

Neural Spectral Clustering Based Voltage Area Partition of Active Distribution Systems

Shengyi Wang, Liang Du, Lianren Zhu, and Yan Li

Abstract—As distribution systems evolve to accommodate large-scale renewable energy sources, maintaining voltage stability becomes increasingly challenging. Network partitioning plays a pivotal role in voltage control tasks, especially in active distribution systems (ADSs). By partitioning the network into manageable small sub-networks, i.e., voltage area partition (VAP), fine-grained, decentralized, and coordinated voltage control can be realized, which prevents over-voltage or under-voltage issues and facilitates the integration and absorption of renewable energies. However, because of the weak ability to extract complicated voltage relationships, existing naive graph clustering VAP methods are likely to suffer a performance bottleneck in voltage cohesiveness for large-sized distribution networks. Therefore, this paper proposes a neural spectral clustering-based VAP method for ADSs. Specifically, a network partition problem is solved by clustering a neural spectral mapping of multi-phase voltage coupling features. Theoretical and experimental results show that the proposed method can partition the network with voltage cohesiveness higher than that of the standard spectral clustering method while bringing certain advantages in computational efficiency.

I. INTRODUCTION

Large-scale integration of renewable energies to distribution networks will become a trend, which transforms the role of distribution systems from passive energy consumers into active energy prosumers. Although such a deployment benefits the environment and human sustainable development, it brings numerous operation challenges to ADSs. Among them, voltage quality issues need to be particular attention. Normally, the high R/X ratio of distribution networks makes buses prone to voltage deviations when active power flows through power distribution lines. Due to the large reverse power flow produced by renewable energy sources, the voltage deviation may deteriorate and even arouses over-voltage issues [1], [2].

Conventional voltage regulation (VR) is achieved by optimizing on-load tap changers, switched capacitors, and voltage regulators, which are mechanical in nature, and thus suffer from slow response speed and limited switching frequencies, making them inadequate in promptly mitigating fast voltage fluctuation and addressing voltage violations in a short time. In contrast, controllable distributed energy resources (DERs) have opened new avenues for more flexible and fast voltage control in ADSs. For example, PV inverters

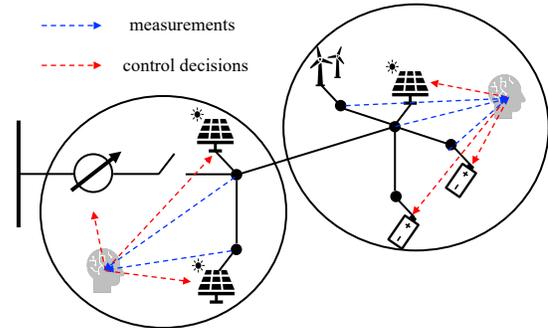


Fig. 1. Illustration for the network partition-based distributed voltage control strategies

can participate in VR while losing some active power-providing capacity [3]. Technically, this type of VR device changes the distribution of power flow by compensating distribution networks with granular active and reactive power.

Based on their communication architectures, existing active-reactive power voltage control strategies via controllable DERs can be categorized into centralized, distributed, and local methods [4]. Centralized methods rely on global information for decision-making, which is computationally costly and may involve privacy and safety issues. Unlike centralized methods, distributed and local methods have fewer communication demands, which are more appropriate for a complicated decision-making environment. To efficiently manage large-scale controllable DERs for VR, inspired by VR in bulk transmission systems, some researchers have made efforts to network partition-based distributed voltage control strategies [5], [6].

As for network partition-based distributed voltage control strategies, as shown in Fig. 1, the entire distribution network is divided into multiple areas, each of which is assigned a voltage control agent who can make control decisions based on measurements within its affiliated area. A critical step involved in this type of voltage control strategy is VAP. In essence, VAP is employed to decompose a centralized VR problem into multiple regional VR problems. Proper VAP results are treated as the premise of achieving satisfying VR performance, which can ensure low interdependence between areas in terms of voltage magnitudes, i.e., bus voltage magnitude changes within one area have minimal impact on others, and buses in the same area have a stronger voltage-coupling level than others. Therefore, it is valuable to develop effective VAP methods.

The majority of existing works convert VAP in ADSs to a

S. Wang and L. Du are with the Department of Electrical and Computer Engineering, Temple University, Philadelphia, USA

L. Zhu is with the Department of Electrical Engineering, Shanghai University, Shanghai, China

Y. Li is with the Department of Electrical Engineering, The Pennsylvania State University, University Park, USA

graph clustering task. Specifically, an electrical distance that reflects the voltage coupling degree is first defined, and then based on it either a modularity metric [7]–[10] or a cut-based metric [11]–[13] is optimized to cluster buses with close electrical distances. However, when the distribution network size is large, these VAP methods may have difficulties in capturing complicated voltage similarity relations between buses and between phases. Therefore, they are likely to suffer a degradation of voltage cohesiveness performance, which means that there may exist a better partition result with much closer voltage relations in each area. It is worth noting that such an improvement is essential because more efficient local adjustment can be found to save the voltage regulation cost, i.e., the same voltage regulation performance with less power compensation. Furthermore, as the number of buses increases to tens of thousands, the computational expenditure in optimizing the graph quality functions, such as the inverse calculation of spectral clustering, is very high.

To improve the adaptability and performance of VAP methods in large-sized distribution networks, this paper proposes a neural spectral clustering-based VAP method for ADSs. The contributions are two-fold. First, unlike standard spectral clustering methods, a neural network is introduced to efficiently learn a graph spectral embedding, which is realized by stochastic batch optimization, with the affordable computational cost in inverse calculation, and used for clustering in VAP later. Secondly, a Siamese neural network is developed to learn the complicated bus voltage affinity relations.

The remainder of this paper is organized as follows. Section II is a formulation of the VAP problem. Section III proposes a neural spectral clustering-based VAP method, with numerical test results demonstrated in Section VI. Finally, Section V concludes this paper.

II. VOLTAGE AREA PARTITION PROBLEM

The voltage-magnitude sensitivity coefficients, i.e., elements of the Jacobian matrix of voltage magnitudes with reference to power injections, and are widely applied in voltage control tasks for ADSs. Therefore, they are utilized here to evaluate the voltage coupling degree. As for distribution networks, voltage-magnitude sensitivity coefficients are relative to both active and reactive injections. Assume that the majority of controllable sources in the system are wye-connected, and thus the effect of delta-connected source power injections is not considered in the voltage coupling degree. Mathematically, for a generic three-phase distribution network with one slack bus and N three-phase PQ buses, the per-phase voltage coupling degree can be defined as

$$\xi_{ij}^{\phi} = \frac{\left| \frac{\partial V_i^{\phi}}{\partial p_j^{\phi}} \right| + \left| \frac{\partial V_i^{\phi}}{\partial q_j^{\phi}} \right|}{\left| \frac{\partial V_j^{\phi}}{\partial p_i^{\phi}} \right| + \left| \frac{\partial V_j^{\phi}}{\partial q_i^{\phi}} \right|} \cdot \frac{\left| \frac{\partial V_j^{\phi}}{\partial p_i^{\phi}} \right| + \left| \frac{\partial V_j^{\phi}}{\partial q_i^{\phi}} \right|}{\left| \frac{\partial V_i^{\phi}}{\partial p_j^{\phi}} \right| + \left| \frac{\partial V_i^{\phi}}{\partial q_j^{\phi}} \right|}, \quad \forall \phi \in \{a, b, c\} \quad (1)$$

where V_i^{ϕ} is voltage magnitude at phase ϕ at bus i , p_i^{ϕ} is active power injection at phase ϕ at bus i , q_i^{ϕ} is reactive

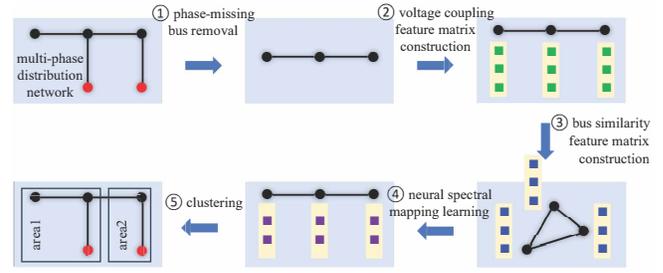


Fig. 2. Illustration for the proposed voltage area partition method

power injection at phase ϕ at bus i , $\frac{\partial V_i^{\phi}}{\partial p_i^{\phi}}$ ($\frac{\partial V_i^{\phi}}{\partial q_i^{\phi}}$) is voltage-active (voltage-reactive) power sensitivity, which can be achieved by the power flow model.

Note that $\xi_{ij}^{\phi} \in [0, 1]$, and the larger the value of ξ_{ij}^{ϕ} is, the closer voltage connection both buses at phase ϕ have. Let matrix $\xi^{\phi} := (\xi_{ij}^{\phi})_{N \times N}$ collect all values of the voltage coupling degree at phase ϕ . Since $\xi_{ij}^{\phi} = \xi_{ji}^{\phi}$, ξ^{ϕ} is symmetric. Define $\xi := (\xi^a, \xi^b, \xi^c)_{N \times 3N}$ as the voltage coupling feature matrix, of which each row ξ_i can be viewed as voltage coupling features of bus i .

For network partition-based distributed voltage control, each voltage area is assigned a regional voltage controller that is responsible for regulating the voltage magnitudes of a set of buses. Therefore, a voltage area partition problem is designed as finding a clustering scheme so that a set of buses sharing similar voltage coupling features can be grouped and partitioned into a voltage area. Although the standard spectral clustering method can be applied to solve this problem, with large distribution networks, direct computation of eigenvectors is computationally expensive. Meanwhile, ADSs tend to bear complicated bus voltage affinity relations and overly simplistic measure of similarity may lead to poor clustering results, which also motivates this work.

III. VOLTAGE AREA PARTITION METHOD

In order to solve the above-designed problem, a neural spectral clustering-based VAP method for ADSs is proposed and illustrated in Fig. 2, which consists of the following five steps:

- 1) First, the phase-missing buses are temporarily removed. For a multi-phase distribution network, there are single-phase and two-phase buses dispersed at the end of the feeder. Most of them could arouse several isolated buses in the VAP results. It is not desired because there is no physical connection among these buses, which makes VAP results less practical. Although removed, the phase-missing buses can still affect the clustering results through voltage-magnitude sensitivity coefficients.
- 2) Second, for the remaining three-phase buses, the voltage coupling feature matrix is constructed based on Eqn. (1).
- 3) Then, a Siamese network is designed to capture the affinity of bus voltage coupling features so that a bus

similarity feature matrix with a graph is constructed. More details can be found in Subsection A.

- 4) After that, the voltage coupling features can be represented in a dimension-reduced space by learning a neural spectral mapping on the graph. More details can be found in Subsection B.
- 5) Finally, as in standard spectral clustering methods, VAP results are determined by clustering algorithms with dimension-reduced features. The phase-missing buses can be naturally partitioned into the nearest voltage area.

A. Bus Similarity Matrix Construction

Denote a bus similarity feature matrix as $\mathbf{W} := (W_{ij})_{N \times N}$, and each element W_{ij} measures the affinity of voltage coupling features of three-phase buses i and j , i.e., ξ_i and ξ_j . In many applications, the Euclidean distance with the Gaussian kernel is utilized to measure the affinity of data points, i.e.,

$$W_{ij} = e^{-\frac{\|\xi_i - \xi_j\|_2^2}{2\sigma^2}} \quad (2)$$

However, this affinity measure is overly simplistic, and it only considers the local affinity relations. In the meantime, it is not easy to choose an appropriate value for bandwidth σ .

In the proposed method, a Siamese network is designed to measure global affinity, i.e., W_{ij} is determined by all bus voltage coupling features. The Siamese network is represented by a neural network that takes voltage coupling feature ξ_i as an input, and maps it into an embedding space, i.e., $\mathbf{z}_i = f_{\theta_{\text{siamese}}}(\xi_i)$ where θ_{siamese} collects all neural network weights. The role of this embedding is to predict the affinity, i.e.,

$$W_{ij} = \frac{2}{1 + e^{\|\mathbf{z}_i - \mathbf{z}_j\|_2^2}} \in [0, 1] \quad (3)$$

When two buses have similar voltage coupling features, the value of the Euclidean distance $\|\mathbf{z}_i - \mathbf{z}_j\|_2$ should be small, and the affinity value W_{ij} should be large and vice versa.

To learn such an embedding, the network can be trained by minimizing the following cross-entropy loss function

$$L_{\text{siamese}}(\theta_{\text{siamese}}; \xi_i, \xi_j) = -\hat{W}_{ij}(\log(W_{ij})) - (1 - \hat{W}_{ij})(\log(1 - W_{ij})) \quad (4)$$

where $\hat{W}_{ij} \in \{0, 1\}$ is a label for a pair of points ξ_i and ξ_j , which constitutes matrix $\hat{\mathbf{W}} := (\hat{W}_{ij})_{N \times N}$. The label value \hat{W}_{ij} can be determined by the following rule

$$\hat{W}_{ij} = \begin{cases} 1 & : \|\xi_i - \xi_j\|_2 \leq \varepsilon \\ 0 & : \|\xi_i - \xi_j\|_2 > \varepsilon \end{cases} \quad (5)$$

where ε is a margin. If the value of the Euclidean distance between the points ξ_i and ξ_j is small, their label is assigned as 1 and vice versa. Note that the labeling via Eqn. (5) is a local affinity measure that is a priori to help the Siamese network learn consistent affinity relations, i.e., global affinity relations should be consistent with local ones.

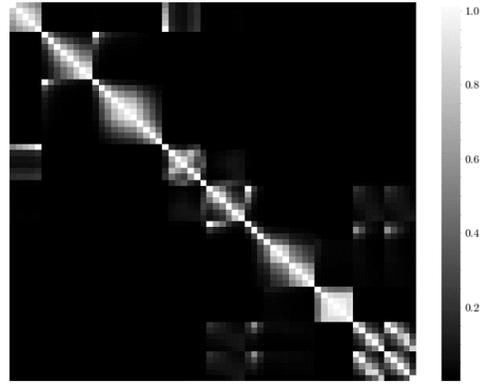


Fig. 3. Illustration of an example bus similarity matrix learned from the proposed Siamese network

In addition, the output value W_{ij} is a real number that provides a more fine-grained measure of voltage affinity relations between buses i and j . Once the network is trained, the desired bus similarity matrix can be achieved, as shown in Fig. 3. It can be shown that this matrix is symmetric, and the diagonal light part indicates that the adjacent buses tend to share more similar voltage coupling features. The specific procedures for constructing the proposed bus similarity matrix are summarized in Algorithm. 1.

Algorithm 1: An algorithm for constructing the proposed bus similarity matrix

Data: matrices ξ and $\hat{\mathbf{W}}$, and learning rates η_{siamese}

Result: matrix \mathbf{W}

- 1 Initialize the neural network weights θ_{siamese} ;
 - 2 **while** not converge **do**
 - 3 **for each pair** (ξ_i, ξ_j) **do**
 - 4 using Eqn. 3 to compute W_{ij} ;
 - 5 using gradient decent to update $\theta_{\text{siamese}} \leftarrow \theta_{\text{siamese}} - \eta_{\text{siamese}} \frac{\partial L_{\text{siamese}}(\theta_{\text{siamese}}; \xi_i, \xi_j)}{\partial \theta_{\text{siamese}}}$;
 - 6 **end**
 - 7 **end**
 - 8 **for each pair** (ξ_i, ξ_j) **do**
 - 9 using Eqn. 3 to update W_{ij} ;
 - 10 **end**
-

B. Neural Spectral Mapping Learning

The afore-learned bus similarity feature matrix involves a fully-connected undirected graph, which measures the affinity of different buses in terms of voltage coupling features. Following the conventional practice of spectral clustering, the next step is to embed the data of voltage coupling features in the eigenspace of the Laplacian matrix, i.e., spectral mapping. While spectral mapping of data points can be achieved by a simple eigen-decomposition of their graph Laplacian matrix, with large datasets direct computation of eigenvectors may be computationally expensive, especially in a multi-phase distribution network with tens of thousands

of buses. Therefore, neural spectral mapping is resorted to strengthen the efficiency of the proposed VAP method.

Given $\mathbf{W}' := (W'_{ij})_{m \times m}$ corresponding a mini-batch that can be silced from \mathbf{W} , feature points ξ_i and ξ_j with large W'_{ij} need to be embedded close and orthonormal to each other. Hence, a neural spectral mapping mini-batch learning problem is formulated by

$$\begin{aligned} \min_{\theta_{\text{spectral}}} L_{\text{spectral}}(\theta_{\text{spectral}}; \xi_i, \xi_j) \quad \text{s.t.} \\ \mathbf{u}_i = f_{\theta_{\text{spectral}}}(\xi_i); \\ \mathbf{u}_j = f_{\theta_{\text{spectral}}}(\xi_j); \\ \frac{1}{m} \mathbf{U}^\top \mathbf{U} = \mathbf{I}_{k \times k} \end{aligned} \quad (6)$$

where $f_{\theta_{\text{spectral}}} : \mathbb{R}^{3N} \rightarrow \mathbb{R}^k$ is a general neural network mapping the input features into a spectral embedding space, k is the number of clusters or voltage areas customized by users, θ_{spectral} collects all neural network weights, \mathbf{u}_i and \mathbf{u}_j are neural spectral mappings, $\mathbf{U} \in \mathbb{R}^{m \times k}$ is a output matrix whose the i -th row are \mathbf{u}_i^\top , m is the number of feature points in a mini-batch, \mathbf{I} is an identity matrix, and the objective function L_{spectral} is designed as

$$\begin{aligned} L_{\text{spectral}}(\theta_{\text{spectral}}; \xi_i, \xi_j) \\ = \frac{1}{m^2} \sum_{i,j=1}^m W'_{ij} \|\mathbf{u}_i - \mathbf{u}_j\|_2^2 = \frac{1}{m^2} \sum_{i,j=1}^m W'_{ij} \sum_{p=1}^k (u_{ip} - u_{jp})^2 \\ = \frac{1}{m^2} \sum_{p=1}^k [\sum_{i=1}^m d_i u_{ip}^2 - 2 \sum_{i,j=1}^m W'_{ij} u_{ip} u_{jp} + \sum_{j=1}^m d_j u_{jp}^2] \\ = \frac{2}{m^2} \sum_{p=1}^k [\sum_{i=1}^m d_i u_{ip}^2 - \sum_{i,j=1}^m W'_{ij} u_{ip} u_{jp}] \\ = \frac{2}{m^2} \sum_{p=1}^k [\check{\mathbf{u}}_p^\top (\mathbf{D} - \mathbf{W}') \check{\mathbf{u}}_p] \\ = \frac{2}{m^2} \text{trace}(\mathbf{U}^\top (\mathbf{D} - \mathbf{W}') \mathbf{U}), \end{aligned} \quad (7)$$

where u_{ip} is the p -th element in \mathbf{u}_i , $d_i = \sum_{j=1}^m W'_{ij}$, $\mathbf{D} = \text{diag}(d_i) \in \mathbb{R}^{m \times m}$ is a diagonal matrix, and $\check{\mathbf{u}}_p$ is the p -th column of \mathbf{U} . Note that the Eqn. 7 is an approximation of the ratio cut objective in the standard spectral clustering, i.e., partitioning a subgraph into disjoint subsets that are internally cohesive and externally separate, where the i -th element in $\check{\mathbf{u}}_p$ can indicate whether bus i belongs to area p , and $\mathbf{D} - \mathbf{W}'$ corresponds to the graph Laplacian.

To solve the above-formulated neural mapping learning problem, the constrained optimization problem needs to be converted to an unconstrained problem. The first two constraints in Eqn. (6) can be removed by inserting them into the objective (7) while the last constraint, i.e., the orthogonality constraint, can be removed by using an orthonormal layer as the output layer in the neural network, as shown in Fig. 4. This orthonormal layer can enforce the orthogonality constraint, which takes k neurons as inputs and acts as a linear layer with k outputs.

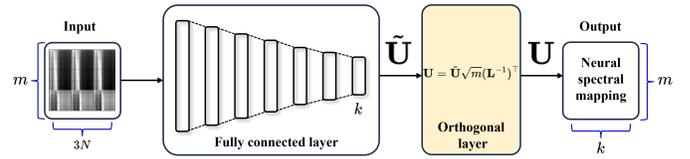


Fig. 4. Illustration of the neural network architecture for spectral mapping

Algorithm 2: An algorithm for learning the neural spectral mapping

Data: matrices ξ and \mathbf{W} , m , k and learning rates

η_{spectral}

Result: neural spectral mappings

- 1 Initialize the neural network weights θ_{spectral} ;
 - 2 **while** *not converge* **do**
 - 3 randomly sample m rows of ξ ;
 - 4 using $f_{\theta_{\text{spectral}}}$ to compute $\tilde{\mathbf{U}}$;
 - 5 using Cholesky decomposition to obtain \mathbf{L} ;
 - 6 using Orthonormal layer to compute \mathbf{U} ;
 - 7 using gradient decent to update
 $\theta_{\text{spectral}} \leftarrow \theta_{\text{spectral}} - \eta_{\text{spectral}} \frac{\partial L_{\text{spectral}}(\theta_{\text{spectral}}; \xi_i, \xi_j)}{\partial \theta_{\text{spectral}}}$;
 - 8 **end**
 - 9 **for** *every* m rows in ξ **do**
 - 10 repeat lines 4,5,6;
 - 11 **end**
-

Unlike fully connected neural network layers, weights in the orthonormal layer can be computed through QR decomposition as $\sqrt{m}(\mathbf{L}^{-1})^\top$ where matrix \mathbf{L} is a lower triangular matrix and can be obtained via the Cholesky decomposition, i.e., $\tilde{\mathbf{U}}^\top \tilde{\mathbf{U}} = \mathbf{L}\mathbf{L}^\top$, and matrix $\tilde{\mathbf{U}} \in \mathbb{R}^{m \times k}$ denotes the inputs to the orthonormal layer. After multiplying $\tilde{\mathbf{U}}$ from the right by weights $\sqrt{m}(\mathbf{L}^{-1})^\top$, columns of $\tilde{\mathbf{U}}$ can be orthogonalized into \mathbf{U} .

The neural network $f_{\theta_{\text{spectral}}}$ can be trained by solving the above-reformulated optimization problem in a gradient descent way. In each step, a minibatch of m feature points is randomly sampled. After forward propagation, the weights of the orthogonal layer are updated first, and then back-propagate the gradients $\frac{\partial L_{\text{spectral}}}{\partial \theta_{\text{spectral}}}$ to update the remaining weights in $f_{\theta_{\text{spectral}}}$. Once this neural network is trained, all the weights are frozen, and the entire voltage coupling feature matrix is divided into multiple mini-batches, sequentially sent to the network as the input to achieve the desired spectral mapping used for clustering. The specific procedures for learning the proposed neural spectral mapping are summarized in Algorithm. 2.

It is noted that the minibatch training involved in neural spectral mapping learning significantly save the computational and storage cost of the inverse calculation or the eigenvectors.

IV. NUMERICAL SIMULATION

The effectiveness of the proposed method in this paper is verified using IEEE 123 bus distribution systems. The

number of voltage areas is set as 6. The maximum number of iterations is set as 50, which is used to control the termination of the training. The batch sizes are set as 120 and 16 for neural networks $f_{\theta_{\text{Siamese}}}$ and $f_{\theta_{\text{spectral}}}$, respectively. In terms of the structure of these two networks, given that the dimension of their input data is the same, i.e., 192, we design a similar network architecture with all fully connected layers. The numbers of neurons in the Siamese network are designed as [192, 168, 128, 64, 32, 16], while the numbers of neurons in the Spectral neural network are designed as [192, 168, 128, 64, 32, 16, 6].

Furthermore, the multi-phase power flow model developed in [14] is used to find voltage-active (voltage-reactive) power sensitivities. The standard spectral clustering method is selected as a benchmark method with Gaussian kernel bandwidth set as 2.4. All the programs are written in Python and performed on a Windows PC equipped with AMD Ryzen 5 5600H, 3.30 GHz, NVIDIA GeForce RTX 3070 GPU, and 16.0 GB RAM.

To evaluate the voltage cohesiveness, an index called voltage magnitude increment is designed as $\frac{V_{i2}^{\phi} - V_{i1}^{\phi}}{V_{i0}^{\phi} - V_{i1}^{\phi}}$ for each phase at each bus, where V_{i1}^{ϕ} is the voltage magnitude at the original operation point, V_{i0}^{ϕ} is the referenced voltage magnitude, and V_{i2}^{ϕ} is the voltage magnitude at the new operating point caused by injecting some active and reactive power at the pilot nodes. Each voltage area has only one pilot node that can be selected as the node with the maximum total voltage coupling degrees with other buses in the situated area. Pilot nodes vary with VAP results. In general, lower the variance of all indexes in a voltage area implies higher voltage cohesiveness.

When the training data is similar to the actual operating point, VAP results achieved by the standard spectral clustering and the proposed neural spectral clustering-based method are shown in Fig. 5 and Fig. 6, respectively. Different areas are colored in different colors. Fig. 7 shows the neural spectral mapping reconstructed in a PCA feature space. It can be seen that most of the feature points in the same cluster are similar, and the spectral mapping learned by the trained neural network is meaningful. Fig. 8 shows that in the same voltage area (with the same color), the variance of all indexes is relatively small. The average index variance can be calculated as the average of all cluster index variances. Table. I shows that the average index variance of the proposed method under different pilot node selections (i.e., taking on either spectral clustering's pilot nodes or neural spectral clustering's ones) is lower, reflecting the overall performance of voltage cohesiveness is better than the standard spectral clustering method.

When the training data is different from the actual operating point, Table. II shows that even if the learned neural spectral mapping is not updated to the new operating point, it still can have a better performance in terms of voltage cohesion. Meanwhile, there is a negligible voltage cohesion performance difference between the previously trained and retrained neural spectral clustering networks.

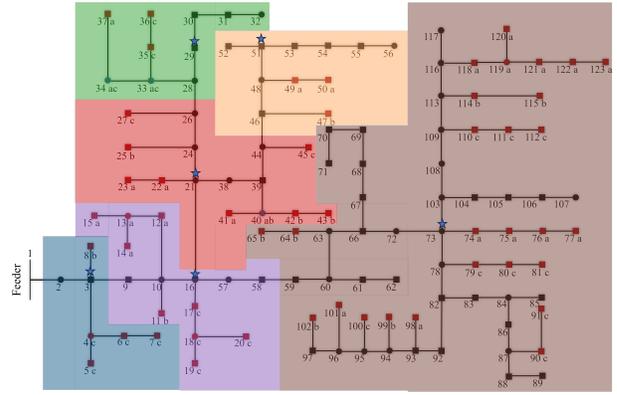


Fig. 5. Illustration of VAP results from the standard spectral clustering method

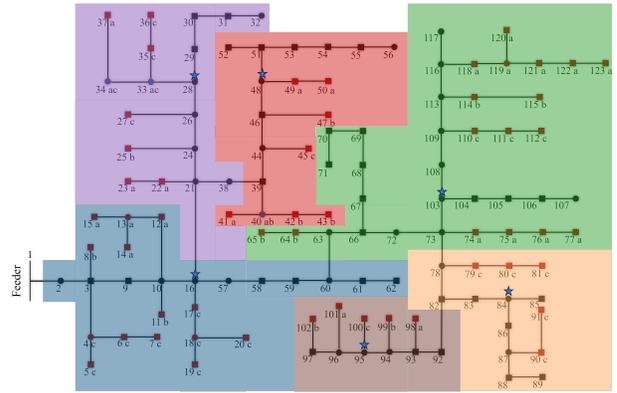


Fig. 6. Illustration of VAP results from the proposed neural spectral clustering-based method

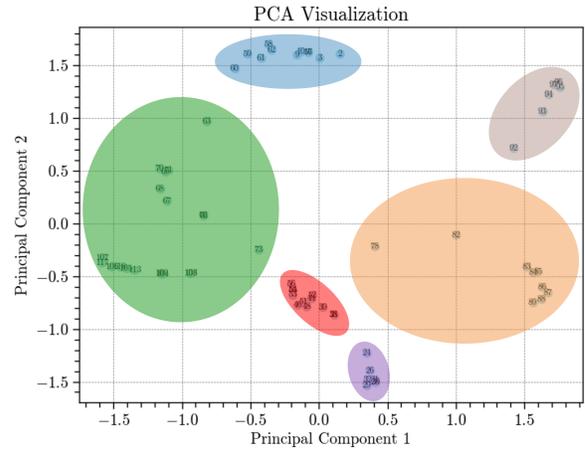


Fig. 7. Illustration of PCA Visualization Results from the proposed neural spectral clustering-based method

V. CONCLUSION

This paper proposes a novel VAP method to enhance its applicability and performance in ADSs. The VAP problem is first formulated as a graph clustering problem, where a voltage coupling feature matrix is defined. Via a Siamese network, complicated bus voltage affinity relations can be captured into a bus similarity matrix. Instead of conducting eigen-decomposition on the entire graph, an unconstrained

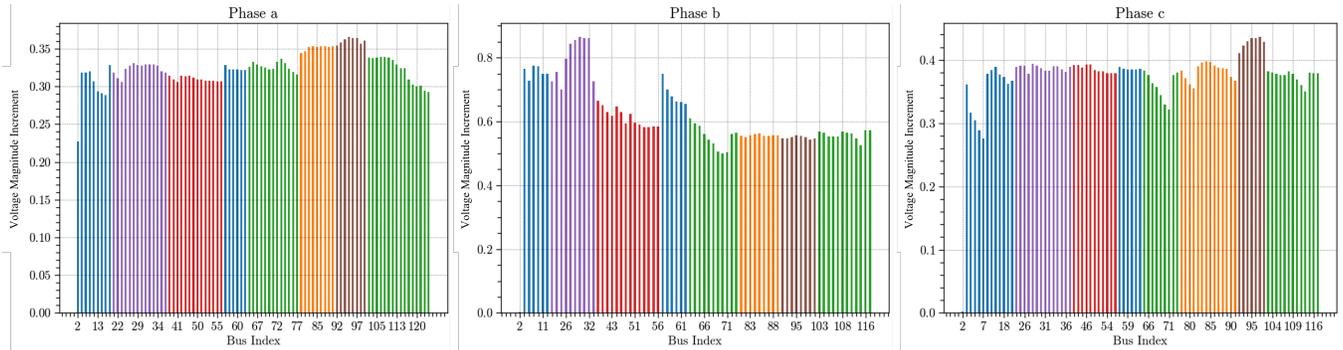


Fig. 8. Illustration of Evaluation Index Results from the proposed neural spectral clustering-based method

TABLE I

AVERAGE VARIANCE PER CLUSTER OF DIFFERENT PHASE

Method and Pilot Nodes	Average Variance among Clusters			
	Phase a	Phase b	Phase c	Phase abc
Spectral clustering + Spectral clustering pilot nodes	0.000296	0.018170	0.001927	0.006798
Neural spectral clustering + Spectral clustering pilot nodes	0.000115	0.006855	0.000922	0.002631
Spectral clustering + Neural spectral clustering pilot nodes	0.000451	0.021425	0.002420	0.008099
Neural spectral clustering + Neural spectral clustering pilot nodes	0.000152	0.007316	0.001369	0.002946

TABLE II

AVERAGE VARIANCE PER CLUSTER OF DIFFERENT PHASE

Method	Average Variance among Clusters			
	Phase a	Phase b	Phase c	Phase abc
Spectral clustering	0.000393	0.018825	0.002063	0.007094
Previously-trained neural spectral clustering	0.000133	0.006445	0.001171	0.002583
Retrained neural spectral clustering	0.000131	0.005469	0.001129	0.002243

optimization problem with an orthogonal layer is designed to learn efficient neural spectral mapping. Experiment results validate the better performance of the proposed method in terms of voltage cohesiveness compared to that of the standard spectral clustering method. Furthermore, mini-batch learning of neural spectral mappings improves the efficiency of optimizing the graph quality functions. Designing a learning objective for neural spectral mapping to make each area have a balanced regulation capacity is left for future research.

REFERENCES

- [1] K. E. Antoniadou-Plytaria, I. N. Kouveliotis-Lysikatos, P. S. Georgilakis, and N. D. Hatziaargyriou, "Distributed and decentralized voltage control of smart distribution networks: Models, methods, and future research," *IEEE Transactions on smart grid*, vol. 8, no. 6, pp. 2999–3008, 2017.
- [2] S. Wang, L. Du, X. Fan, and Q. Huang, "Deep reinforcement scheduling of energy storage systems for real-time voltage regulation in unbalanced lv networks with high pv penetration," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 4, pp. 2342–2352, 2021.
- [3] L. Chen, N. Liu, S. Yu, and Y. Xu, "A stochastic game approach for distributed voltage regulation among autonomous pv prosumers," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 776–787, 2022.
- [4] H. Sun *et al.*, "Review of challenges and research opportunities for voltage control in smart grids," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2790–2801, 2019.
- [5] P. Li, C. Zhang, Z. Wu, Y. Xu, M. Hu, and Z. Dong, "Distributed adaptive robust voltage/var control with network partition in active distribution networks," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2245–2256, 2019.
- [6] S. Wang, J. Duan, D. Shi, C. Xu, H. Li, R. Diao, and Z. Wang, "A data-driven multi-agent autonomous voltage control framework using deep reinforcement learning," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4644–4654, 2020.
- [7] H. Ruan, H. Gao, Y. Liu, L. Wang, and J. Liu, "Distributed voltage control in active distribution network considering renewable energy: A novel network partitioning method," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4220–4231, 2020.
- [8] S. Wang, B. Cui, and L. Du, "An efficient power flexibility aggregation framework via coordinate transformation and chebyshev centering optimization," in *2022 IEEE Power & Energy Society General Meeting (PESGM)*, 2022, pp. 1–5.
- [9] D. Cao, J. Zhao, W. Hu, F. Ding, Q. Huang, Z. Chen, and F. Blaabjerg, "Data-driven multi-agent deep reinforcement learning for distribution system decentralized voltage control with high penetration of pvs," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4137–4150, 2021.
- [10] T. M. Aljohani, A. Saad, and O. A. Mohammed, "Two-stage optimization strategy for solving the vvo problem considering high penetration of plug-in electric vehicles to unbalanced distribution networks," *IEEE Trans. Industry Applications*, vol. 57, no. 4, pp. 3425–3440, 2021.
- [11] D. Cao *et al.*, "Attention enabled multi-agent drl for decentralized voltage control of active distribution system using pv inverters and svcs," *IEEE Trans. Sustainable Energy*, vol. 12, no. 3, pp. 1582–1592, 2021.
- [12] —, "Deep reinforcement learning enabled physical-model-free two-timescale voltage control method for active distribution systems," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 149–165, 2021.
- [13] A. Bedawy, N. Yorino, K. Mahmoud, Y. Zoka, and Y. Sasaki, "Optimal voltage control strategy for voltage regulators in active unbalanced distribution systems using multi-agents," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1023–1035, 2019.
- [14] A. Bernstein and E. Dall'Anese, "Linear power-flow models in multiphase distribution networks," in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*. IEEE, 2017, pp. 1–6.