Energy Engineering

Tech Science Press

Doi:10.32604/ee.2025.072828

ARTICLE

Distribution Network Partitioning and Distributed Voltage Coordinated Optimization Method under High-Proportion Photovoltaic Penetration

Jian Wang[1], Gongqiang Yang[1,*], Yufeng Sun[2], Gangui Yan[1] and Jie Long[3]

[1]Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Northeast Electric Power University, Ministry of Education, Jilin, 132012, China

[2]State Grid Shandong Electric Power Co., Ltd. Feixian Power Supply, Linyi, 273400, China

[3]Automation Department of Power Dispatching Center, State Grid Gansu Electric Power Co., Ltd., Lanzhou, 730050, China

*Corresponding Author: Gongqiang Yang. Email: 2202300177@neepu.edu.cn

Received: 04 September 2025; Accepted: 03 November 2025; Published: 27 May 2026

ABSTRACT: Given that the power grid partitioning method relying mainly on line reactive power flow information sees frequent changes in partitioning results with reactive power flow fluctuations under high-proportion fixed-power-factor PV-connected distribution networks, and traditional distributed PV collaborative optimization fails to adapt due to such changes, a stable partitioning and distributed PV collaborative optimization method for this scenario is proposed. Firstly, the Gaussian mixture model (GMM) is used to characterize the characteristics of PV reactive power output, obtaining the typical curve of PV reactive power output. Secondly, the Monte Carlo Simulation (MCS) probabilistic power flow calculation is performed to obtain the node voltage distribution of the distribution network. Thirdly, based on the node voltage distribution, the Earth Mover's Distance (EMD) is used to obtain the statistical distance between any two nodes, and this statistical distance is combined with the electrical distance defined by node voltage sensitivity to form a comprehensive electrical distance. Then, the affinity propagation clustering algorithm is applied, and considering the dynamic reactive power margin requirement, the reactive power/voltage partitioning result is obtained. Based on the reactive power partitioning result, a reactive power optimization model is established with the minimum active power loss of the system as the objective function. The optimization model is convexified using the LinDistFlow equation, and the Alternating Direction Multiplier Method (ADMM) is adopted to coordinate the reactive power output of PV inverters in each partition, achieving global optimal voltage control in the distribution network. Finally, the proposed method is verified using the IEEE 33-bus system. The application of this method reduces the system power loss by 35.94%. Compared with the traditional partitioning method, the partitioning variation rate under Scenario 1 is reduced by 54.17% and that under Scenario 2 is reduced by 70.85% when this method is adopted. This fully demonstrates that the partitioning results of the proposed method are stable, and the collaborative optimization method can improve the system voltage stability and reduce the system power loss.

KEYWORDS: Gaussian mixture model; Monte Carlo trend; distribution network partition; distributed voltage control; synchronous ADMM

1 Introduction

1.1 Background and Motivation

As distributed photovoltaic systems are gradually integrated into the grid, the network structure of traditional distribution grids has undergone fundamental changes, presenting numerous challenges to the stable operation of distribution grids [1,2]. On the one hand, since the output of distributed photovoltaic

Copyright © 2026 The Authors. Published by Tech Science Press.
This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

systems is highly influenced by natural conditions, its randomness can impact the power flow in distribution grids, resulting in complex and variable power flow distributions [3]. On the other hand, due to the flexibility in site selection and scale, distributed PV systems often connected to the local grid to reduce network losses and improve energy efficiency, which leads to a growing number of nodes in the distribution grid and a more complex network topology [4–6]. These two factors make it increasingly difficult for traditional centralized voltage control to meet the requirements of precision and speed in regulation, thereby posing challenges to maintaining the stability of distribution grid voltage [7].

1.2 Literature Review

At present, there are mainly three types of voltage control in distribution networks, namely centralized control, local control, and distributed control. The first two have limitations in addressing the voltage stability issue in scenarios with a high proportion of photovoltaic penetration in distribution networks [8]. However, distributed voltage control, which does not rely on a central processing unit, enables each regional controller to independently solve optimization problems and can cope with the frequent fluctuations in regional voltage when a high proportion of distributed photovoltaic is connected to the distribution network [9]. Distributed voltage control involves dividing the distribution network into zones according to certain zoning principles, and then conducting voltage control on each zone independently. Therefore, the prerequisite for achieving distributed voltage control is to carry out reasonable partitioning. In terms of reactive power/voltage zoning in power grids, there are mainly two types. (1) Electrical distances are defined based on the topological structure of the power grid or reactive power/voltage sensitivity, and then zoning is carried out with the goal of minimizing the electrical distance. This paper refers to such methods as "traditional distance methods". (2) A new type of electrical distance is defined based on the system power flow variation, and this distance is used as the weight. By applying the complex network theory and aiming for the optimal modularity function index, partitioning is carried out. This paper refers to such a method as the "reactive power flow distance method". Both types of methods define the electrical distance of the network, but due to the different considerations of electrical distance, the complexity of partitioning and the applicable scenarios vary significantly.

For the traditional distance method, reference [10] defined the electrical distance based on the reactive power/voltage sensitivity between nodes through the full-state embedding method, and partitioned the distribution network using the sensitivity matrix of the full-state embedding method. The reactive power flow distance method takes into account the power flow changes of the line compared with the traditional distance method, and automatically partitions based on the improved modularity function index, without the need for manual pre-specification of the number of partitions [11]. References [12,13] defined the electrical distance based on the reactive power/voltage sensitivity between nodes and the negative logarithm of the sensitivity, and finally partitioned it in the form of Ward distance based on the hierarchical clustering method. Traditional distance methods are basically based only on the topological structure of the power grid for partitioning. The partitioning algorithms are mostly hierarchical clustering method, K-means method, immune algorithm, etc. The partitioning basis is simple and the calculation speed is fast. However, most algorithms require manual pre-designation of the number of zones and do not take into account the changes in power flow in the power grid. This results in a deterioration of the regional voltage control effect based on traditional electrical distance zoning under frequent fluctuations in power flow, making it difficult to adapt to the scenario where a high proportion of distributed photovoltaic power is connected to the distribution network.

For the reactive power flow distance method, reference [14] combines the electrical distance defined by the reactive power/voltage sensitivity between nodes with the reactive power flow, assigns values to

the line weights, and establishes a reactive power transmission network topology diagram based on the complex network theory, and performs partitioning based on the simplified modularity function index. Reference [15], based on the indicators of "node coupling degree and active power matching degree", partitions the distribution network using the spectral clustering algorithm, constructs a two-layer model for energy storage configuration optimization, and improves the operational flexibility and economy of the distribution network.

Compared with the traditional distance method, the reactive power flow distance method takes into account the power flow changes of the line and automatically performs partitioning based on the improved modularity function index, without requiring manual pre-specification of the number of partitions. Therefore, it has received considerable attention in scenarios where a high proportion of distributed photovoltaic power is integrated into the distribution network.

In terms of distributed voltage coordination optimization, in reference [16], in order to achieve the optimal reactive power compensation and active power reduction of photovoltaic power within the region, the autonomous optimization sequence of each region is arranged according to the size of voltage offset, and a distributed voltage control method based on particle swarm optimization algorithm is adopted. However, the inter-zone coordination speed of this method is relatively slow and the communication burden is heavy. Reference [17] aims to reduce the time gap between the steady-state optimization of the system and the real-time control of distributed photovoltaic power generation. The dual sub-gradient method is adopted to address the time gap issue and achieve the optimal power flow within the region. Ref. [18] established an optimization model with the goal of economy, convexized the model through the second-order cone relaxation Method, and adopted the ADMM for distributed computing. However, the traditional ADMM algorithm is serial computing, which has a limited computing speed in solving problems, and the second-order cone relaxation increases the computational load of model solving.

1.3 Research Gaps and Contributions

This paper proposes a collaborative optimization method for distribution network zoning and distributed voltage. This method considers the impact of power flow variations in the distribution network caused by the uncertainty of distributed photovoltaic output on grid zoning, and on this basis, conducts distributed voltage control. Firstly, the random variation of reactive power output at the grid connection point caused by the random active power output of PV power is fitted through the Gaussian mixture model. Then, the probabilistic power flow operation is carried out through the Monte Carlo simulation method to obtain the random parameters of the voltage at each node. Secondly, the discrete distribution formed by the random parameters of the node voltage is used to define the comprehensive electrical distance under statistical meaning by combining the bulldozer distance with the electrical connection of the node voltage. Then, by using the Affinity Propagation (AP) clustering algorithm, this electrical distance is taken as the input for automatic partitioning. A reactive power optimization model is established within the partition with the goal of minimizing active power loss. The LinDistFlow equation is adopted to ignore the quadratic terms in the equality constraints, and the optimization model is converted into a convex one, which improves the solution efficiency. By adopting the synchronous ADMM algorithm and coordinating the reactive power output of the PV inverter, the optimal control of the entire distribution network is achieved. The effectiveness of the proposed partitioning method and distributed voltage control method was verified based on the IEEE 33-node system. The flowchart of the method proposed in this paper is shown in Fig. 1.

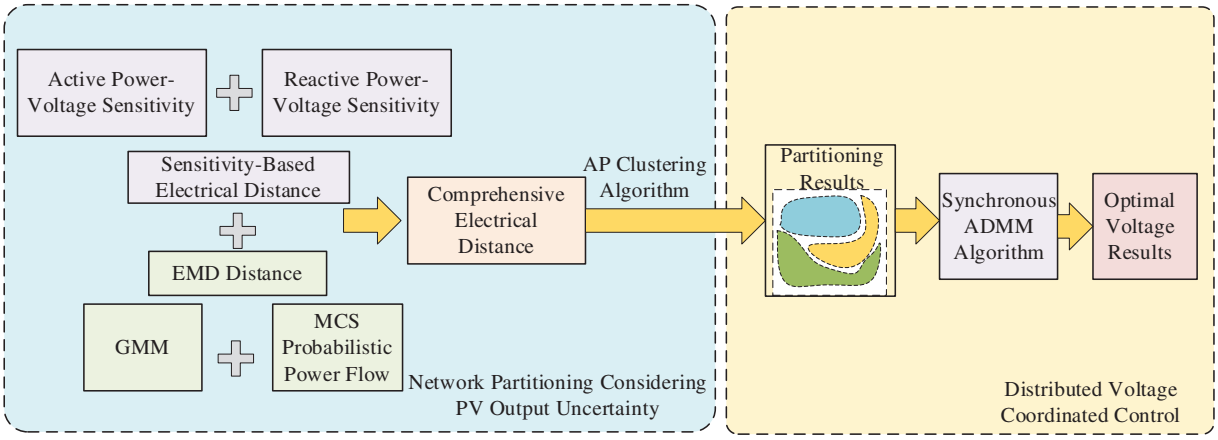

Figure 1: Distribution network partitioning and distributed voltage coordinated method

1.4 Paper Organization

The work of this paper is as follows. Section 2 introduces the Gaussian mixture model and the Monte Carlo simulation method, and constructs the distribution network node voltage dataset. Section 3 introduces the reactive power zoning method based on AP clustering and comprehensive electrical distance. Section 4 introduces the distributed reactive power optimization method based on reactive power partitioning. Section 5 provides cases and conducts an analysis of them. Section 6 summarizes the full text and outlines the future research directions.

2 Construction of a Distribution Network Nodal Voltage Dataset under Fluctuations in Active and Reactive Power Output from Distributed PV

Given the high R/X ratio typically present in practical distribution networks, the stochastic fluctuations in both active and reactive power outputs of distributed PV systems—operating under constant power factor control—can lead to significant voltage variations at the point of common coupling [19]. In this section, a Gaussian mixture model is used to represent the stochastic reactive power output of the grid-connected PV system [20,21]. The active power data are derived from this reactive power GMM based on the strong coupling relation $P = Q/\tan\varphi$ under constant power factor control. Using Monte Carlo Simulation-based probabilistic power flow computation, a mapping is constructed between the stochastic active and reactive power outputs at the distributed PV grid connection point and the resulting node voltages in the distribution network. This approach yields the variation intervals of voltage magnitudes across distribution grid nodes—under random fluctuations in active and reactive power injection—which form the dataset of nodal voltages referred to in this study.

2.1 Gaussian Mixture Model for the Reactive Power Output of Photovoltaic Systems

Since the reactive power output of distributed photovoltaic systems fluctuates randomly with changes in active power output under constant power factor control, this paper utilizes historical data on the active power output of distributed photovoltaic systems and power factor to obtain historical data on photovoltaic reactive power output. The GMM obtained from the historical data of PV reactive power output is a fitting model that includes the characteristics of photovoltaic reactive power output. Based on the GMM, the typical curve of photovoltaic reactive power output can be obtained. When the output characteristics of photovoltaic power remain unchanged, the typical curve of photovoltaic reactive power output also remains unchanged.

The GMM expression is shown in Eq. (1).

$$\begin{cases} f_x(X) = \sum_{m=1}^{M} \omega_m N_m(x; \mu_m, \sigma_m) \\ \sum_{m=1}^{M} \omega_m = 1, \omega_m > 0 \\ N_m(x; \mu_m, \sigma_m) = \dfrac{e^{-\frac{1}{2}(x-\mu_m)^T \sigma_m^{-1}(x-\mu_m)}}{(2\pi)^{w/2} \det(\sigma_m)^{1/2}} \end{cases} \tag{1}$$

In the formula: ω_m, μ_m, and σ_m represent the weight, mean vector, and covariance matrix of the mth Gaussian component, respectively; M denotes the number of Gaussian components; $f_x(X)$ is the joint probability density of X, x is an element of the random variable X; $N_m(\cdot)$ is a W-dimensional multivariate Gaussian distribution; det(·) denotes the determinant operation; T denotes the transpose of a matrix.

Let the joint parameter set of the GMM be denoted as $\Omega = \{\omega_m, \mu_m, \sigma_m; m = 1, 2, \ldots, M\}$. Then, Eq. (1) can also be expressed in probabilistic notation as:

$$X \sim \mathrm{GMM}(\omega, \mu, \sigma) \tag{2}$$

In the formula: $\omega = \{\omega_m; \omega_m \in R, m = 1, 2, \ldots, M\}$ represents a set of M Gaussian component weights; $\mu = \{\mu_m; \mu_m \in R^W, m = 1, 2, \ldots, M\}$ is the set of M Gaussian component mean vectors; $\sigma = \{\sigma_m; \sigma_m \in R^W \times R^W, m = 1, 2, \ldots, M\}$ is the set of M Gaussian component covariance matrices.

The active power output of distributed PV systems can be directly measured. Under constant power factor control, the reactive power output data can be calculated from the active power data and the power factor. Following the computational procedure outlined in Ref. [22], a GMM for the reactive power output at the PV grid connection point can be derived, denoted as $Q_s \sim \mathrm{GMM}(\omega_Q, \mu_Q, {}_Q)$.

2.2 Probabilistic Power Flow Based on Monte Carlo Simulation

Based on the GMM of the stochastic reactive power output at the distributed PV grid connection point, this paper employs MCS for probabilistic power flow calculation to establish a mapping between the stochastic active and reactive power outputs at the PV point of common coupling and the nodal voltages in the distribution network. As a result, a distribution network nodal voltage dataset is derived, denoted as $V_i = \{v_{i1}, v_{i2}, \cdots, v_{is}\}$, where i represents the number of nodes in the distribution network and s indicates the sampling count in the MCS process. Here, v_{is} refers to the voltage value obtained from the power flow calculation of the s-th MCS sample at the i-th node.

The probabilistic power flow equations based on MCS are given as follows:

$$v_{is} \begin{cases} x = g(y) \\ z = h(y) \end{cases}, \quad x = \{x_1, x_2, \cdots, x_s\} \tag{3}$$

In the formula: g represents the nodal power balance equations; h denotes the branch power flow equations; x indicates the input random variables, which may include stochastic outputs of distributed generation, random variations in load, and changes in line operational status, among others; s signifies the number of input random variables, i.e., the sample size; y refers to the nodal voltage magnitudes and phase angles; z represents the active and reactive power flows in the branches.

Based on the active and reactive power outputs of the distributed photovoltaic systems and Eq. (3), we obtain:

$$V_i = \{v_{i1}, v_{i2}, \cdots, v_{is}\} \tag{4}$$

The probabilistic power flow calculation based on MCS utilizes typical reactive power output profiles of PV systems and the corresponding active power profiles derived via the relation $P = Q/\tan\varphi$. This process establishes a mapping between the active/reactive power outputs of PV systems and the nodal voltages in the distribution network, thereby incorporating the characteristics of PV power generation under high penetration scenarios. Since the typical reactive power output profile remains unchanged, the resulting reactive power flow distribution—derived from these typical output curves—also remains invariant as long as the network topology of the distribution system remains unaltered. This invariant reactive power flow distribution can be further employed to analyze the characteristics of nodal voltage distributions.

3 A Reactive Power Partitioning Method Based on Comprehensive Electrical Distance and Affinity Propagation Clustering

Partitioning refers to the division of a power grid into several regions according to specific principles, with the fundamental principle being "strong coupling within regions and weak coupling between regions" [23]. Most existing reactive power flow distance methods aim to minimize the electrical distance when partitioning, meaning that a smaller electrical distance between nodes indicates stronger connectivity and higher coupling, justifying their grouping into the same region.

The electrical distance metric employed in existing reactive power flow distance methods comprises two components. The first is determined by the grid topology and is referred to as the sensitivity distance, which remains largely invariant under changing power flow conditions. The second component is influenced by the reactive power flow on transmission lines. These two parts are combined through a ratio or multiplicative operation to form the composite electrical distance. Significant fluctuations in reactive power flow can lead to considerable variations in this flow-dependent component, consequently causing frequent changes in the partitioning results. Within a fixed grid topology, nodes that are strongly interconnected exhibit closer voltage magnitudes, while those with weaker electrical connectivity show more significant differences in voltage values. Thus, under a constant network structure but varying power flow conditions, if an electrical distance metric can be constructed utilizing the relatively stable relationships between nodes—even amid power flow variations—the resulting partition scheme based on this metric will demonstrate enhanced stability.

In summary, when constructing the electrical distance in this paper, in addition to considering the sensitivity distance, the probability statistical distance is also used to quantify the proximity of the node voltage values. As the relatively stable electrical distance between the two nodes in a statistical sense, this paper refers to it as the statistical distance. Then, by combining the statistical distance with the sensitivity-based electrical distance, a stable and comprehensive electrical distance is obtained. This composite metric accounts for both changes in line power flow and grid structure, and its stability is reinforced by the incorporation of the statistical distance. This paper refers to it as the comprehensive electrical distance.

3.1 Comprehensive Electrical Distance Based on Statistical Distance

3.1.1 Statistical Distance between Nodal Voltages

The probabilistic statistical distance is a mathematical measure used to quantitatively describe the distance between two distinct probability distributions in a probability space. Among these measures, the EMD is one of the most widely applied methods [24].

To quantitatively evaluate the proximity between the discrete distributions formed by the nodal voltage datasets V_i obtained in Section 2, this paper employs EMD to compute the statistical distance between the voltage distributions of any two nodes.

EMD is the minimum average distance required to move data from one distribution to another under optimal path planning. It is used to describe the probability and statistical distance between two discrete distributions P and Q as:

$$D_E(P, Q) = \left(\inf_{\forall(u,v)\in\Gamma(P,Q)} E\left[d(u, v)^2\right]\right)^{1/2} \tag{5}$$

In the formula: $D_E(P, Q)$ is the probability statistical distance; $\inf(\cdot)$ is the lower bound of the function; $E[\cdot]$ is the expectation of the random variable; $d(\cdot)$ is the Euclidean distance, that is $d(x, y) = \|x - y\|_2$, $(u, v) \in \Gamma(P, Q)$ indicates that the edge distribution of (u, v) is P and Q.

Based on the EMD, the statistical distance between any two nodes i and j is calculated according to the following procedure:

(1) The voltage datasets V_i and V_j for nodes i and j, obtained from Section 2.2, are arranged in ascending order to form two discrete distributions, denoted as A and B, where all elements in A and B satisfy:

$$\begin{cases} P_A(v_{ik}) = a_k, & \sum_{k=1}^{Y} a_k = 1 \\ P_B(v_{jk}) = b_k, & \sum_{k=1}^{Y} b_k = 1 \end{cases} \tag{6}$$

In the formula: Y denotes the number of data points in each dataset, which is equal to the number of MCS iterations; v_{ik} and v_{jk} represent the k-th elements in distributions A and B, respectively; $P_A(v_{ik})$ and $P_B(v_{jk})$ respectively represent the probabilities of v_{ik} and v_{jk} occurring in A and B. a_k and b_k represent the probability values of v_{ik} and v_{jk} appearing in A and B, respectively.

(2) The EMD distance between two distributions A and B can be expressed as:

$$D_E(A, B) = \min \sum_{k=1}^{Y} \left\|v_{ik} - v_{jk}\right\|_2^2 \cdot f_{ijk} \tag{7}$$

In the formula: $\|\cdot\|_2$ denotes the Euclidean distance; f_{ijk} represents the transfer quantity from v_{ik} to v_{jk}, which is the variable to be solved. f_{ijk} satisfies:

$$\begin{cases} \sum_{k=1}^{S} f_{ijk} = a_k, & \sum_{k=1}^{S} f_{ijk} = b_k \\ f_{ijk} \geq 0 \end{cases} \tag{8}$$

The statistical distance is the solution of Eq. (6) under the constraint conditions of Eq. (7). The smaller the value, the smaller the statistical gap between nodes i and j, and *vice versa*.

Since $D_E(A, B)$ reflects the statistical distance between two nodes i and j, for ease of understanding, it will be rewritten as, that is $D_E(A, B) = D_E(i, j)$.

3.1.2 Electrical Distance Defined by Voltage Sensitivity between Arbitrary Nodes

Partitioning an electrical grid involves dividing it based on a specific network topology. Therefore, the network structure inherently influences the electrical distance between nodes. Under varying power flow conditions, cases may occur where two nodes are geographically far apart yet exhibit very similar voltage distributions. If statistical distance alone is used as the basis for partitioning, the connectivity of the resulting regions may fail to meet operational requirements. Voltage sensitivity between nodes can reflect the electrical coupling determined by the network architecture [25]. Hence, by combining the sensitivity-based electrical distance with the statistical distance—and appropriately weighting their contributions in the partitioning criterion—the connectivity issues arising from the sole use of statistical distance can be mitigated. Thus, this paper adopts nodal voltage sensitivity to further account for the impact of varying power flow conditions on the electrical distance between nodes.

In distribution networks, the voltage sensitivity at each system node can be derived from the inverse of the Jacobian matrix obtained during power flow calculations:

$$\begin{bmatrix}\Delta P\\ \Delta Q\end{bmatrix} = \begin{bmatrix}A^{P\delta} & B^{PU}\\ C^{Q\delta} & D^{QU}\end{bmatrix}\begin{bmatrix}\Delta\delta\\ \Delta U\end{bmatrix} = J\begin{bmatrix}\Delta\delta\\ \Delta U\end{bmatrix} \tag{9}$$

In the formula: ΔP and ΔQ represent the matrices of changes in active and reactive power injections at the nodes, respectively; $\Delta\delta$ and ΔU represent the matrices of changes in voltage angles and magnitudes at the nodes, respectively; J represents the Jacobian matrix; $A^{P\delta}$, B^{PU}, $C^{Q\delta}$ and D^{QU} are elements of the Jacobian matrix, which are used to characterize the relationship between the changes in node voltage caused by variations in node injection power.

From Eq. (9), the inverse relationship between the variation of node injection power and the variation of node voltage is obtained as:

$$\begin{bmatrix}\Delta\delta\\ \Delta U\end{bmatrix} = J^{-1}\begin{bmatrix}\Delta P\\ \Delta Q\end{bmatrix} = \begin{bmatrix}S^{P\delta} & S^{PU}\\ S^{Q\delta} & S^{QU}\end{bmatrix}\begin{bmatrix}\Delta P\\ \Delta Q\end{bmatrix} \tag{10}$$

In the formula: the sensitivity factors S^{PU} and S^{QU} represent the changes in voltage magnitude caused by a unit active power injection and a unit reactive power injection at the node, respectively; the sensitivity factors $S^{P\delta}$ and $S^{Q\delta}$ denote the changes in voltage phase angle resulting from a unit active power injection and a unit reactive power injection at the node, respectively.

The electrical distance between system nodes can be expressed using the sensitivity factors S^{PU} and S^{QU} as follows:

$$\begin{cases} l_{ij}^{P} = \Delta U_i \Delta U_j^{-1} = S_{ij}^{PU} S_{jj}^{PU^{-1}} \\ l_{ij}^{Q} = \Delta U_i \Delta U_j^{-1} = S_{ij}^{QU} S_{jj}^{QU^{-1}} \end{cases} \tag{11}$$

In the formula: ΔU_i and ΔU_j represent the variations in voltage magnitude at nodes i and j, respectively; l_{ij}^{P} and l_{ij}^{Q} represent the changes in voltage at node i caused by a unit change in voltage magnitude at node j; The influence of node j on node i is positively correlated with the values of l_{ij}^{P} and l_{ij}^{Q}. Higher values indicate stronger coupling between the two nodes.

Considering the mutual influence among all nodes, this paper defines the sensitivity electrical distance as:

$$L(i,j)=\sqrt{\sum_{k=1}^{N}\left[\left(\lg\frac{1}{l_{ik}^{P}}+\lg\frac{1}{l_{ik}^{Q}}\right)-\left(\lg\frac{1}{l_{jk}^{P}}+\lg\frac{1}{l_{jk}^{Q}}\right)\right]^{2}} \tag{12}$$

In the formula: N denotes the total number of nodes; A smaller value of $L(i,j)$ indicates stronger electrical coupling between nodes i and j.

3.1.3 The Comprehensive Electrical Distance between All Node Pairs

Combining the statistical distance $D_E(i,j)$ between nodes i and j in Section 3.1.1 with the electrical distance $L(i,j)$ defined by the voltage sensitivity between nodes i and j in Section 3.1.2, the comprehensive electrical distance $d(i,j)$ between nodes i and j is defined as:

$$d(i,j)=\eta_1 D_E(i,j)+\eta_2 L(i,j) \tag{13}$$

In the formula: η_1 and η_1 are the weight coefficients of $D_E(i,j)$ and $L(i,j)$, respectively. If the values of η_1 and η_2 are not selected appropriately, the connectivity of the partition results will fail to meet the requirements. To avoid this problem, this paper uses the fractional scale method to determine the values of η_1 and η_2. For details of the fractional scale method, please refer to Ref. [26].

3.2 AP Clustering Partitioning Based on Comprehensive Electrical Distance

Most existing partitioning methods employ clustering algorithms to achieve network division. Classical clustering techniques such as the K-means algorithm and fuzzy C-means algorithm [12] require manual pre-specification of cluster centers and the number of partitions. While such approaches are suitable for small-scale and structurally simple power grids, they become less effective in the context of increasingly large and complex distribution networks. The manual specification of parameters often relies heavily on engineering experience, introducing strong subjectivity into the partitioning process. Applying empirical knowledge from original scenarios to new contexts may not guarantee reasonable partitioning results, and the voltage control performance under such partitions is often unsatisfactory.

For the problems existing in classical clustering algorithms, there are already many improved algorithms or new algorithms that do not require manual designation of cluster centers and the number of partitions. Among them, the AP clustering algorithm has better interpretability in data center selection and global data processing. It holds that the number of clusters is naturally determined by the data distribution, and the cluster centers naturally exist in the data to be clustered. By analyzing the information of each data point, the best clustering center points are selected [27]. Therefore, the AP clustering algorithm has been widely applied in voltage partitioning.

The AP clustering algorithm constructs a similarity matrix S to quantify the degree of similarity between each pair of nodes. In this paper, the similarity matrix is formulated using the negative value of the comprehensive electrical distance defined in Eq. (14), as follows:

$$S(i,j)=-d(i,j) \tag{14}$$

In the formula: $S(i,j)$ represents an $N\times N$ order matrix, where N is the number of nodes, used to indicate the similarity between nodes i and j.

$s(j, j)$ is the element on the main diagonal of the similarity matrix $S(i, j)$, which represents the bias parameter $P(j)$, also known as the reference degree P. A higher value of the preference parameter P indicates a greater likelihood that node j will be chosen as a cluster center. Because the AP algorithm considers each node as a potential cluster center, the reference degree P of each point is the same.

The AP clustering algorithm regards each data point as a node. Nodes pass on their respective attraction and fitness information, that is, through iterative calculation, to select the best clustering center. The formulas for the attraction matrix R and the fitness matrix A are as follows:

$$\begin{cases} R(i, j) = S(j, j') - \max\limits_{j' \neq j} \{\{A(i, j') + S(i, j')\} \\ A(i, j) = \min\left\{0, R(j, j) + \sum\limits_{i' \in (i, j)} \max\{0, R(i', j)\}\right\} \end{cases} \tag{15}$$

In the formula: $R(i, j)$ indicates the attraction degree for node j to serve as the intra-class center for node i. This metric is directed from node i to j. $A(i, j)$ represents the suitability of node i in selecting node j as the center point within its class, pointing from j to i.

A higher value of $R(i, j)$ suggests a greater likelihood that node j will become the cluster center for node i; similarly, a higher value of $A(i, j)$ indicates a stronger tendency for node j to be selected as the cluster center for node i. Upon convergence of the iterative process, the cluster center for node i is determined by selecting the node j that maximizes the sum$\{R(i, j) + A(i, j)\}$.

3.3 Dynamic Adjustment of Partitions Based on Dynamic Reactive Power Reserve

After obtaining the initial partitioning results, to ensure the subsequent distributed voltage control effect, the partitioning results should guarantee that there is sufficient reactive power margin in each partition for adjustment. Therefore, it is necessary to test the reactive power margin in each partition of the initial partitioning results. If it does not meet the requirements, partition adjustment needs to be carried out to obtain the final partitioning results.

When testing the reactive power reserve margin of partition j, the dynamic reactive power margin α_j of partition j is adopted, as shown in Eq. (16).

$$\alpha_j = \left(1 - \frac{|Q_{Lj}|}{\sum\limits_{j,k=1}^{Z} Q_{PVj,k}}\right) \times 100\% \tag{16}$$

In the formula: Q_{Lj} denotes the reactive power demand of the load in partition j; $Q_{PVj,k}$ represents the adjustable reactive power from the k-th PV inverter in partition j; Z indicates the total number of PV inverters within partition j.

It is generally believed that when $\alpha_j > 0$, the partition meets the reactive power margin requirement, that is, it meets the dynamic reactive power reserve requirement. If $\alpha_j \leq 0$, the reactive power margin requirement is not met. Then, partition adjustment is carried out. The partitions that do not meet the requirements are merged with the adjacent partitions with sufficient dynamic reactive power reserves. If the merged partitions meet the dynamic reactive power margin requirement, the partition adjustment is completed. The partition adjustment is verified again to see if each partition meets the dynamic reactive power margin requirement. If not, the partition adjustment continues. The final partitioning result is obtained until all partitions meet the

dynamic reactive power margin requirements. The flowchart of partitioning in this paper is shown in Fig. A1 of Appendix A.

3.4 Metrics for Evaluating the Quality of Partitioning Results

For the same network structure and power flow distribution, different partitioning methods usually lead to different partitioning results. Therefore, it is necessary to evaluate the partitioning quality of the partitioning results. This paper adopts the following two indicators to evaluate the quality of the partitioning results.

(1) Modularity Index

Modularity is a general indicator for evaluating the quality of partitions, reflecting the tightness of nodes within a partition. Its calculation formula is as follows:

$$\begin{cases} \rho = \dfrac{1}{2m}\sum\limits_{i}\sum\limits_{j}\left(A_{ij} - \dfrac{k_i k_j}{2m}\right)\delta(i,\ j) \\ A_{ij} = 1 - d(i,\ j)/\max d(i,\ j) \\ k_i = \sum\limits_{j} A_{ij} \\ m = \left(\sum\limits_{i}\sum\limits_{j} A_{ij}\right)/2 \end{cases} \tag{17}$$

In the formula: ρ denotes the modularity index, satisfying $0 \le \rho \le 1$; A_{ij} represents the weight of the edge connecting nodes i and j; k_i represents the sum of the weights of all edges connected to node i; m represents the total sum of all edge weights in the network; If nodes i and j are in the same partition, $\delta(i, j) = 1$; otherwise, $\delta(i, j) = 0$.

In the grid partition, when $0.3 \le \rho \le 0.7$, the partition can be evaluated as a better partition. When $\rho < 0.3$, the connection between nodes within a partition is weak, while the connection between nodes between partitions is strong. The mutual influence of reactive power and voltage control between partitions increases, which is not conducive to the local balance of reactive power and the stable control of voltage. When $\rho > 0.7$, the nodes within the partition are overly closely connected. Voltage adjustment of a certain node will cause voltage fluctuations in other nodes, making voltage adjustment of nodes within the partition difficult. The partition lacks the necessary flexibility and adaptability [28].

(2) Partition Variation Rate Metric

To reflect the stability of the partitioning results under frequent variations in reactive power flow—that is, the sensitivity of the partitioning outcome to changes in reactive power flow—this paper introduces a partition variation rate metric, denoted as K, which is calculated as follows:

$$K = (1 - n_x/n_T) \times 100\% \tag{18}$$

In the formula: where: n_T denotes the number of selected time intervals; n_x represents the frequency of the most frequently occurring partition result across the selected T time intervals. The larger K means the more frequent changes in the partition results under the scenario of frequent power flow changes, and the more unstable the partition results are, which may lead to poor node voltage control effect and even difficulty in ensuring voltage stability. The smaller the K, the less frequently the partitioning result changes in scenarios with frequent power flow variations, and the more stable the partitioning result is, which is conducive to ensuring voltage stability.

4 A Distributed Cooperative Voltage Optimization Method for Active Distribution Networks

In recent years, distributed voltage control has been widely applied in the voltage optimization of distribution networks, and this method is suitable for scenarios with obvious network partition structures [29]. This section mainly discusses the distributed voltage collaborative optimization method based on reactive power partitioning. Firstly, with the goal of minimizing active power loss in the network, a reactive power optimization model is established, and the LinDistFlow equation is used to transform the reactive power optimization model into a convex optimization model. Then, based on the results of voltage/reactive power partitioning, the optimization model is decomposed into several sub-models within each partition. Through the synchronous ADMM algorithm for distributed optimization calculation, the distributed voltage collaborative optimization method for distributed photovoltaic within each partition is obtained.

4.1 The Objective Function and Constraints

4.1.1 Objective Function

To achieve reactive power control and voltage optimization in active distribution networks while ensuring nodal voltages remain within secure operational limits and maintaining economic efficiency in practical system operation, the minimization of total active power loss is adopted as the objective function. The mathematical expression for the system active power loss is provided in Eq. (19).

$$P_{\text{loss}} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} I_{ij}^2 r_{ij} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{P_{ij}^2 + Q_{ij}^2}{V_i^2} r_{ij} \tag{19}$$

In the formula: V_i denotes the voltage magnitude at node i; I_{ij}^2 denotes the current transmitted by line ij; P_{ij} and Q_{ij} denote the active and reactive power flow, respectively, at the sending end of line ij; r_{ij} denotes the resistance of line ij.

4.1.2 Constraints

The equation constraint established in this paper is the system power flow constraint, and the inequality constraints are the reactive power regulation range constraint of the photovoltaic inverter and the safe operation range constraint of the voltage at each node.

(1) The system power flow constraint in this paper is described by the branch power flow equation. From the radiative active distribution network shown in Fig. 2, assuming the influence of the line on ground capacitance and admittance is ignored, according to Kirchhoff's voltage-current law, the power flow equation can be obtained, as shown in Eq. (20):

$$\begin{aligned}
&\sum_{k:j\to k} P_{jk} = P_{ij} - R_{ij}\frac{P_{ij}^2 + Q_{ij}^2}{V_i^2} - p_j^l + p_j^{pv} \\
&\sum_{k:j\to k} Q_{jk} = Q_{ij} - X_{ij}\frac{P_{ij}^2 + Q_{ij}^2}{V_i^2} - q_j^l + q_j^{pv} \\
&V_j^2 = V_i^2 - 2\left(R_{ij}P_{ij} + X_{ij}Q_{ij}\right) + \left(R_{ij}^2 + X_{ij}^2\right)\frac{P_{ij}^2 + Q_{ij}^2}{V_i^2} \\
&V_1 = V_{\text{ref}}
\end{aligned} \tag{20}$$

In the formula: V_1 represents the voltage of the balance node; V_{ref} represents the reference voltage of the node; V_j represents the voltage of node j; $k:j\rightarrow k$ represents the set of lines from the starting node j to the ending node k. Assuming the direction of $j\rightarrow k$ is positive, P_{jk} and Q_{jk} respectively represent the active power and reactive power passing through the starting end of line jk. p_j^{pv} and q_j^{pv} respectively represent the active power and reactive power generated by the distributed photovoltaic at node j; p_j^l and q_j^l respectively represent the active power and reactive power required by the load at node j; X_{ij} represents the impedance on line ij, and R_{ij} represents the resistance on line ij.

(2) The reactive power regulation range constraint of distributed PV power. The reactive power regulation capability of a PV system is constrained by its active power output and the rated capacity of its inverter, with the specific expression shown in Eq. (21).

$$|q_j^{pv}| \leqq q_{j\max}^{pv} = \sqrt{s_j^2 - \left(p_j^{pv}\right)^2} \tag{21}$$

In the formula: s_j represents the inverter capacity of distributed PV; $q_{j\max}^{pv}$ represents the maximum value of the currently available reactive power of the distributed PV at node j; p_j^{pv} represents the current active power of distributed PV at node j.

(3) During the normal operation of an active distribution network, all nodes are required to operate within the safe operation range, and thus the voltage safety constraint must be satisfied.

$$1-\varepsilon \leq V_j \leq 1+\varepsilon \tag{22}$$

In the formula: ε represents the deviation tolerance for nodal voltage, which is set to 0.05 p.u.

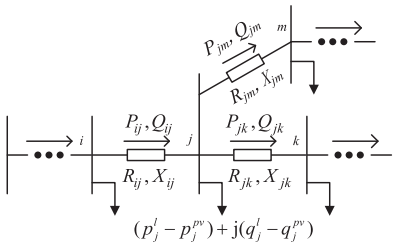

Figure 2: Schematic diagram of branch power flow

4.1.3 Model Convexification Processing

Eq. (23) describes the above optimization model. Among them, the nonlinear and non-convex equality constraints render the established optimization model non-convex. In the solution process of distributed optimization algorithms, non-convex optimization models face two key issues: difficulty in finding the global optimal solution, and inability to guarantee convergence during the solution of non-convex models. To address the above problems, existing methods typically involve the convexification of the optimization model. Currently, the main methods for convexifying optimization models involve convexifying the model through second-order cone programming (SOCP), semidefinite programming (SDP), and other approaches [30].

$$\begin{cases} \min \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \dfrac{P_{ij}^2 + Q_{ij}^2}{V_i^2} r_{ij} \\ s.t. \begin{cases} \sum\limits_{k:j\to k} P_{jk} = P_{ij} - R_{ij} \dfrac{P_{ij}^2 + Q_{ij}^2}{V_i^2} - p_j^l + p_j^{pv} \\ \sum\limits_{k:j\to k} Q_{jk} = Q_{ij} - X_{ij} \dfrac{P_{ij}^2 + Q_{ij}^2}{V_i^2} - q_j^l + q_j^{pv} \\ V_j^2 = V_i^2 - 2\left(R_{ij}P_{ij} + X_{ij}Q_{ij}\right) + \left(R_{ij}^2 + X_{ij}^2\right) \dfrac{P_{ij}^2 + Q_{ij}^2}{V_i^2} \\ V_1 = V_{\text{ref}} \\ |q_j^{pv}| \leqq q_{j\max}^{pv} = \sqrt{s_j^2 - \left(p_j^{pv}\right)^2} \\ 1 - \varepsilon \leq V_j \leq 1 + \varepsilon \end{cases} \end{cases} \tag{23}$$

The voltage difference between adjacent nodes can be ignored compared with their voltage amplitudes. Similarly, the line power loss can be ignored compared with the transmission power. Based on the above reasonable assumptions, the LinDistFlow equation can simplify the quadratic terms in the equality constraints, convexify the model, and significantly reduce the required computational load. Therefore, this paper adopts the LinDistFlow equation to reasonably simplify Eq. (23); meanwhile, it sets the node voltage $V_i = V_1$ in the objective function. Linearization is achieved by using variable substitution for the square of the voltage. Finally, the model is simplified to the form shown in Eq. (24).

$$\begin{cases} \min \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{P_{ij}^2 + Q_{ij}^2}{U_1} r_{ij} \\ s.t. \begin{cases} \sum\limits_{k:j\to k} P_{jk} = P_{ij} - p_j^l + p_j^{pv} \\ \sum\limits_{k:j\to k} Q_{jk} = Q_{ij} - q_j^l + q_j^{pv} \\ U_j = U_i - 2\left(R_{ij}P_{ij} + X_{ij}Q_{ij}\right) \\ |q_j^{pv}| \leqq q_{j\max}^{pv} = \sqrt{s_j^2 - \left(p_j^{pv}\right)^2} \\ 1 - \varepsilon \leq V_j \leq 1 + \varepsilon \\ U_1 = {V_{ref}}^2 \end{cases} \end{cases} \tag{24}$$

In the formula: U_i is represented as the square of the voltage at node i, and U_j is represented as the square of the voltage at node j. Therefore, Eq. (23) is simplified into a linear convex optimization model. According to optimization theory, the local optimal solution of a convex optimization problem is equal to the global optimal solution. Therefore, by adopting the LinDistFlow equation, the model is guaranteed to possess a globally optimal solution, which facilitates subsequent distributed cooperative voltage optimization.

4.2 Distributed Cooperative Voltage Optimization Based on Reactive Power Partitioning

Due to the specific network structure of the active distribution network, after reactive power partitioning, adjacent partitions are connected only through a single branch. Herein, this network can be characterized by a directed graph $G = (N, W)$, in the formula: G denotes the directed graph; W denotes the set of all branches, and $w_{ij} \in W$ denotes a specific branch in the active distribution network system; N

denotes the set of all nodes in the distribution network system, where $\{i, j\} \in N$ represents nodes i and j in the distribution network system. Therefore, the optimization model (24) can be decomposed into multiple partition-based optimization models according to the results of physical partitioning. Assuming that the current distribution network is divided into p connected reactive power partitions $p1$, $p2$, $p3\ldots$, then N_{p1} denotes the node set of partition $p1$, W_{p1} denotes the branch set of partition $p1$, $w_{ij} \in W$ represents a branch in partition $p1$; there is one coupling branch between two adjacent partitions $p1$ and $p2$, through which the two partitions exchange variables. The state variables of the coupling branch w_{ij} include the squared voltages U_i and U_j at the head-end node i and tail-end node j of the coupling branch, as well as the active power P_{ij} and reactive power Q_{ij}. As shown in Eq. (25), the set O corresponds to the coupling branches between different partitions, while the state variable of a coupling branch originating from partition $p1$ is denoted as $X_{p1,ij}$.

$$\begin{aligned} O &= \left\{ w_{ij} | W_{p1} \cap W_{p2},\ \forall p1,\ p2 \in R,\ p1 \neq p2 \right\} \\ X_{p1,ij} &= \left\{ P_{p1,ij},\ Q_{p1,ij},\ U_{p1,i},\ U_{p1,j} \right\} \end{aligned} \tag{25}$$

In this paper, Fig. 3 is used as an example. It can be seen from the figure that the coupling branch between Partition 1 and Partition 2 is w_{45}, and the coupling branch between Partition 2 and Partition 3 is w_{78}. Therefore, the set of coupling branches in the figure is $O = \{w_{45}, w_{78}\}$.

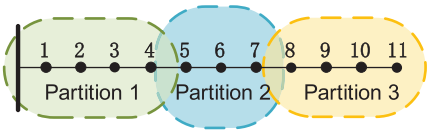

Figure 3: Schematic diagram of regional division

During the normal operation of the system, the state variables of the coupling branch between two adjacent partitions are equal. Therefore, to ensure the equivalence between the decomposed sub-problems and the original problem, the state variable $X_{p1,ij}$ of the coupling branch in Partition $p1$ must be equal to the state variable $X_{p2,ij}$ of the coupling branch in the adjacent Partition $p2$. Thus, when the above conditions are satisfied, the optimization model can be decomposed into sub-optimization models for each partition. The distributed reactive power optimization model based on reactive power partition coordination proposed in this paper is shown in Eq. (26).

$$\begin{aligned} & min \sum_{p1=1}^{p} f_{p1}\left(x_{p1}\right) \\ & \text{s.t.} \begin{cases} h_{p1}\left(x_{p1}\right) = 0 \\ g\left(x_{p1}\right) = 0 \\ X_{p1,ij} = X_{p2,ij}, w_{ij} \in O \end{cases} \end{aligned} \tag{26}$$

In the formula: $f_{p1}\left(x_{p1}\right)$ denotes the objective function in Partition $p1$; $h_{p1}\left(x_{p1}\right)$ and $g\left(x_{p1}\right)$ respectively denote the equality constraints and inequality constraints in Partition $p1$; x_{p1} denotes the set of all variables in Partition $p1$; and Partition $p2$ denotes the partition adjacent to Partition $p1$. On this basis, decomposing the reactive power optimization model of the entire system into p sub-optimization models is conducive to distributed collaborative computing.

4.3 A Synchronous ADMM Approach for Distributed Computation

The linear and decomposable optimization model obtained in Section 3.2 can be efficiently solved in a distributed manner using the synchronous ADMM. The original optimization problem is decomposed into two coordinated sub-problems for Partitions $p1$ and $p2$ below, followed by a brief description of the distributed solution procedure using synchronous ADMM [31].

Firstly, the controller of each partition establishes the optimization model for its respective partition. The augmented Lagrangian functions corresponding to the objective functions of the optimization models for Partitions $p1$ and $p2$ are $L_{p1}\left(x_{p1},\ X_{p1}^{t},\ \lambda_{p1}^{t}\right)$ and $L_{p2}\left(x_{p2},\ X_{p2}^{t},\ \lambda_{p2}^{t}\right)$, respectively, and through appropriate simplification, they are transformed into:

$$\begin{aligned} L_{p1}\left(x_{p1},\ X_{p1}^{t},\ \lambda_{p1}^{t}\right) &= f_{p1}\left(x_{p1}\right) + \frac{\rho}{2} \parallel X_{p1,ij} - X_{p1}^{t} + \lambda_{p1}^{t} \parallel_2^2 \\ L_{p2}\left(x_{p2},\ X_{p2}^{t},\ \lambda_{p2}^{t}\right) &= f_{p2}\left(x_{p2}\right) + \frac{\rho}{2} \parallel X_{p2,ij} - X_{p2}^{t} + \lambda_{p2}^{t} \parallel_2^2 \end{aligned} \tag{27}$$

In the formula: t denotes the iteration number; λ_{p1}^{t} and λ_{p2}^{t} respectively denote the vectors composed of dual variables of Partitions $p1$ and $p2$, which correspond to the state variables of the coupling branch; ρ denotes the penalty parameter of the ADMM algorithm; X_{p1}^{t} and X_{p2}^{t} respectively denote the fixed reference values of Partitions $p1$ and $p2$ in the $t+1$-th iteration, which are taken as the average values of the coupling branch states obtained from the t-th iteration of Partitions $p1$ and $p2$, as shown in Eq. (28).

$$X_{p1}^{t} = X_{p2}^{t} = \left(X_{p1,ij}^{t} + X_{p2,ij}^{t}\right)/2 \tag{28}$$

After establishing the two partition sub-models, the controllers within each partition are solved according to the following steps, and the specific process is shown in Fig. A2 in Appendix A.

During the $(t+1)$-th iterative process, the controllers belonging to Partitions $p1$ and $p2$ respectively solve their own partition's optimization models, and perform parallel computing to obtain the decision variable values for their respective partitions that minimize the augmented Lagrangian functions $L_{p1}\left(x_{p1},\ X_{p1}^{t},\ \lambda_{p1}^{t}\right)$ and $L_{p2}\left(x_{p2},\ X_{p2}^{t},\ \lambda_{p2}^{t}\right)$, as shown in Eq. (29). Meanwhile, they obtain the coupling branch state variables X_{p1}^{t+1} and X_{p2}^{t+1} for their respective partitions.

$$\begin{aligned} x_{p1}^{t+1} &= argminL_{p1}\left(x_{p1},\ X_{p1}^{t},\ \lambda_{p1}^{t}\right) \\ x_{p2}^{t+1} &= argminL_{p2}\left(x_{p2},\ X_{p2}^{t},\ \lambda_{p2}^{t}\right) \end{aligned} \tag{29}$$

The controllers of Partitions $p1$ and $p2$ exchange the coupling branch state variables X_{p1}^{t+1} and X_{p2}^{t+1}, and the mean value of the coupled branches' state variables is solved based on Eq. (30), with this value adopted as the state variable reference for subsequent iterations.

$$X_{p1}^{t+1} = X_{p2}^{t+1} = \left(X_{p1,ij}^{t+1} + X_{p2,ij}^{t+1}\right)/2 \tag{30}$$

The controllers of Partitions $p1$ and $p2$ respectively update the dual variables within their partitions, as shown in Eq. (31).

$$\begin{aligned} \lambda_{p1}^{t+1} &= \lambda_{p1}^{t} + \left(X_{p1,ij}^{t+1} - X_{p1}^{t+1}\right) \\ \lambda_{p2}^{t+1} &= \lambda_{p2}^{t} + \left(X_{p2,ij}^{t+1} - X_{p2}^{t+1}\right) \end{aligned} \tag{31}$$

The convergence criterion of this algorithm is that the dual residual is less than the convergence accuracy. The dual residual refers to the square of the 2-norm of the difference between the coupling branch states calculated by adjacent partitions, as shown in Eq. (32).

$$\left\| X_{p1,ij}^{t+1} - X_{p2,ij}^{t+1} \right\|_2^2 \leq \delta \tag{32}$$

It can be clearly seen from Fig. A2 in Appendix A that throughout the ADMM calculation process, the only variable that must be exchanged between the two partitions is the state variable associated with the coupled branch. This is conducive to iterative computation and reduces the hardware requirements for data transmission in each partition. During each iteration, all partitions can perform parallel computing, which effectively improves the computational efficiency.

5 Case Studies

A Matlab simulation platform is used to build the IEEE 33-bus test system, and simulation analysis is conducted to verify the effectiveness of the partitioning method and distributed voltage control method proposed in this paper.

5.1 Case Setting

This paper mainly focuses on the partitioning and distributed voltage control methods for distribution networks with high penetration of distributed PV. Therefore, under the premise that the static reactive power demand in the distribution network is satisfied, this paper assumes that the only dynamic reactive power compensation device in the distribution network is the PV inverter.

There are two scenarios regarding the relationship between the dynamically adjustable reactive power capacity of PV inverters and the dynamic reactive power demand within each partition:

Scenario 1: The dynamically adjustable reactive power capacity of PV inverters in all partitions can meet the dynamic reactive power demand of their respective partitions.

Scenario 2: The dynamically adjustable reactive power capacity of PV inverters in some partitions can meet the dynamic reactive power demand of their respective partitions.

Thus, this paper constructs simulation scenarios corresponding to the above two scenarios by adjusting the number of PV inverters connected to the IEEE 33-bus system:

Simulation Scenario 1: Distributed PV systems with a capacity of 0.3 MVA each are connected to 10 nodes (Nodes 3, 6, 8, 11, 14, 16, 20, 22, 25, and 29) of the IEEE 33-bus system, corresponding to Scenario 1.

Simulation Scenario 2: Distributed PV systems with a capacity of 0.3 MVA each are connected to 9 nodes (Nodes 3, 6, 11, 14, 16, 20, 22, 25, and 29) of the IEEE 33-bus system, corresponding to Scenario 2.

Since the partitioning method proposed in this paper merges partitions with insufficient reactive power margin with adjacent partitions to ensure that the reactive power margin within each partition meets the requirements, this paper selects a typical time in Scenario 1 as the scenario for distributed voltage coordinated optimization.

Simulation Parameters

Due to the strong similarity in photovoltaic output among those geographically close, the distributed photovoltaic in the scene is divided into three groups based on geographical location, namely Group 1: Nodes 3, 6, 20, and 22; Group 2: Nodes 8, 11, 14, and 16; Group 3: Nodes 25 and 29.

Each group of photovoltaic applications uses the same output data. The three groups of photovoltaic output data are derived from the measured data in a certain area. The three groups of photovoltaic data are obtained by randomly selecting 50 groups of measured data, respectively. In this paper, the power factor cosφ is taken as 0.95. Under the constant power factor control mode, based on three sets of photovoltaic output data and power factor, three sets of photovoltaic reactive power output data $Q = P\tan\varphi$ can be obtained.

The node load data are obtained by multiplying the IEEE 33-node system load by the daily load variation coefficient. The load variation coefficient of each node is the same, as shown in Fig. A3 of Appendix B.

Since the photovoltaic output of each group of nodes is the same, one node is selected for each group, and the photovoltaic output curve is given. Group 1 selects node 3, Group 2 selects node 16, and Group 3 selects node 25. The photovoltaic output results at nodes 3, 16, and 25 are shown in Fig. A3.

As can be seen from Fig. A3, the PV output varies greatly at different times. During the period from 0:00 to 8:00, the load is relatively light and the active power output of the photovoltaic is basically zero, resulting in sufficient reactive power margin for the photovoltaic inverter. During the 8–10 period, the system load is relatively light and the active power output of the photovoltaic is small, so the reactive power margin of the photovoltaic inverter is relatively sufficient. During the period from 10 to 13, the system load was relatively heavy and the active power output of the photovoltaic system was large, resulting in a relatively scarce reactive power margin for the photovoltaic inverter. The period from 13:00 to 24:00 is roughly the same as the above-mentioned period. Therefore, this paper selects three typical moments, 4:00, 9:00, and 13:00, for analysis.

5.2 Analysis of Reactive Power Partitioning Results

5.2.1 GMM Results

Based on the reactive power output data of three groups of PV systems, three sets of GMMs for reactive power output can be obtained, as shown in Figs. A4–A6 of Appendix B, respectively.

It can be observed from the three figures that the GMMs of the three groups of reactive power output are all composed of four distinct Gaussian components in terms of probabilistic characteristics.

5.2.2 MCS Results

Based on the calculation times of the IEEE 33-node system MCS in Ref. [32], the calculation times of the MCS power flow in this paper are set to 10,000 times. To further verify the sufficiency of choosing MCS calculation as 10,000 times, the statistical distance result matrix of 20,000 MCS calculations is taken as the benchmark matrix, and the statistical distance result matrix of MCS obtained from 1000 to 20,000 times with a resolution of 1000 times is used as the comparison matrix. Calculate the Average Relative Difference Rate (ARD), and the formula is:

$$\mathrm{ARD} = \frac{1}{M}\sum_{k=1}^{M}\left|\frac{D_{\mathrm{base},k} - D_{\mathrm{compare},k}}{D_{\mathrm{base},k} + \varepsilon}\right| \times 100\% \tag{33}$$

In the formula: $D_{\mathrm{base},k}$ denotes the k-th element in the base distance matrix; $D_{\mathrm{compare},k}$ denotes the k-th element in the comparison distance matrix; ε represents an extremely small value, which is used to avoid division by zero; M is the number of off-diagonal elements in the matrix.

The convergence curve of the average relative difference rate of statistical distance results obtained under 1000 to 20,000 MCS calculations based on the above content is shown in Fig. A7 of Appendix B. It can be seen from Fig. A7 that the average relative difference rate between the statistical distance under

10,000 MCS calculations and that under 20,000 MCS calculations is less than 0.05%. Additionally, the average relative difference rate of statistical distance results obtained from 10,000 to 20,000 MCS calculations changes minimally. Therefore, it is reasonable to select 10,000 as the number of MCS calculations.

5.2.3 Results of Electrical Distance

The electrical distance defined by the "reactive power flow distance method" is denoted as d'. The weights η_1 and η_2 in the comprehensive electrical distance proposed in this paper are determined using the fractional scaling method, yielding results of $\eta_1 = 0.4$ and $\eta_2 = 0.6$. The effects of different weights on the partitioning results are shown in Table A1 of Appendix B.

The sensitivity electrical distance obtained from Eq. (11), the statistical distance between nodes obtained from the node voltage calculation results and Eq. (6), the comprehensive electrical distance obtained from Eq. (12), and the reactive power flow distance obtained from the method in Ref. [14] are shown in Fig. 4, respectively.

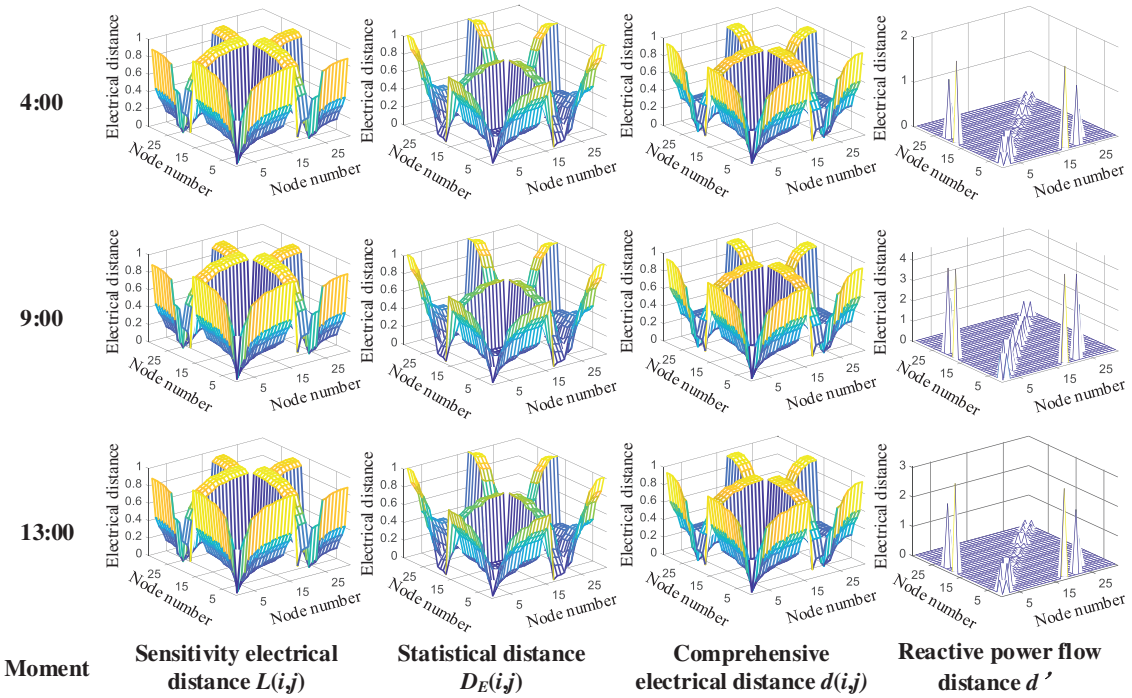

Figure 4: The three typical moments of $L(i, j)$, $D_E(i, j)$, $d(i, j)$ and d'

As shown in Fig. 4, the sensitivity distance and statistical distance do not change with time and remain essentially the same at three typical time points. This is because the sensitivity distance is mainly determined by the grid structure—if the grid structure remains unchanged, the sensitivity distance will not change. By modeling the characteristics of PV output, a node voltage dataset under these output characteristics is obtained. Then, the statistical distance is used to quantify the similarity between node voltages under these characteristics, thereby deriving the statistical electrical distance. Since this distance is obtained based on PV output characteristics, it remains unchanged when the PV output characteristics are stable. Therefore, the partitioning results derived from this distance are stable. Consequently, the comprehensive electrical distance composed of the sensitivity distance and statistical distance also does not change with time. It is identical at the three typical time points and exhibits stability.

Due to the frequent fluctuations in power flow in scenarios where photovoltaic power is highly integrated into the distribution network, the reactive power flow distance based on reactive power flow changes with changes in reactive power flow. The reactive power flow distance varies between nodes at the three typical time points, so it does not exhibit stability.

5.2.4 Partitioning Results

In this paper, four validity constraint indicators commonly used in power grid partitioning are adopted to verify the rationality of the partitioning results.

(1) Node connectivity metric within a partition (hereinafter referred to as "Indicator 1"): Nodes within the same partition should be directly or indirectly connected to each other, but should not be indirectly connected through nodes in other partitions.

(2) Reasonableness of the number of partitions metric (hereinafter referred to as "Indicator 2"): If the requirement that the optimal number of partitions for an n-node system does not exceed $\sqrt{n}$ is satisfied, the number of partitions is considered reasonable.

(3) Reactive power source constraint metric within a partition (hereinafter referred to as "Indicator 3"): Each partition should contain at least one reactive power source. That is, the following condition should be satisfied:

$$N_{i,\min} \le N_i \le N_{i,\max} \tag{34}$$

In the formula: N_i represents the number of reactive power sources in partition i; $N_{i,\min}$ and $N_{i,\max}$ respectively represent the lower limit and upper limit of the number of reactive power sources in partition i. When $N_{i,\min} > 1$, that is, there is at least one reactive power source in the partition, the requirement is satisfied.

(4) Reactive power margin constraint within a zone (hereinafter referred to as "Indicator 4"): Each zone must have sufficient reactive power margin to regulate node voltage fluctuations, i.e., the dynamic reactive power margin within each zone must satisfy the constraints specified in Section 3.3.

The zone division results must satisfy all four of the above constraints; otherwise, the zone division is invalid.

(a) Result of Scenario 1

The partitioning results under Scenario 1 are consistent across three typical time points, as shown in Fig. 5. The validity of the partitioning results is verified as follows:

(1) Indicator 1: As shown in the partitioning results of Fig. 5, the partitioning results in this paper meet the connectivity requirement.

(2) Indicator 2: As shown in the partitioning results of Fig. 5, the number of partitions is 5, which is less than $\sqrt{33}$, satisfying the validity check for the number of partitions.

(3) Indicator 3: Each partition should contain at least one reactive power source. As shown in Fig. 5, all partitions contain at least one distributed photovoltaic power source, meeting the requirement.

(4) Indicator 4: The dynamic reactive power reserve margin tables for each partition in this paper are shown in Table 1. The dynamic reactive power reserve margins in all partitions meet the constraints specified in Section 3.3, ensuring a certain reactive power margin.

(b) Result of Scenario 2

The initial partitioning results and final partitioning results under Scenario 2 are shown in Figs. 6 and 7, respectively. The validity of the partitioning results is verified as follows:

(1) Indicator 1: As shown in the final partitioning results in Fig. 7, the partitioning results in this paper meet the connectivity requirements.
(2) Indicator 2: As shown in the final partitioning results in Fig. 7, the number of partitions is 4, which is less than $\sqrt{33}$, satisfying the validity check for the number of partitions.
(3) Indicator 3: Each partition should contain at least one reactive power source. As shown in the final partitioning results in Fig. 7, each partition contains at least one distributed photovoltaic power source, satisfying the requirement.
(4) Indicator 4: The dynamic reactive power reserve margin table for each partition in this paper is shown in Table 2, which shows the results after partition adjustment. The dynamic reactive power reserve margin of all partitions meets the constraints specified in Section 3.3 and has a certain reactive power margin

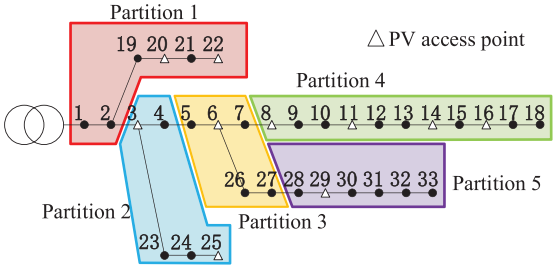

Figure 5: Topology diagram of partition results under Scenario 1 in this paper

Table 1: The dynamic reactive power margin at three typical moments in scenario 1 for the proposed method in this paper

Scenario	Partition	4:00	9:00	13:00
Scenario 1	Partition 1: 1, 2, 19–22	0.4286	0.2421	0.0498
	Partition 2: 3–4, 23–25	0.3991	0.1926	0.0522
	Partition 3: 5–7, 26, 27	0.1435	0.0812	0.0373
	Partition 4: 8–18	0.1041	0.0763	0.0401
	Partition 5: 28–33	0.0896	0.0668	0.0346

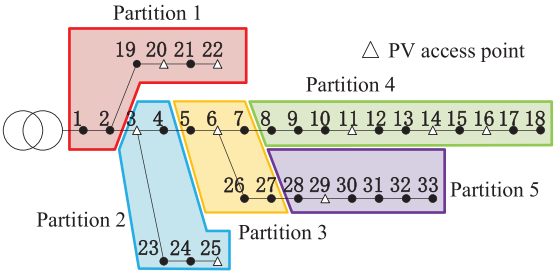

Figure 6: Topological diagram of the initial partition result under Scenario 2 in this paper

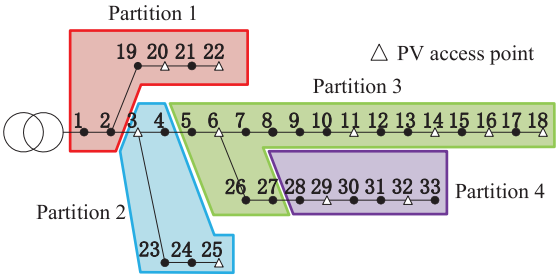

Figure 7: Topological diagram of the final partition result under Scenario 2 in this paper

Table 2: The dynamic reactive power margin at three typical moments in scenario 2 for the proposed method in this paper

Scenario	Partition	4:00	9:00	13:00
Scenario 2 (Initial Partitioning Result, i.e., before adjustment)	Partition 1: 1, 2, 19–22	0.4286	0.2421	0.0498
	Partition 2: 3–4, 23–25	0.3991	0.1926	0.0522
	Partition 3: 5–7, 26, 27	0.1435	0.0812	0.0373
	Partition 4: 8–18	0.0722	0.0122	−0.0137
	Partition 5: 28–33	0.0896	0.0668	0.0346
Scenario 2 (Final Partitioning Result, i.e., after adjustment)	Partition 1: 1, 2, 19–22	0.4286	0.2421	0.0498
	Partition 2: 3, 4, 23–25	0.3991	0.1926	0.0522
	Partition 3: 5–18, 26, 27	0.1206	0.0633	0.0106
	Partition 4: 28–33	0.0896	0.0668	0.0346

The initial partition results under Scenario 2 are shown in Fig. 6, and the reactive power margin of the partitions under Scenario 2 is shown in Table 2. From Table 2, the dynamic reactive power margin value α_j of partition 4 in the initial partition results is $\alpha_j = -0.0137 < 0$. Therefore, to meet the reactive power margin requirements within the zone, partition adjustments are needed to merge partition 3 and 4, reducing the number of partitions to 4, as shown in Fig. 7. This is achieved by merging with adjacent partitions to ensure all partitions meet the dynamic reactive power margin requirements.

5.2.5 Comparative Analysis of the Proposed Partitioning Method and Other Methods

This paper employs the reactive power flow distance method and the proposed method under Scenarios 1 and 2, respectively, to obtain the partitioning results for both scenarios. The comparison results of the modularity index and partitioning change rate index between the two methods are shown in Table 3.

Table 3: Comparison of average modularity and partition change rate metrics under two scenarios

Scenario	Method	Average modularity	Partition variation rate %
Scenario 1	Proposed method	0.6691	0
	Reactive power flow distance method	0.6572	54.17

(Continued)

Table 3 (continued)

Scenario	Method	Average modularity	Partition variation rate %
Scenario 2	Proposed method	0.6599	12.50
	Reactive power flow distance method	0.6361	83.33

As shown in Table 3, the average modularity of the partitioning results obtained using the proposed partitioning method in Scenarios 1 and 2 are 0.6691 and 0.6599, respectively, which fall within the optimal range of 0.3 to 0.7 for the modularity metric and are greater than the values obtained using the other two methods. Furthermore, the partitioning variation rates of the proposed method in Scenarios 1 and 2 are 0% and 12.50%, respectively, while those of the reactive power flow distance method are 54.17% and 83.33%, respectively, which are larger than those of the proposed method. This aligns with the conclusions from the typical time electrical distance changes in Fig. 4. Therefore, the proposed partitioning method demonstrates stability in partitioning results under scenarios with high photovoltaic penetration in distribution grids.

5.3 Results and Analysis of the Distributed Voltage Coordinated Optimization Method

5.3.1 Voltage Results and Analysis of the Distributed Voltage Coordinated Optimization Method

Since the partitioning method mentioned in this paper merges partitions with insufficient reactive power margin with adjacent partitions to ensure that the reactive power margin within each partition meets the requirements, this paper selects the 13 typical time points in Scenario 1 (where the adjustable reactive power margin is minimal) as the optimization scenario for reactive power optimization. Based on the partitioning results, the optimization method proposed in this paper is applied to optimize the IEEE 33-node system. The synchronous ADMM algorithm is set as follows: the penalty parameter ρ is 0.01, and the convergence accuracy δ is 10^{-5}. An analysis of voltage optimization for the distribution network under Scenario 1 is conducted, yielding the voltage values of the IEEE 33-node system before and after optimization, as shown in Fig. 8.

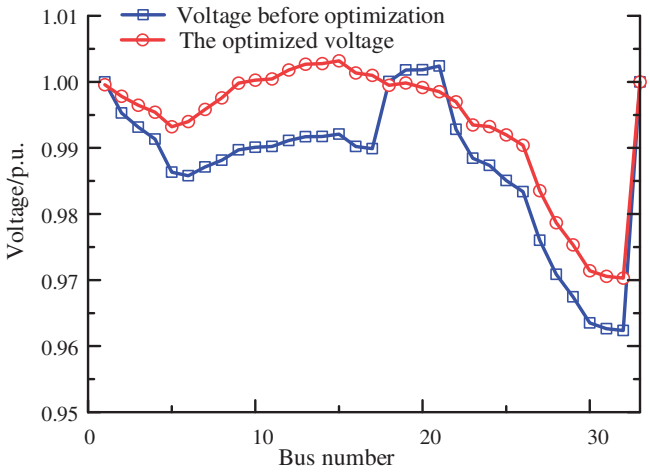

Figure 8: Voltage optimization diagram of IEEE 33 nodes

As shown in Fig. 8, the minimum voltage before optimization was approximately 0.96 p.u., posing a risk of exceeding the lower limit for safe voltage operation. After optimization, the minimum voltage was approximately 0.97 p.u., and the voltage distribution across the entire system was more uniform compared to before optimization, with the voltage at all nodes remaining within the safe operating range. Additionally, the total network loss of the original system was 3.9093 MW, which was reduced to 2.5041 MW after applying the optimization method. The more network loss consumed by the distribution grid, the greater the economic loss incurred by the distribution grid. The optimization method proposed in this paper can effectively reduce network loss and improve the economic efficiency of distribution grid operation.

5.3.2 Validation of the Effectiveness of the Proposed Distributed Optimization Method

To validate the correctness and effectiveness of the proposed partition-based distributed reactive power optimization method, the results of this method are compared with those of centralized optimization strategies. The centralized optimization methods are categorized into two scenarios: the centralized a optimization method solves the reactive power optimization problem Eq. (23), while the centralized b optimization method solves the reactive power optimization problem Eq. (24). The centralized optimization method is a traditional centralized reactive power optimization problem, which is converted into a convex optimization problem using a second-order cone relaxation method, with network constraints based on unsimplified branch flow equations; the centralized b optimization method equates the optimization problem to a distributed reactive power optimization problem, with network constraints based on simplified branch flow equations.

The reactive power output results and total results calculated by the distributed method and centralized methods are shown in Tables 4 and 5. As shown in Table 4, the reactive power output results of centralized b and distributed optimization are similar, while there are minor deviations in the power output results of centralized a optimization and distributed optimization. The primary reason for the deviation is the simplification of the branch flow equations in Section 3.1.3, which omits the quadratic terms in the flow equations. As shown in Table 5, the distributed method results in a higher minimum voltage and a lower maximum voltage compared to the centralized method in the IEEE 33-node system. The distributed method effectively reduces the voltage difference between nodes in the system.

Table 4: Reactive power output data for centralized and distributed photovoltaic systems

PV access node	Centralized a optimization/kvar	Centralized b optimization/kvar	Distributed optimization/kvar
3	300.0000	300.0000	299.9998
6	300.0000	300.0000	300.0000
8	300.0000	300.0000	300.0000
11	192.1733	192.0896	188.1415
14	118.2607	118.1259	118.1259
16	84.6236	84.4193	84.4194
20	69.5227	84.0264	84.2826
22	54.7217	54.6459	54.6475
25	300.0000	299.6824	299.8341
29	300.0000	300.0000	299.9997

Table 5: Comparison between centralized optimization and distributed optimization

	Centralized a optimization	Centralized b optimization	Distributed optimization
Active power loss of line/Mw	2.51258	2.4964	2.5041
Total reactive power/kvar	2019.302	2032.9895	2029.4505
Maximum voltage/p.u.	1.0045	1.0032	1.0030
Minimum voltage/p.u.	0.9697	0.9702	0.9704

5.3.3 ADMM Convergence Analysis

This section analyzes the convergence of a distributed optimization model using synchronous ADMM. This algorithm simplifies the serial ADMM algorithm by averaging the reference values of the coupled branch state variables in the interaction, enabling parallel computation and reducing computational load. Compared to serial ADMM, it offers higher computational efficiency and is more suitable for practical engineering applications. In distributed iterative computation, the number of iterations required to achieve convergence often determines the total communication volume of the entire distributed control system. That is, the fewer the iterations, the smaller the communication burden.

By applying the optimization method proposed in this paper to the IEEE 33-node system, we obtained the changes in dual residuals across regions with respect to the number of iterations Fig. 9 and the changes in dual residuals across regions with respect to the number of iterations Fig. 10. As shown in Fig. 9, the IEEE 33-node system converged after 54 iterations. This indicates that the proposed distributed voltage coordination optimization method can achieve convergence quickly. Additionally, since the information exchange volume in distributed voltage control is limited to coupled branch state variables, it better protects information privacy and security compared to centralized control.

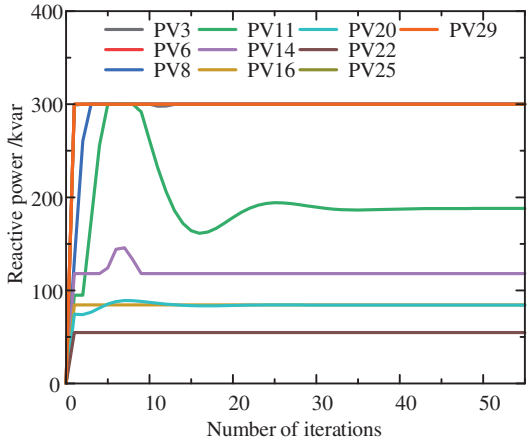

Figure 9: The dual residuals of each region vary with the number of iterations

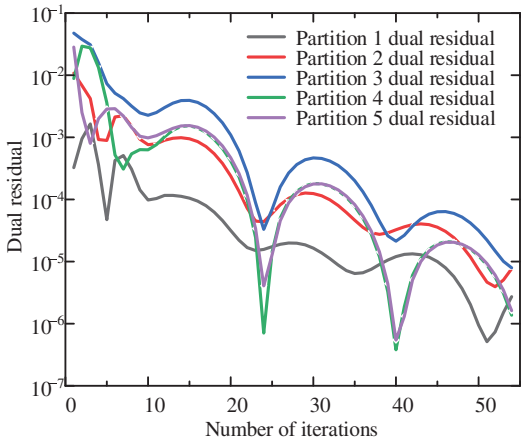

Figure 10: The PV reactive power output of the IEEE 33-node system varies with the number of iterations

6 Conclusion

To address the issues in distribution network scenarios with a high proportion of PV penetration—where frequent changes in power flow lead to frequent variations in voltage zoning results, thereby degrading the performance of traditional centralized voltage control and, in severe cases, even failing to ensure the effectiveness of voltage control and voltage stability—this paper proposes a stable zoning method and a distributed voltage collaborative optimization method that account for the uncertainty of distributed PV output. Firstly, a reactive power zoning approach for distribution networks based on comprehensive electrical distance is proposed. To tackle the problem of zoning changes caused by power flow fluctuations resulting from the uncertainty of distributed PV output, a comprehensive electrical distance combining statistical distance and electrical distance is adopted for zoning. Using the AP clustering algorithm and considering the dynamic reactive power margin of zones simultaneously, stable zoning results are obtained. Secondly, this paper further conducts distributed voltage collaborative optimization based on the established zones. The distributed collaborative optimization based on zoning divides the optimization problem into several sub-problems according to the zones. The use of synchronous ADMM for solving these sub-problems can effectively reduce the number of transmissions of boundary coupling variables, accelerate the convergence of distributed solution, and reduce communication overhead.

The main conclusions are as follows:

(1) The method combining GMM and Monte Carlo probabilistic power flow is used to fully characterize the uncertainty of distributed PV output.

(2) By constructing a comprehensive electrical distance that integrates electrical distance and statistical distance, and applying AP clustering, stable zoning can be effectively achieved.

(3) The zone-based distributed voltage optimization method can effectively reduce the active power loss of the system network and improve voltage distribution. Through distributed voltage optimization calculations, the system communication pressure can be reduced.

(4) The use of the linDistFlow equation enables convex optimization of the zone-based distributed voltage optimization model, while effectively reducing the computational load and improving computational efficiency.

The research methods proposed in this paper provide a new idea for voltage management in distribution networks and have significant application potential. However, the research on distributed PV output in this paper only relies on historical output data, and this study only considers distributed PV as a single reactive power regulation resource. In future research, distributed PV output data can be further obtained through prediction methods to improve the zoning method that accounts for PV uncertainty. In addition, reactive power control can be combined with energy storage scheduling, with research carried out around multi-time-scale control.

Acknowledgement: I would like to express my sincere gratitude to all the other authors for their outstanding contributions and strong support to this research.

Funding Statement: This study was funded by the Science and Technology Project of the Headquarters of State Grid Corporation of China (Project No. 5100-202306384A-2-3-XG).

Author Contributions: The authors confirm contribution to the paper as follows: study conception, task division, content planning, model development, simulation validation, and draft manuscript preparation: Jian Wang, Gongqiang Yang; project introduction, progress and direction oversight, funding acquisition, and final manuscript review: Gangui Yan, Yufeng Sun; data support, practical validation, and final manuscript review: Jie Long. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Due to the nature of this research, participants of this study did not agree for their data to be shared publicly, so supporting data is not available.

Ethics Approval: Not applicable.

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

Abbreviations

EMD	Earth Mover's Distance
GMM	Gaussian mixture mode
MCS	Monte Carlo Simulation
PV	Photovoltaic
ADMM	Alternating Direction Multiplier Method
AP	Affinity propagation
SOCP	Second-order cone programming
SDP	Semidefinite programming

Appendix A

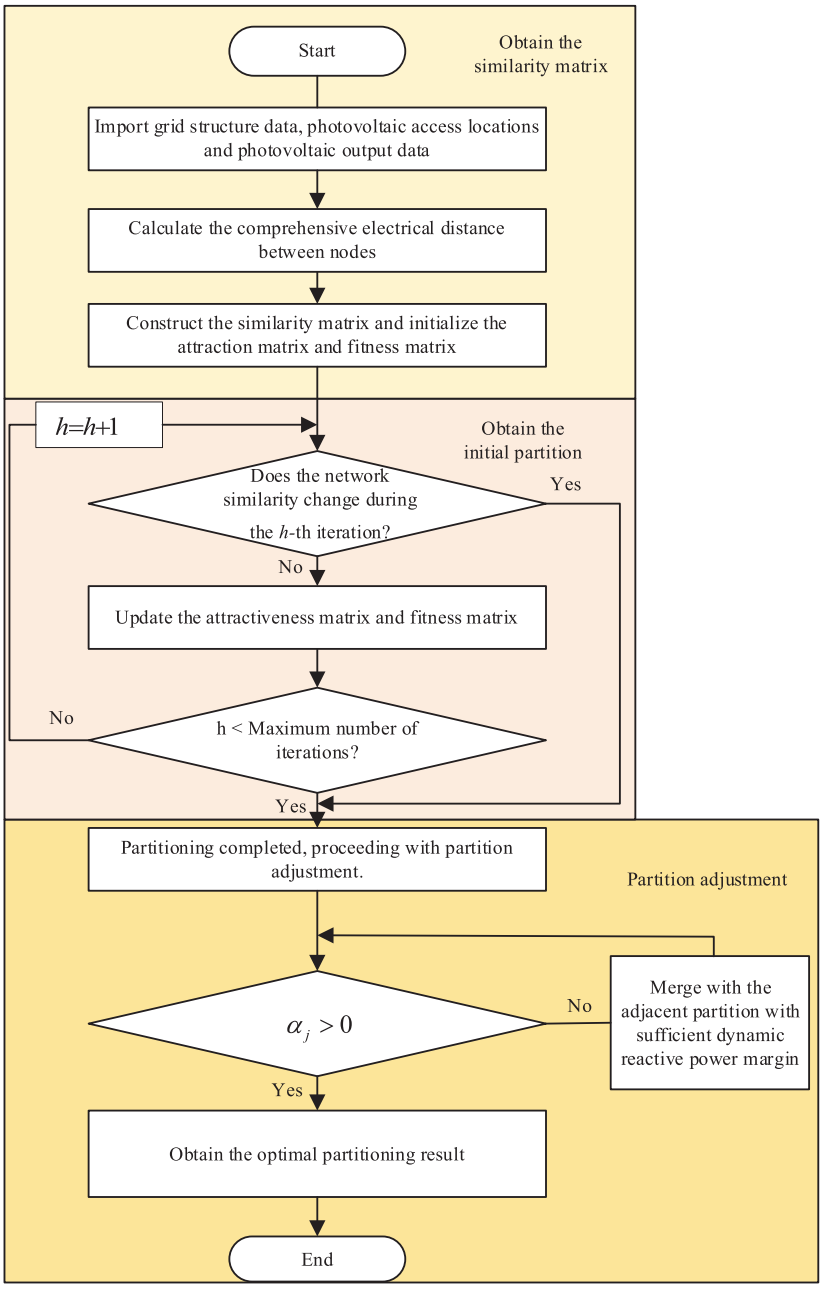

Figure A1: Flowchart of partitioning

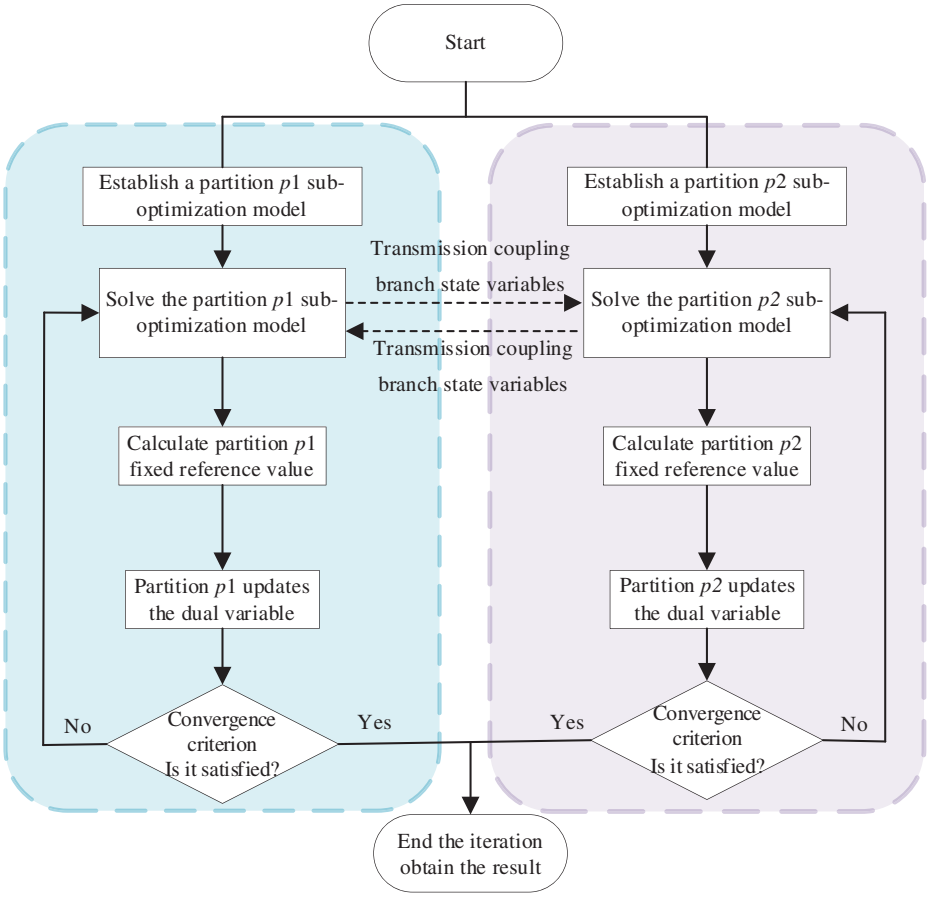

Figure A2: Distributed coordination optimization flowchart

Appendix B

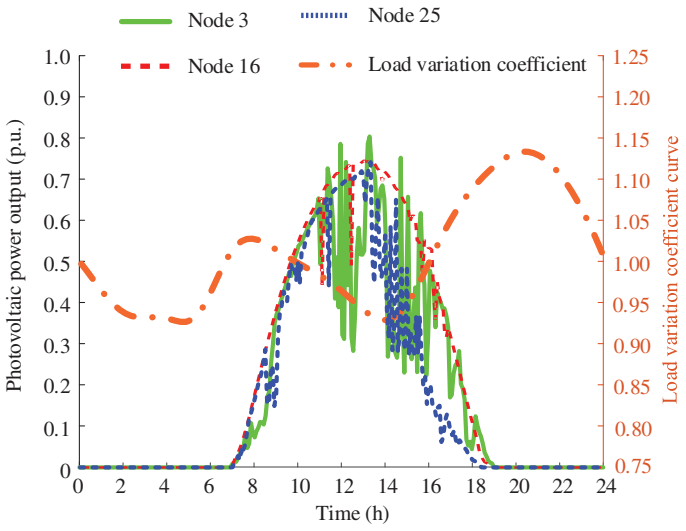

Figure A3: Photovoltaic active power output and load variation curve

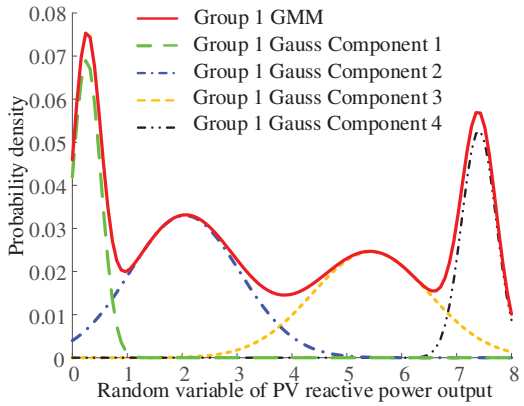

Figure A4: GMM probability density function of photovoltaic reactive power output in group 1

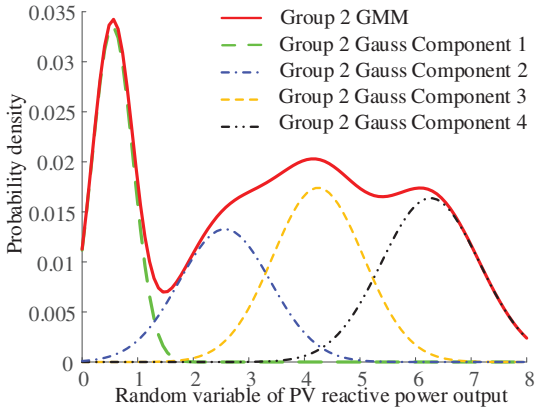

Figure A5: GMM probability density function of photovoltaic reactive power output in group 2

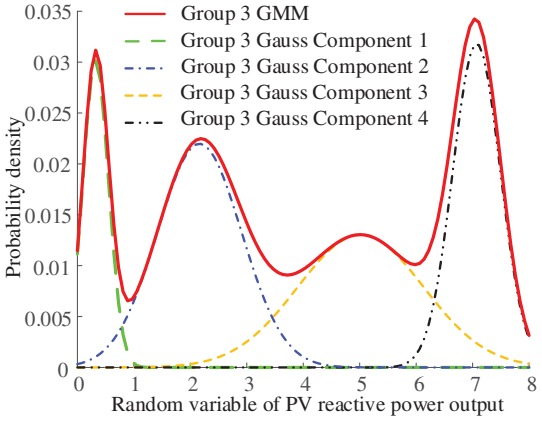

Figure A6: GMM probability density function of photovoltaic reactive power output in group 3

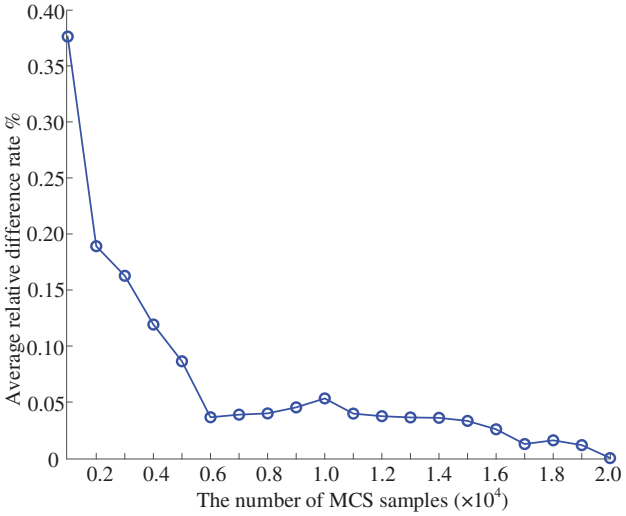

Figure A7: Convergence analysis under different numbers of MCS sampling times (with 2000 times as the benchmark)

Table A1: The impact of different weights on partitioning results

Serial number	η_1	η_2	Connectivity	Modularity
1	0.9	0.1	Not satisfied	0.5635
2	0.8	0.2	Not satisfied	0.5843
3	0.7	0.3	Not satisfied	0.6135
4	0.6	0.4	Not satisfied	0.6197
5	0.5	0.5	Not satisfied	0.6212
6	0.4	0.6	Satisfied	0.6691
7	0.3	0.7	Satisfied	0.6556
8	0.2	0.8	Satisfied	0.6556
9	0.1	0.9	Satisfied	0.6681

References

1. Luo C, Wu H, Zhou Y, Qiao Y, Cai M. Network partition-based hierarchical decentralised voltage control for distribution networks with distributed PV systems. Int J Electr Power Energy Syst. 2021;130:106929. doi:10.1016/j.ijepes.2021.106929.
2. Sun X, Qiu J, Yi Y, Tao Y. Cost-effective coordinated voltage control in active distribution networks with photovoltaics and mobile energy storage systems. IEEE Trans Sustain Energy. 2022;13(1):501–13. doi:10.1109/tste.2021.3118404.
3. Fotopoulou M, Rakopoulos D, Malamaki KN, Andriopoulos N, Lampsidis G, Kaousias K. Photovoltaic penetration potential in the Greek island of Ikaria. Sol Compass. 2024;12:100080. doi:10.1016/j.solcom.2024.100080.
4. Tian S, Jia Q, Cui Y, Xue S, Yu H, Liu W. Multi-objective collaborative optimization of VDAPFs and SVGs allocation considering MFGCIs contribution for voltage partitioning mitigation in distribution networks. Electr Power Syst Res. 2022;207:107830. doi:10.1016/j.epsr.2022.107830.
5. Wang J, Yao L, Liang J, Wang J, Cheng F. Distributed optimization strategy for networked microgrids based on network partitioning. Appl Energy. 2025;378:124834. doi:10.1016/j.apenergy.2024.124834.

6. Li J, Wang X, Guo Q, Yang C, Chen Y, Zhao R, et al. A cluster-optimized regulation method for distribution networks considering distributed photovoltaic heterogeneous characteristics. Int J Electr Power Energy Syst. 2025;170:110857. doi:10.1016/j.ijepes.2025.110857.
7. Nazir FU, Pal BC, Jabr RA. A two-stage chance constrained volt/var control scheme for active distribution networks with nodal power uncertainties. IEEE Trans Power Syst. 2019;34(1):314–25. doi:10.1109/pesgm40551.2019.8974082.
8. Antoniadou-Plytaria KE, Kouveliotis-Lysikatos IN, Georgilakis PS, Hatziargyriou ND. Distributed and decentralized voltage control of smart distribution networks: models, methods, and future research. IEEE Trans Smart Grid. 2017;8(6):2999–3008. doi:10.1109/tsg.2017.2679238.
9. Jiao W, Wu Q, Chen J, Tan J, Song G, Huang S, et al. Analytical target cascading based real-time distributed voltage control for MV and LV active distribution networks. Int J Electr Power Energy Syst. 2024;159:110024. doi:10.1016/j.ijepes.2024.110024.
10. Zhang R, Liu H, Yang M, Wang J, Shang J. Network partitioning and hierarchical voltage regulation for distribution networks using holomorphic embedding method-based sensitivity. IET Generation Trans Dist. 2023;17(3):604–20. doi:10.1049/gtd2.12745.
11. Cotilla-Sanchez E, Hines PDH, Barrows C, Blumsack S. Comparing the topological and electrical structure of the North American electric power infrastructure. IEEE Syst J. 2012;6(4):616–26. doi:10.1109/jsyst.2012.2183033.
12. Guo QL, Sun HB, Zhang BM, Wu WC. Power network partitioning based on clustering analysis in mvar control space. Autom Electr Power Syst. 2005;29(10):36–40,54. (In Chinese).
13. Bao W, Zhu T, Zhao C, Wu T, Guo RP. A three-stage network partition method for secondary voltage control based on agglomerative analysis. Autom Electr Power Syst. 2016;40(5):127–32. (In Chinese). doi:10.7500/AEPS20150506012.
14. Zheng JX, Zhong J. A complex network theory fast partition algorithm of reactive voltage based on node type and coupling of partitions. Power Syst Technol. 2020;44(1):223–30. (In Chinese). doi:10.13335/j.1000-3673.pst.2018.2654.
15. Shi J, Zhou J, Zhang Z, Liu Q, Cheng Q, Cheng S, et al. Energy storage system configuration in power distribution network considering partitioned resource coordination. Electr Power Syst Res. 2025;248:111961. doi:10.1016/j.epsr.2025.111961.
16. Zhao B, Xu Z, Xu C, Wang C, Lin F. Network partition-based zonal voltage control for distribution networks with distributed PV systems. IEEE Trans Smart Grid. 2018;9(5):4087–98. doi:10.1109/tsg.2017.2648779.
17. Dall'Anese E, Dhople SV, Giannakis GB. Photovoltaic inverter controllers seeking AC optimal power flow solutions. IEEE Trans Power Syst. 2016;31(4):2809–23. doi:10.1109/pesgm.2017.8274362.
18. Zheng W, Wu W, Zhang B, Sun H, Liu Y. A fully distributed reactive power optimization and control method for active distribution networks. IEEE Trans Smart Grid. 2016;7(2):1021–33. doi:10.1109/tsg.2015.2396493.
19. Gui Y, Nainar K, Bendtsen JD, Diewald N, Iov F, Yang Y, et al. Voltage support with PV inverters in low-voltage distribution networks: an overview. IEEE J Emerg Sel Topics Power Electron. 2024;12(2):1503–22. doi:10.1109/jestpe.2023.3280926.
20. Gogebakan M. A novel approach for Gaussian mixture model clustering based on soft computing method. IEEE Access. 2021;9:159987–60003. doi:10.1109/access.2021.3130066.
21. Dridi N, Hadzagic M. Akaike and Bayesian information criteria for hidden Markov models. IEEE Signal Process Lett. 2019;26(2):302–6. doi:10.1109/lsp.2018.2886933.
22. Li P, Zhang H, Zhao S, Wang F. An optimal allocation method for power distribution network partitions based on improved spectral clustering algorithm. Eng Appl Artif Intell. 2023;123:106497. doi:10.1016/j.engappai.2023.106497.
23. Liu F, Gu B, Qin S, Zhang K, Cui L, Xie G. Power grid partition with improved biogeography-based optimization algorithm. Sustain Energy Technol Assess. 2021;46:101267. doi:10.1016/j.seta.2021.101267.
24. Zhang C, Cai Y, Lin G, Shen C. DeepEMD: differentiable earth mover's distance for few-shot learning. IEEE Trans Pattern Anal Mach Intell. 2023;45(5):5632–48. doi:10.1109/tpami.2022.3217373.
25. Mao X, Zhu W, Wu L, Zhou B. Comparative study on methods for computing electrical distance. Int J Electr Power Energy Syst. 2021;130:106923. doi:10.1016/j.ijepes.2021.106923.

26. Isabelle Flora FM. Site suitability analysis for integrating large-scale hybrid photovoltaic and concentrated solar power plants in Cameroon using Monte Carlo, analytic hierarchy process, and geographic information system methods. Energy Convers Manage X. 2025;25:100871. doi:10.1016/j.ecmx.2025.100871.
27. Huang D, Wang CD, Peng H, Lai J, Kwoh CK. Enhanced ensemble clustering via fast propagation of cluster-wise similarities. IEEE Trans Syst Man Cybern Syst. 2021;51(1):508–20. doi:10.1109/tsmc.2018.2876202.
28. Yan L, Sheikholeslami M, Gong W, Shahidehpour M, Li Z. Architecture, control, and implementation of networked microgrids for future distribution systems. J Mod Power Syst Clean Energy. 2022;10(2):286–99. doi:10.35833/mpce.2021.000669.
29. Di Fazio AR, Risi C, Russo M, De Santis M. Coordinated optimization for zone-based voltage control in distribution grids. IEEE Trans Ind Appl. 2022;58(1):173–84. doi:10.1109/tia.2021.3129731.
30. Novoa L, Flores R, Brouwer J. Optimal DER allocation in meshed microgrids with grid constraints. Int J Electr Power Energy Syst. 2021;128:106789. doi:10.1016/j.ijepes.2021.106789.
31. Li Z, Wu Q, Chen J, Huang S, Shen F. Double-time-scale distributed voltage control for unbalanced distribution networks based on MPC and ADMM. Int J Electr Power Energy Syst. 2023;145:108665. doi:10.1016/j.ijepes.2022.108665.
32. Hashish MS, Hasanien HM, Ji H, Alkuhayli A, Alharbi M, Akmaral T, et al. Monte Carlo simulation and a clustering technique for solving the probabilistic optimal power flow problem for hybrid renewable energy systems. Sustainability. 2023;15(1):783. doi:10.3390/su15010783.