economy (B30934200003).

REFERENCES

Multiple flexible resources in low carbon distribution networks