Probabilistic Physics-Informed Graph Convolutional Network for Active Distribution System Voltage Prediction

Tong Su, *Student Member, IEEE*, Junbo Zhao, *Senior Member, IEEE*, Yansong Pei, *Student Member, IEEE*, Fei Ding, *Senior Member, IEEE*

***Abstract*—This paper proposes a novel data-driven probabilistic physics-informed graph convolutional network (GCN) for active distribution system voltage prediction with PVs and EVs. It leverages both measurements and network topology to accurately and efficiently predict node voltages without the need for an accurate distribution system power flow model. The dropout-enabled Bayesian inference is developed to achieve uncertainty quantification of the voltage prediction. Thanks to the network model embedding, it also has robustness against topology changes, a key difference with existing machine learning-based approaches. Comparison results with other state-of-the-art machine learning methods on a realistic 759-node distribution system demonstrate that the proposed method can achieve better accuracy and robustness under different scenarios.**

***Index Terms*—Physics-informed graph convolutional network (GCN), probabilistic voltage prediction, distribution system.**

I. Introduction

THE increasing penetration of PVs and EVs has introduced more uncertainties into the active distribution system (ADS), especially in voltage variations [1]. Accurate voltage predictions can enable predictive voltage control to maintain distribution system security. Thanks to the wide deployment of smart meters, the acquisition of extensive historical data become available for voltage predictions. Traditionally, the voltage magnitudes can be calculated via power flow models using predicted loads and distributed energy resources. However, the distribution system model is subject to various errors and difficult to maintain good accuracy. To this end, data-driven voltage prediction methods have been proposed, including random forests [2], long short-term memory network [3], graph convolutional network (GCN) [4] etc. These methods are all subject to prediction errors that can lead to control issues. Therefore, quantifying prediction uncertainty is crucial to mitigate such risks. Probabilistic voltage prediction methods include probabilistic extreme learning machine [5], federated learning [6], Gaussian process [7]-based ones, etc. Compared to these methods, GCN has advantages in handling topological changes, but it is challenging to derive its probabilistic form. This paper proposes a probabilistic physics-informed GCN-based voltage prediction method. The main contributions are as follows: 1) A physics-informed node voltage prediction method is proposed, which firstly utilizes the improved version of GCN, i.e., GCN via initial residual and identity mapping (GCNII) [8]. 2) The dropout is integrated with GCNII to qualify the prediction uncertainty. 3) The missing data of ADS loads and EV measurements are also considered. Results on a real 759-node ADS demonstrate that the proposed method outperforms other state-of-the-art machine learning methods in accuracy and robustness under various scenarios.

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Solar Energy Technologies Office Agreement Number 37770. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes (*Corresponding Author*: Junbo Zhao).

T. Su, J. Zhao and Y. Pei are with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT, 06269 (e-mail: tongsu@uconn.edu, junbo@uconn.edu, yansong.pei@uconn.edu).

F. Ding is with the National Renewable Energy Laboratory, Golden, CO 80401. (e-mail: fei.ding@nrel.gov).

II. Proposed Probabilistic Prediction Method

A. Problem Statement

Due to the challenge of obtaining accurate physical parameters of ADS and the widespread deployment of smart meters, data-driven voltage prediction has emerged as a model-based alternative and has considerable attention. Data-driven voltage prediction can be expressed as

$$\boldsymbol{v}^{\mathrm{PRE}} = \mathbb{M}(\boldsymbol{S}, \boldsymbol{C}) \tag{1}$$

where $\boldsymbol{v}^{\mathrm{PRE}}$ is the predicted node voltage vector; $\boldsymbol{S}$ is the measured operating data of ADS; $\mathbb{M}$ is the data-driven prediction methods; $\boldsymbol{C}$ is the hyperparameters of $\mathbb{M}$. Due to measurement data inaccuracies and inherent uncertainties in data-driven methods, the $\boldsymbol{v}^{\mathrm{PRE}}$ in (1) contains prediction errors. Accurately quantifying this uncertainty is essential for subsequent planning and operation, such as Volt-VAR management, voltage regulation, etc.

B. Proposed Physics-Informed GCNII

GCN can handle graph-structured data by applying graph convolution operation to all neighboring nodes of a given node via the K-th order polynomial of Laplacians [8]. This approximation relies on fixed coefficients, which limit the expressive power of the multi-layer GCN model and result in over-smoothing. To address that, GCNII is proposed by using initial residual connection and identity mapping [8]. Formally, we have

$$\begin{aligned}\boldsymbol{H}^{(\ell+1)} =& \sigma\left(\left(1-\alpha^{\ell}\right)\tilde{\boldsymbol{P}}\boldsymbol{H}^{(\ell)}\left(\left(1-\beta^{\ell}\right)\mathbf{I}_n+\beta^{\ell}\boldsymbol{W}_1^{(\ell)}\right)\right.\\ &\left.+\alpha^{\ell}\boldsymbol{H}^{(0)}\left(\left(1-\beta^{\ell}\right)\mathbf{I}_n+\beta^{\ell}\boldsymbol{W}_2^{(\ell)}\right)\right)\end{aligned} \tag{2}$$

where $\boldsymbol{H}^{(\ell)}$ is the ℓ-th layer of GCNII; σ denotes the ReLU activation function; $\boldsymbol{W}^{(\ell)}$ is the ℓ-th layer weight matrix; $\tilde{\boldsymbol{P}} = \tilde{\boldsymbol{D}}^{-1/2}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-1/2}$ is the graph convolution matrix with the renormalization trick; $\boldsymbol{A}$ is the adjacency matrix; $\boldsymbol{D}$ is the diagonal degree matrix; $\tilde{\boldsymbol{D}} = \boldsymbol{D}+\mathbf{I}_n$; $\tilde{\boldsymbol{A}} = \boldsymbol{A}+\mathbf{I}_n$; $\mathbf{I}_n$ is the identity matrix; α and β are predefined residual related hyperparameters. The first item in (2) represents the smoothed representation, while the second item represents the initial residual connection to the first layer $\boldsymbol{H}^{(0)}$. GCNII-based online voltage prediction can be represented as

$$\boldsymbol{v}^{\text{PRE}} = \text{GCNII}(\boldsymbol{x},\boldsymbol{A},\boldsymbol{W}) \tag{3}$$

where $\boldsymbol{x} = (\boldsymbol{p}^{\text{PV}},\boldsymbol{q}^{\text{PV}},\boldsymbol{p}^{\text{Battery}},\boldsymbol{p}^{\text{EV}},\boldsymbol{p}^{\text{Load}},\boldsymbol{q}^{\text{Load}})$ denotes the operating data vectors ($\boldsymbol{p}^{\text{PV}}$, $\boldsymbol{p}^{\text{Battery}}$, $\boldsymbol{p}^{\text{EV}}$ and $\boldsymbol{p}^{\text{LOAD}}$ are active power generation or load vectors of PV, Battery, EV or load, respectively; $\boldsymbol{q}^{\text{PV}}$ and $\boldsymbol{q}^{\text{LOAD}}$ are reactive power injections of PV or load, respectively). The loss function $\mathcal{L}$ is

$$\mathcal{L} = \sum_{i\in\mathcal{T}}\sum_{j\in\mathcal{N}}\lambda_j\left|v_{ij}^{\text{PRE}}-v_{ij}^{*}\right| \tag{4}$$

where * denotes the true value; $\mathcal{T}$ denotes the test set; $\mathcal{N}$ is the set of all nodes; λ_j is the weight of voltage deviation at j-th node, and usually set to 1 for all nodes. However, for large ADS, $\boldsymbol{x}$ is a sparse vector. If the same value of λ is used for all nodes, it may decrease the prediction accuracy for nodes with actual voltage problems, resulting in smoothed voltage predictions. To address this issue, the λ_j is redefined as

$$\lambda_j = N_{v,j}+1 \tag{5}$$

where $N_{v,j}$ denotes the number of voltage violations at node j in all samples.

C. Dropout-based Probabilistic Voltage Prediction

The Bayesian probability theory is typically used to analyze model uncertainty, but it has a high computational cost. Dropout can be incorporated into neural networks of any depth and with any non-linearities to approximate the Bayesian inference [9]. In this paper, dropout is applied in GCNII to model the voltage prediction uncertainty. Suppose M is the number of samples for dropout, the dropout-based probabilistic voltage prediction is shown in Fig. 1. It shows that M Monte

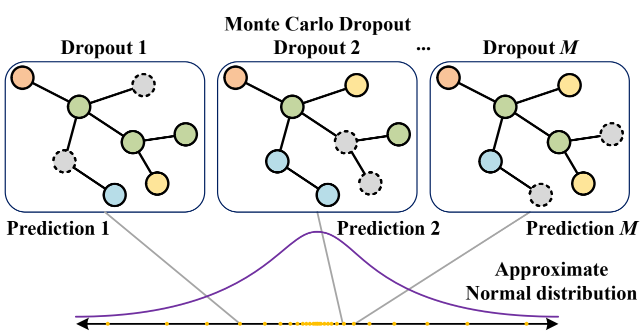


Fig. 1. Dropout-based probabilistic voltage prediction.

Carlo dropouts are performed to obtain M voltage predictions, which can be used to approximate the distribution of voltages. The expectation and variance of predicted voltage are

$$\mathbb{E}_{q(\boldsymbol{v}^*|\boldsymbol{x}^*)}\left(\boldsymbol{v}^*\right) \approx \frac{1}{M}\sum_{m=1}^{M}\boldsymbol{V}_m^{\text{PRE}} \tag{6}$$

$$\begin{aligned}\mathbb{V}^{\text{Cov}}_{q(\boldsymbol{v}^*|\boldsymbol{x}^*)}\left(\boldsymbol{v}^*\right) \approx& \frac{1}{M}\sum_{m=1}^{M}\left(\boldsymbol{V}_m^{\text{PRE},T}\boldsymbol{V}_m^{\text{PRE}}\right.\\ &\left.-\mathbb{E}_{q(\boldsymbol{v}^*|\boldsymbol{x}^*)}\left(\boldsymbol{v}^*\right)^T\mathbb{E}_{q(\boldsymbol{v}^*|\boldsymbol{x}^*)}\left(\boldsymbol{v}^*\right)\right)\end{aligned} \tag{7}$$

where $\mathbb{E}_{q(\boldsymbol{v}^*|\boldsymbol{x}^*)}\left(\boldsymbol{v}^*\right)$ and $\mathbb{V}^{\text{Cov}}_{q(\boldsymbol{v}^*|\boldsymbol{x}^*)}\left(\boldsymbol{v}^*\right)$ are expectation vector and covariance variance matrix, respectively; $\boldsymbol{V}^{\text{PRE}} = \{\boldsymbol{v}_m^{\text{PRE}}\}_{m=1}^{M}$ denotes the set of M predicted voltage vectors; superscript T denotes the transpose of a vector. The Bernoulli distribution is used to generate the binary vector $\boldsymbol{z}_m$, which is then multiplied with weight matrix $\boldsymbol{W}$ to achieve dropout, i.e., $\boldsymbol{v}_m^{\text{PRE}} = \text{GCNII}(\boldsymbol{x},\boldsymbol{A},\boldsymbol{z}_m\boldsymbol{W})$. The confidence probability $\mathbb{P}$ of voltage predictions can be calculated by the integral of the probability density function of the Normal distribution.

$$\mathbb{P} = \frac{1}{\sqrt{2\pi}\mathbb{S}_j}\int_{\mathbb{E}_j-\gamma\mathbb{S}_j}^{\mathbb{E}_j+\gamma\mathbb{S}_j}\exp\left(-\frac{\left(v_j^{\text{PRE}}-\mathbb{E}_j\right)^2}{2\mathbb{S}_j^2}\right)dv_j^{\text{PRE}}, j\in\mathcal{N} \tag{8}$$

where $\mathbb{E}$ denotes $\mathbb{E}_{q(\boldsymbol{y}^*|\boldsymbol{x}^*)}(\boldsymbol{V}^*)$; $\mathbb{V}$ denotes $\mathbb{V}_{q(\boldsymbol{y}^*|\boldsymbol{x}^*)}(\boldsymbol{V}^*)$; $\mathbb{S}_{q(\boldsymbol{y}^*|\boldsymbol{x}^*)}(\boldsymbol{v}^*) = \sqrt{\mathbb{V}_{q(\boldsymbol{y}^*|\boldsymbol{x}^*)}(\boldsymbol{v}^*)}$ denotes the standard deviation; When the confidence probability for voltage is chosen as 80%, 90%, 95% and 99%, the corresponding γ are 1.282, 1.645, 1.960 and 2.575, respectively.

III. Numerical Results

The proposed method is tested on a real unbalanced three-phase 759-node ADS, which has 120 PVs, 99 batteries, 71 EVs, 163 loads and 20 switches. We assume that only 50% of the PVs and load data are known, specifically, 50% of $\boldsymbol{p}^{\text{PV}}$, $\boldsymbol{p}^{\text{Load}}$, and $\boldsymbol{q}^{\text{Load}}$ are zeros. The active power outputs of PVs and loads vary between 0% and 100%. The EV is either in an on or off state. The battery has three states: charging, discharging, and idle, and the active power output ranges from the maximum charging power to the maximum discharging power. Seven different feasible topologies are generated by adjusting different switches. Each topology corresponds to 10,000 random operating scenarios. A total of 70,000 samples were generated by using OpenDSS, of which 80% were used as the training set and 20% were used as the test set. The one-line diagram of the 759-node ADS is shown in Fig. 2.

A. GCNII Training and Performance Evaluation

GCNII is built on the PyTorch Geometric (PyG) library. According to (4), the comparisons of voltage prediction results of different λ settings are shown in Fig. 3. It can be seen from Fig. 3 that when λ is set to 1, the predicted voltages tend to be smooth, i.e., distributed between 1.00 and 1.05. When λ is set according to (5), the predicted voltage v^{pre} is very close to v^*.

Apart from GCNII, we train several other models to compare their performances, including GCN, graph attention network (GAT), artificial neural networks (ANN), deep belief network (DBN), and deep Gaussian process (DGP). Several

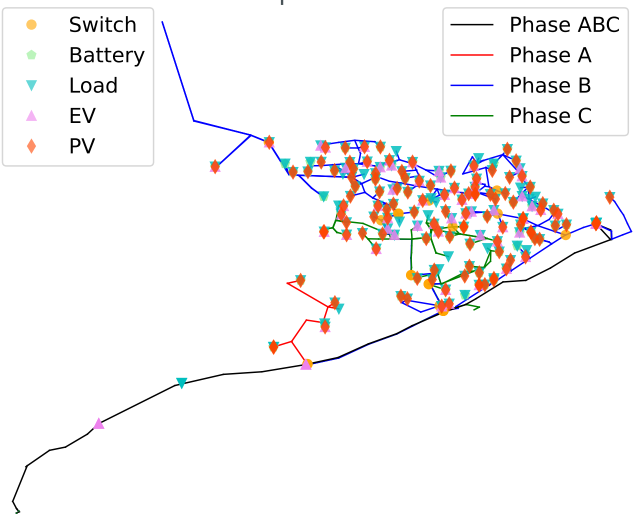
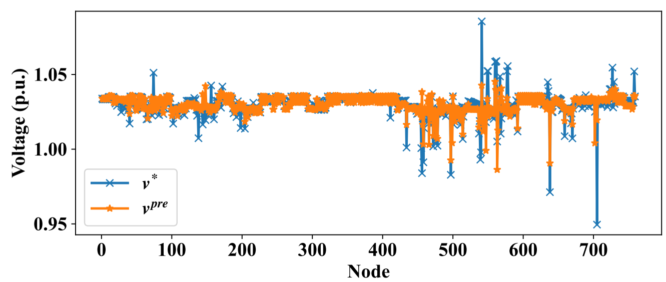
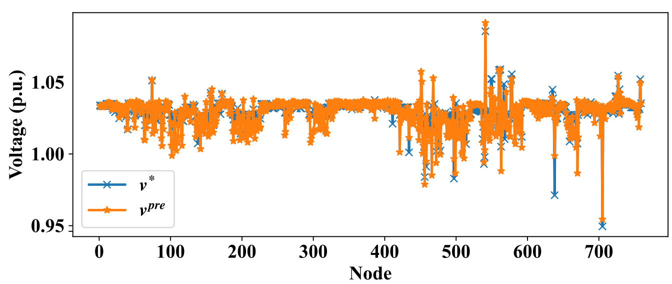


Fig. 2. One-line diagram of the real 759-node ADS.



(a)



(b)

Fig. 3. Comparison of voltage prediction results of different λ settings. (a): Voltage prediction for nodes in $\mathcal{N}$ ($\lambda = 1$). (b): Voltage prediction for nodes in $\mathcal{N}$ (λ is set according to (5)).

metrics are utilized, including the miss detection rate (MD) and false alarm rate (FA) [10]. Furthermore, to specifically evaluate the performance on nodes with voltage violation, additional metrics called MD$_V$ and FA$_V$ are computed. Typically, the importance ranking of these accuracy metrics is MD $>$ FA. This is because MD may cause voltage safety risks, which are critical, while FA increases control frequency. The comparison results are shown in Table I. Table I demonstrates that the GCNII outperforms other deep learning models in terms of prediction accuracy, especially in MD, which is the most critical metric related to voltage security risks. It is worth noting that DGP incurs a significantly higher computational burden and time requirement compared to GCNII.

TABLE I
COMPARISON OF PREDICTION ACCURACY WITH OTHER MODELS

Model	Layer	HC	MD (%)	FA (%)	MD$_V$ (%)	FA$_V$ (%)
GCN	8	300	0.5625	0.4607	3.746	1.936
GAT	7	300	0.7771	0.1084	4.828	0.8191
GCNII*	**8**	**300**	**0.2824**	**0.1497**	**1.880**	**0.9968**
ANN	8	300	0.5661	0.1066	3.768	0.5940
DBN	8	300	0.4988	0.1542	3.033	1.026
DGP	6	300	2.639	0.2707	9.861	1.182

*: Proposed method; HC: Hidden channel column only shows the best hidden channel for each number of layers; Head is a hyperparameter of GAT, which refers to the number of attention heads, and is set to 5; The hyperparameters of all models have been selected to be optimal.

Additionally, DGP faces challenges in training when dealing with multiple outputs.

B. Dropout-based Voltage Chance Constraint

According to (8), four confidence levels are set, including 80%, 90%, 95%, and 99%. N_d is set to 300, and the corresponding confidence intervals for nodes with historical voltage violation records are presented in Fig. 4 [9]. From Fig. 4, it

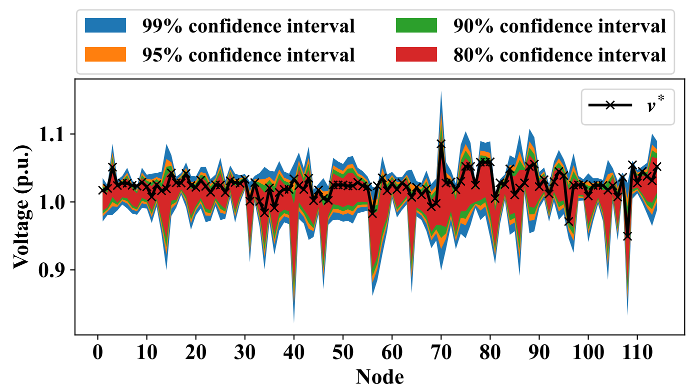


Fig. 4. Confidence interval for different probabilities.

can be observed that as the confidence probability increases, the corresponding confidence intervals become larger, and the probability of v^* falling within the confidence interval also increases. The probabilistic voltage prediction time is 5.3s, which satisfies the voltage prediction requirement (either fast time-scale, such as 1 min or longer time-scale) and can provide valuable information for further voltage control purposes.

C. Performance under Topology Changes

The distribution system operating points are always changing and induce different patterns of voltage variations. One particular challenge is to deal with topology change while maintaining good voltage predictions. This requires high robustness of the proposed method. Thanks to the incorporation of physical information regarding topologies in our proposed GCNII method, it naturally gains robustness to topology changes. Fig. 5 presents a comparison of prediction results between GCNII and DBN in the presence of different topology changes. From Fig. 5, we can see that GCNII can adapt

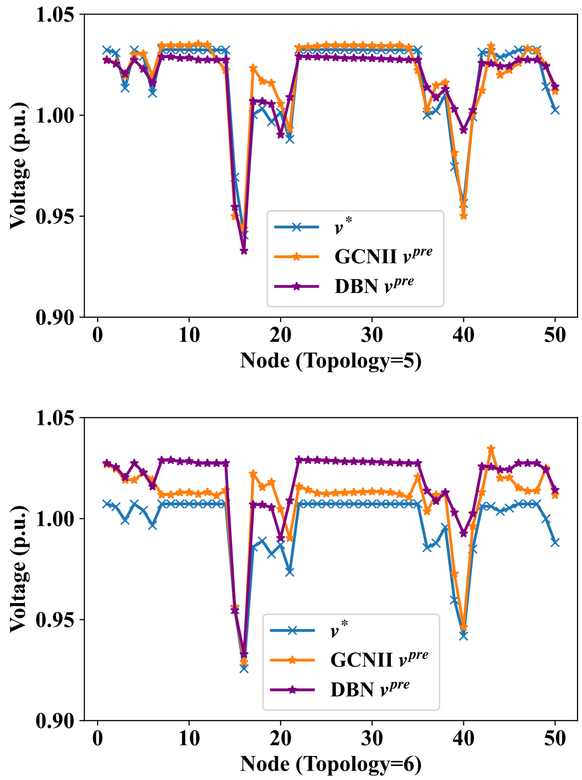


Fig. 5. Comparison of voltage predictions for nodes around changed switches that induce topology changes.

to topology changes, while DBN, which only incorporates a flag indicating different topologies can not deal with this. For example, there is in fact low voltage issue around node 40 under topology change. The proposed GCNII is still able to capture this and inform the system operator for taking corrective actions to deal with that. By contrast, DBN fails to predict this and thus can yield voltage security issues for the distribution system.

IV. Conclusion

This paper proposes a novel probabilistic physics-informed GCNII-based voltage prediction method for ADS. The key innovation of this paper is the integration of GCNII and dropout techniques to approximate the distribution of node voltages and evaluate their uncertainty, providing valuable information for further voltage control. Numerical results on a real 759-node ADS demonstrate that the proposed method outperforms other state-of-the-art machine learning methods in accuracy and robustness across various scenarios.

References

[1] A. F. Bastos, S. Santoso, V. Krishnan, and Y. Zhang, "Machine learning-based prediction of distribution network voltage and sensor allocation," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2020, pp. 1–5.

[2] M. Marković, A. Sajadi, A. Florita, R. Cruickshank III, and B.-M. Hodge, "Voltage estimation in low-voltage distribution grids with distributed energy resources," *IEEE Trans. Sustain. Energy*, vol. 12, no. 3, pp. 1640–1650, Jul. 2021.

[3] Y. Seyedi, H. Karimi, and J. Mahseredjian, "A data-driven method for prediction of post-fault voltage stability in hybrid ac/dc microgrids," *IEEE Trans. Power Syst.*, vol. 37, no. 5, pp. 3758–3768, Sep. 2022.

[4] T. Wu, I. L. Carreño, A. Scaglione, and D. Arnold, "Spatio-temporal graph convolutional neural networks for physics-aware grid learning algorithms," *IEEE Trans. Smart Grid*, 2023.

[5] H.-Y. Su and H.-H. Hong, "An intelligent data-driven learning approach to enhance online probabilistic voltage stability margin prediction," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3790–3793, Jul. 2021.

[6] J.-F. Toubeau and et. al, "Privacy-preserving probabilistic voltage forecasting in local energy communities," *IEEE Trans. Smart Grid*, vol. 14, no. 1, pp. 798–809, Jan. 2023.

[7] P. Pareek, W. Yu, and H. D. Nguyen, "Optimal steady-state voltage control using gaussian process learning," *IEEE Trans. Ind. Informat.*, vol. 17, no. 10, pp. 7017–7027, Oct. 2020.

[8] M. Chen, Z. Wei, Z. Huang, B. Ding, and Y. Li, "Simple and deep graph convolutional networks," in *Proc. ICML*, 2020, pp. 1725–1735.

[9] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *Proc. ICML*, 2016, pp. 1050–1059.

[10] T. Su, J. Zhao, and X. Chen, "Deep sigma point processes-assisted chance-constrained power system transient stability preventive control," *IEEE Trans. Power Syst.*, 2023.