

Physics-informed Graph Neural Networks for Collaborative Dynamic Reconfiguration and Voltage Regulation in Unbalanced Distribution Systems

Jingtao Qin, *Student Member, IEEE*, Rui Yang, *Member, IEEE*, and Nanpeng Yu, *Senior Member, IEEE*

Abstract—Network reconfiguration has long been employed as a strategic approach to minimize power distribution system losses and effectively regulate voltage levels. Tap-changing voltage regulators are also critical for controlling bus voltages, especially in accommodating the increasing integration of distributed energy resources (DERs) with intermittent outputs. This paper introduces novel methodologies to address the challenges of dynamic reconfiguration and optimal tap setting in unbalanced three-phase distribution systems. We propose an approximated mixed-integer quadratically constrained program (MIQCP) to model dynamic reconfiguration, along with a pioneering formulation for voltage regulator (VR) tap-setting based on Special Ordered Set type 1 (SOS1). To mitigate computational complexity, we propose a physics-informed spatial-temporal graph convolutional network (STGCN) with an integrated link classifier. The proposed approach enables efficient solution generation by fixing specific variables in the MIQCP instance and solving the simplified sub-MIP using an MIP solver. Numerical studies demonstrate the superior prediction accuracy of our STGCN model compared to baseline neural network models, resulting in reduced DER curtailment and voltage deviation with shorter computation time.

Index Terms—Physics-informed networks, dynamic reconfiguration, voltage regulation, unbalanced distribution systems.

NOMENCLATURE

A. Sets, Constants, and Coefficients

T	Set of operation periods.
\mathcal{G}, \mathcal{S}	Set of DERs and substations.
$\mathcal{N}, \mathcal{N}_3$	Set of all buses and three-phase buses.
$\mathbb{N}(i)$	Set of neighbor buses of bus i .
$\mathcal{E}, \mathcal{E}^s, \mathcal{E}^r$	Set of all lines, switches and voltage regulators.
Φ_i	Set of phases of bus i .
Φ_{ij}	Set of phases of line (i, j) .
N, N_s	Number of buses and substations.
i_s	Virtual bus used for the fomulation of voltage regulator (i, j) .
c_g	Cost coefficient of DER curtailment, $k\$/\text{KWh}$.
c_d	Cost coefficient of voltage deviation, $k\$/(\text{KV})^2$.

B. Decision Variables and Parameters

$Pg_{i,t}^\phi, Qg_{i,t}^\phi$	Active and reactive output of DER i at time t on phase ϕ .
$\overline{Pg}_{i,t}^\phi, Sg_{i,t}^\phi$	Maximum active power output and capacity of DER i at time t on phase ϕ .

$Ps_{n,t}^\phi, Qs_{n,t}^\phi$	Active and reactive power of substation i at time t on phase ϕ .
$Pd_{i,t}^\phi, Qd_{i,t}^\phi$	Active and reactive load of bus i at time t on phase ϕ .
$P_{ij,t}^\phi, Q_{ij,t}^\phi$	Active and reactive power flow from bus i to bus j at time t on phase ϕ .
$\mathbf{P}_{ij,t}, \mathbf{Q}_{ij,t}$	Active and reactive power flow vector from bus i to bus j at time t on phase ϕ , $\mathbb{R}^{3 \times 1}$.
$\mathbf{r}_{ij}, \mathbf{x}_{ij}$	Resistance and reactance matrix of line (i, j) , $\mathbb{R}^{3 \times 3}$.
$\tilde{\mathbf{r}}_{ij}, \tilde{\mathbf{x}}_{ij}$	Transformed resistance and reactance matrix of line (i, j) , $\mathbb{R}^{3 \times 3}$.
\bar{S}_{ij}	Power flow capacity of line (i, j) .
$U_{i,t}^\phi$	Squared voltage magnitude of bus i at time t on phase ϕ .
$Ud_{i,t}^\phi$	Deviation in squared voltage magnitude of bus i from the reference value at time t on phase ϕ .
$\bar{V}_i, \underline{V}_i$	Upper and lower limit of voltage magnitude of bus i .
$\bar{V}_i^r, \underline{V}_i^r$	Upper and lower voltage magnitude reference of bus i .
$\hat{U}_{i,t}$	Average voltage magnitude of three phases of bus i at time t .
δ	Maximum voltage imbalance limit for three-phase buses.
$\alpha_{ij,t}$	Binary variable that is set to 1 if switch (i, j) is closed at time t , and 0 otherwise.
$\beta_{ij,t}$	Binary variable that is set to 1 if bus j is the parent of bus i at time t , and 0 otherwise.
$\gamma_{ij,t}$	Auxiliary variable that represents the operation for switch (i, j) at time t , 1 denotes closing the switch, and 0 otherwise.
$\tau_{ij,t}$	Integer variable that indicates the tap position of voltage regulator (i, j) at time t .
$z_{ij,t}$	Auxiliary binary variable for tap position of voltage regulator (i, j) at time t .
$\lambda_{ij,t}$	Tap position variation from last period of voltage regulator (i, j) at time t .
$\gamma_{\max}, \lambda_{\max}$	Maximum number of switching/tap-changing actions over all period T .

C. Operations

\odot	Element-wise Hadamard product.
$(\cdot)^H$	Conjugate transpose.

The authors gratefully acknowledge the financial support provided by the University of California Office of the President, under award number L22CR4556.

$\sigma(\cdot)$ Sigmoid function.
 $\lfloor x \rfloor$ Round down the number x to the nearest integer.

I. INTRODUCTION

DYNAMIC network reconfiguration (DNR) has been an effective method in minimizing power distribution system losses, regulating bus voltages, and improving system resilience [1]–[3]. The value of dynamic reconfiguration in the presence of distributed energy resources (DERs) has been studied in [4]. The findings reveal that the implementation of DNR yields a proportional decrease in the system's active power loss with the increasing penetration of DERs, such as wind and solar photovoltaic (PV). Voltage regulation in unbalanced distribution systems has become increasingly complex and crucial due to the higher penetration of DERs. Several other investigations [5], [6] show that the voltage volatility is strongly influenced by the topology of power distribution systems.

DNR is accomplished by adjusting the status of tie switches and sectionalizers to change the topology of the system. This problem suffers from the combinatorial explosion as the number of binary decision variables that denote the open/close status of tie switches or sectionalizers increases significantly [7]. Moreover, the complexity of this problem is further amplified by the nonlinearity of the power flow constraints. Thus, it is challenging to improve both computational efficiency and solution quality for the DNR problem in power distribution networks [8]. DNR problems can be solved by heuristic algorithms such as the branch exchange method [9], [10], genetic algorithm [11], [12], and particle swarm optimization [13], [14]. Optimization techniques have also been used to solve the DNR problem that considers the intermittent DERs. Wang et al. [15] introduced an approximate dynamic programming (ADP) approach aimed at minimizing curtailment of DERs and load shedding. Zhan et al. [16] proposed a switch opening and exchange (SOE) method to effectively manage variable and uncertain load and solar PV output. Additionally, a distributionally robust model for three-phase unbalanced DNR is presented in [17], which seeks to determine the optimal configuration under the worst-case probability distribution of DER outputs and loads within an ambiguity set. In [18], a dynamic reconfiguration approach for a three-phase unbalanced distribution network is proposed to minimize daily power loss, following the method outlined by B.A. Robbins in [19], which neglects higher-order power flow constraints.

Tap-changing voltage regulators (VRs) are frequently employed in distribution systems to control voltage levels and minimize power losses. In addition to control-based strategies [20]–[22], optimization-based methods for VRs are widely studied and employed. In [23], the tap positions are modeled as discrete variables within the optimal power flow (OPF) problem. This formulation leads to a mixed-integer program (MIP) formulation. The complexity of the problem also grows exponentially with the number of VRs. To address this challenge, various initial studies [24]–[26] have proposed representing transformer tap positions as continuous variables.

Subsequently, the solutions are rounded to the nearest discrete variables. In [26], the optimal configuration of tap positions in unbalanced three-phase systems is formulated as a rank-constrained semidefinite program (SDP). By relaxing the non-convex rank-1 constraint, a convex SDP problem is obtained and solved using the alternating direction method of multipliers (ADMM). In the study by Savasci et al. [27], an automated tap-changing process for VRs is proposed, utilizing a bang-bang type control rule parameterized by a dead band parameter. Accounting for the unbalanced nature of multi-phase feeders, an optimization-based dead band tuning is proposed and formulated as a mixed-integer linear program (MILP) aimed at minimizing both the total number of tap switching and the curtailment of DERs. However, there is very little work exploring the joint voltage regulation and network reconfiguration problem in three-phase unbalanced distribution systems. In [8], the optimal network reconfiguration of unbalanced distribution systems with renewable DERs and VRs is explored. The co-optimization problem is framed as a mixed-integer semidefinite programming (MISDP) with chordal relaxation. However, it only focuses on a single time period, and its computational demands become overwhelming when addressing distribution systems of substantial size.

The recent advances in deep learning (DL) have enabled deep neural networks (DNN)-based applications in power systems. There is a growing interest in the power community for the application of physics-informed neural networks (PINN) [28], which has found applications in state/parameter estimation, system dynamic analysis [29]–[31] and optimal power flow. Lately, there has been a surge in research exploring the application of PINN to address network reconfiguration challenges in distribution systems. J. Authier et al. [32] proposed an end-to-end physics-informed graph neural network framework for dynamic reconfiguration problems where the operational and connectivity constraints are directly incorporated into the learning framework. In [33], a leaning-assisted physics-informed graph convolution network is introduced to predict the connectivity of tie lines. Nevertheless, these two papers concentrated solely on balanced distribution systems and did not take VRs into account.

To fill the knowledge gap mentioned earlier, we first introduce an approximated mixed-integer quadratically constrained program (MIQCP) to solve the DNR problem in unbalanced three-phase distribution systems. Additionally, we introduce a novel formulation for VR tap-setting based on Special Ordered Set type 1 (SOS1). Subsequently, in an effort to mitigate computational complexity, we propose a physics-informed spatial-temporal graph convolution network (STGCN) with an integrated link classifier, where the physical structure of the distribution system is utilized to construct the topology of the spatial-temporal graph. This model is designed to predict both the connectivity of tie switches/sectionalizers and the tap position of VRs. As illustrated in Fig. 1, the workflow begins with utilizing topology data, electric load, and DER outputs to construct spatial-temporal graphs. These graphs are subsequently fed into a physics-informed STGCN and a link classifier to produce predictions for tie switches/sectionalizers and tap positions of VRs. Leveraging these predictions, we can

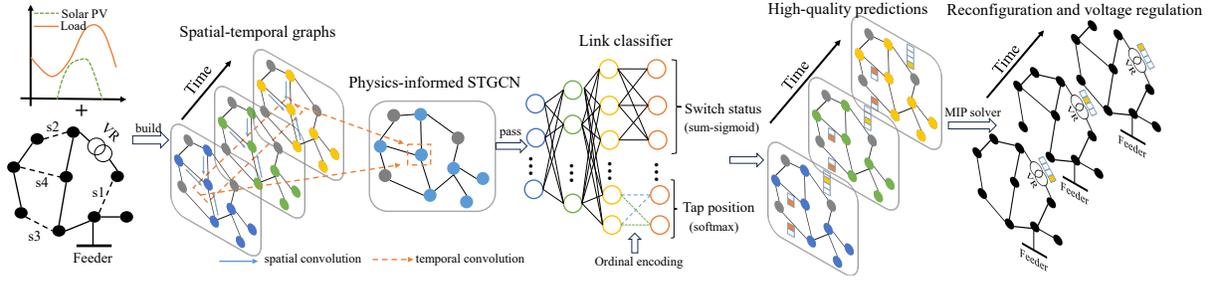


Fig. 1. The complete workflow of physics-informed STGCN for dynamic reconfiguration and voltage regulation.

then fix specific decision variables within the original MIQCP and solve the simplified sub-MIP using an MIP solver. Numerical studies demonstrate that our proposed STGCN model outperforms the baseline neural network models in prediction accuracy. Leveraging the predictions generated by STGCN enables us to obtain superior solutions with reduced DER curtailment and voltage deviation with shorter computation time.

The main contributions of this paper are:

- 1) We construct an MIQCP model for collaborative dynamic network reconfiguration and voltage control with VRs. In the proposed approach, we introduce a novel formulation for tap positions using SOS1 without relaxing the integrality constraint.
- 2) We propose a spatial-temporal graph convolutional network to extract spatial and temporal information from unbalanced distribution systems. The output of this model is then fed into a novel link classifier to deliver accurate predictions for the connectivity of tie switches, sectionalizers, and the optimal tap positions of VRs.
- 3) We develop an algorithm to fix binary decision variables by using the probabilistic predictions for the status of tie switches/sectionalizers and tap positions of VRs. By solving the simplified sub-MIP problem, the proposed algorithm achieves superior network reconfiguration and voltage regulation solutions than both the baseline machine learning models and the commercial MIP solver.

The subsequent sections of this paper are structured as follows: Section II presents the MIQCP formulation for co-optimizing dynamic reconfiguration and voltage regulation. In Section III, we introduce the STGCN model along with the algorithm for sub-MIP solving. Section IV presents the numerical studies. Section V provides the conclusions.

II. RECONFIGURATION AND VOLTAGE REGULATION FORMULATION FOR THREE-PHASE DISTRIBUTION SYSTEM

In this section, we present the mathematical formulation of joint reconfiguration and voltage regulation in unbalanced three-phase distribution networks. The problem formulation is derived based on the linear approximation technique for unbalanced three-phase distribution systems [19]. The objective function is introduced in Subsection A. The constraints are presented in Subsection B to Subsection E.

A. Objective Function

The objective of the collaborative network reconfiguration and voltage regulation is two-fold as shown in (1)-(2). First, it aims to minimize the curtailment of DER generation for both economic and environmental benefits. Second, it tries to reduce voltage deviation beyond the normally accepted bounds to prevent unintentional damage to the electric equipment of end-use customers.

$$\min_{Pg_{i,t}^\phi, Qg_{i,t}^\phi, U_{i,t}^\phi, \alpha_{ij,t}, \beta_{ij,t}, \tau_{ij,t}, z_{ij,t}} \sum_{t \in T} \mathcal{F}_t \quad (1)$$

$$\mathcal{F}_t = \sum_{i \in \mathcal{G}} \sum_{\phi \in \Phi_i} c_g (\overline{Pg}_{i,t}^\phi - Pg_{i,t}^\phi) + \sum_{i \in \mathcal{N}} \sum_{\phi \in \Phi_i} c_d U d_{i,t}^\phi \quad (2)$$

Subject to constraints (3)-(8), (12)-(17), (20)-(35).

B. Power Flow Constraints

By assuming the voltages of each bus are nearly balanced and neglecting the higher-order terms [19], we obtain the active and reactive power balance equations (3)-(4) for bus $i \in \mathcal{N}/\mathcal{S}$, and equations (5)-(6) for bus $i \in \mathcal{S}$. Then, $\forall \phi \in \Phi_i, \forall t \in T$, we have:

$$Pg_{n,t}^\phi - Pd_{i,t}^\phi = \sum_{n:n=i,n \in \mathcal{G}} \sum_{(i,j) \in \mathcal{E}} P_{ij,t}^\phi - \sum_{(k,i) \in \mathcal{E}} P_{ki,t}^\phi \quad (3)$$

$$Qg_{n,t}^\phi - Qd_{i,t}^\phi = \sum_{n:n=i,n \in \mathcal{G}} \sum_{(i,j) \in \mathcal{E}} Q_{ij,t}^\phi - \sum_{(k,i) \in \mathcal{E}} Q_{ki,t}^\phi \quad (4)$$

$$Ps_{n,t}^\phi - Pd_{i,t}^\phi = \sum_{m:m=i,n \in \mathcal{S}} \sum_{(i,j) \in \mathcal{E}} P_{ij,t}^\phi - \sum_{(k,i) \in \mathcal{E}} P_{ki,t}^\phi \quad (5)$$

$$Qs_{n,t}^\phi - Qd_{i,t}^\phi = \sum_{m:m=i,n \in \mathcal{S}} \sum_{(i,j) \in \mathcal{E}} Q_{ij,t}^\phi - \sum_{(k,i) \in \mathcal{E}} Q_{ki,t}^\phi \quad (6)$$

Similarly, for each line $(i, j) \in \mathcal{E}$, by neglecting the higher-order terms and adding/subtracting a large positive number [4], we obtain the voltage magnitude constraint as shown in (7)-(8). Then, $\forall \phi \in \Phi_{ij}, \forall t \in T$, we have:

$$U_{j,t}^\phi \leq U_{i,t}^\phi - 2(\tilde{r}_{ij}^\phi P_{ij,t}^\phi + \tilde{x}_{ij}^\phi Q_{ij,t}^\phi) + M(2 - \alpha_{ij,t} - e_{i,j}^\phi), \quad (7)$$

$$U_{j,t}^\phi \geq U_{i,t}^\phi - 2(\tilde{r}_{ij}^\phi P_{ij,t}^\phi + \tilde{x}_{ij}^\phi Q_{ij,t}^\phi) - M(2 - \alpha_{ij,t} - e_{i,j}^\phi), \quad (8)$$

where M is a large positive constant. $e_{i,j}^\phi$ is the composition indicator which equals 1 if line (i, j) has phase ϕ and 0 otherwise. $\tilde{r}_{ij}, \tilde{x}_{ij}$ are the transformed resistance and reactance of line (i, j) which are calculated by:

$$\tilde{r}_{ij} = \text{Re}(\mathbf{a}\mathbf{a}^H) \odot \mathbf{r}_{ij} + \text{Im}(\mathbf{a}\mathbf{a}^H) \odot \mathbf{x}_{ij} \quad (9)$$

$$\tilde{x}_{ij} = \text{Re}(\mathbf{a}\mathbf{a}^H) \odot \mathbf{x}_{ij} - \text{Im}(\mathbf{a}\mathbf{a}^H) \odot \mathbf{r}_{ij} \quad (10)$$

$$\mathbf{a} = [1 \ e^{-j2\pi/3} \ e^{j2\pi/3}]^T, \quad (11)$$

where \odot is the element-wise Hadamard product. \mathbf{a}^H is the conjugate transpose of \mathbf{a} .

C. Voltage Constraints

In this subsection, we formulate the voltage constraints. In addition to voltage magnitude constraints (12)-(13), voltage imbalance is also a major concern in the distribution system. Here we adopt the formulation in [18] to build the phase imbalance constraints (14)-(15). Finally, the deviation in squared voltage magnitude from the reference value is computed utilizing (16) and (17).

$$U_{i,t}^\phi = 1, \ i \in \mathcal{S}, \phi \in \Phi_i, t \in T \quad (12)$$

$$\underline{V}_i^2 \leq U_{i,t}^\phi \leq \overline{V}_i^2, \ i \in \mathcal{N}/\mathcal{S}, \phi \in \Phi_i, t \in T \quad (13)$$

$$-\delta \leq \frac{U_{i,t}^\phi - \hat{U}_{i,t}}{\hat{U}_{i,t}} \leq \delta, \ i \in \mathcal{N}_3, \phi \in \Phi_i, t \in T \quad (14)$$

$$\hat{U}_{i,t} = \frac{1}{3} \sum_{\phi \in \Phi_i} U_{i,t}^\phi, \ i \in \mathcal{N}_3, t \in T \quad (15)$$

$$Ud_{i,t}^\phi \geq U_{i,t}^\phi - (\underline{V}_i^r)^2 \quad (16)$$

$$Ud_{i,t}^\phi \geq -U_{i,t}^\phi + (\overline{V}_i^r)^2 \quad (17)$$

D. Voltage Regulator Formulation

To obtain a linear formulation of voltage regulators, we start from the equivalent voltage model developed in [26]. Assume there is a three-phase, wye-wye solidly grounded voltage regulator between bus i and bus j . It can be equivalently modeled as an ideal transformer in series with an impedance as shown in Fig. 2, where bus i_s is a virtual node between bus i and bus j that is connected to the secondary side of the ideal transformer.

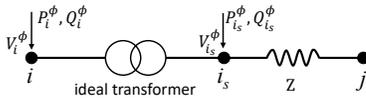


Fig. 2. Equivalent voltage regulator model.

The ideal transformer can be formulated as (18) for each phase, where η_{ij} is the tap ratio.

$$V_{i_s}^\phi = \eta_{ij} V_i^\phi, \ P_i^\phi = -P_{i_s}^\phi, \ Q_i^\phi = -Q_{i_s}^\phi \quad (18)$$

Then, assume τ_{ij} is an integer variable that represents the tap position and $\eta_{ij} = \eta_{ij} + \tau_{ij} \Delta \eta_{ij}$, $0 \leq \tau_{ij} < M_{ij}$, where M_{ij} is the number of tap positions for voltage regulator between bus i and j , $\underline{\eta}_{ij}$ is the lower bound of the tap ratio, and

$\Delta \eta_{ij} > 0$ is the change in tap ratio for each tap adjustment. Then we have:

$$U_{i_s,t}^\phi = \eta_{ij}^2 U_{i,t}^\phi = (\underline{\eta}_{ij} + \tau_{ij} \Delta \eta_{ij})^2 U_{i,t}^\phi \quad (19)$$

Since (19) has a nonlinear term that contains the product of a continuous variable and a squared integer variable, it cannot be handled by most of the MIP solvers. So we apply Special Ordered Set type 1 (SOS1) to tackle this problem. Formally, for a set of binary variables $\{z_1, z_2, \dots, z_{M_{ij}}\}$ with associated non-negative weights $\{w_1, w_2, \dots, w_{M_{ij}}\}$, we give the linear formulation of voltage regulator as:

$$P_i^\phi = -P_{i_s}^\phi, \ Q_i^\phi = -Q_{i_s}^\phi \quad (20)$$

$$U_{i_s}^\phi = U_i^\phi \cdot \sum_{m=1}^{M_{ij}} z_m \cdot w_m \quad (21)$$

$$\sum_{m=1}^{M_{ij}} z_m \cdot w_m = \eta_{ij}^2, \ \sum_{m=1}^{M_{ij}} m \cdot z_m = \tau_{ij}, \ \sum_{m=1}^{M_{ij}} z_m = 1 \quad (22)$$

E. The Other Constraints

The other constraints considered here include the network radiality constraints, DER output constraints, branch capacity constraints and switch operation constraints.

1) *Network Radiality Constraints*: We employ the following equations to enforce the radiality and connectivity of the system, ensuring that all buses, except for the substation buses, have precisely one parent.

$$\sum_{(i,j) \in \mathcal{E}} \alpha_{ij,t} = N - N_s, \ t \in T \quad (23)$$

$$\beta_{ij,t} + \beta_{ji,t} = \alpha_{ij,t}, \ (i,j) \in \mathcal{E}, \ t \in T \quad (24)$$

$$\sum_{j \in \mathbb{N}(i)} \beta_{ij,t} = 1, \ i \in \mathcal{N}/\mathcal{S}, \ t \in T \quad (25)$$

$$\beta_{ij,t} = 0, \ i \in \mathcal{S}, \ t \in T \quad (26)$$

2) DER Output Constraints:

$$0 \leq Pg_{i,t}^\phi \leq \overline{Pg}_{i,t}^\phi, \ i \in \mathcal{G}, \phi \in \Phi_i, t \in T \quad (27)$$

$$(Pg_{i,t}^\phi)^2 + (Qg_{i,t}^\phi)^2 \leq (Sg_{i,t}^\phi)^2, \ i \in \mathcal{G}, \phi \in \Phi_i, t \in T \quad (28)$$

3) Branch Capacity Constraints:

$$(P_{ij,t}^\phi)^2 + (Q_{ij,t}^\phi)^2 \leq \alpha_{ij,t} (\overline{S}_{ij})^2, \ (i,j) \in \mathcal{E}, \phi \in \Phi_{ij}, t \in T \quad (29)$$

4) Switch and Voltage Regulator Operation Constraints:

To mitigate excessive operations and prolong the lifespan of switches and voltage regulators, we introduce maximum operation time constraints for each switch and voltage regulator as follows:

$$\gamma_{ij,t} \geq \alpha_{ij,t} - \alpha_{ij,t-1}, \ (i,j) \in \mathcal{E}^s, t \in T \quad (30)$$

$$\gamma_{ij,t} \geq \alpha_{ij,t-1} - \alpha_{ij,t}, \ (i,j) \in \mathcal{E}^s, t \in T \quad (31)$$

$$\sum_{t \in T} \gamma_{ij,t} \leq \gamma_{\max}, \ \forall (i,j) \in \mathcal{E}^s \quad (32)$$

$$\lambda_{ij,t} \geq \tau_{ij,t} - \tau_{ij,t-1}, \ (i,j) \in \mathcal{E}^r, t \in T \quad (33)$$

$$\lambda_{ij,t} \geq \tau_{ij,t-1} - \tau_{ij,t}, \ (i,j) \in \mathcal{E}^r, t \in T \quad (34)$$

$$\sum_{t \in T} \lambda_{ij,t} \leq \lambda_{\max}, \ \forall (i,j) \in \mathcal{E}^r \quad (35)$$

III. PHYSICS-INFORMED GRAPH CONVOLUTIONAL MODEL FOR RECONFIGURATION AND VOLTAGE REGULATION

In this section, we introduce the proposed physics-informed graph convolutional model for collaborative distribution system reconfiguration and voltage regulation. The overall framework of the proposed model is shown in Fig. 1. It begins by feeding the constructed physics-informed graphs into a spatial-temporal graph convolutional network (STGCN). Subsequently, the resulting graph is passed to a link classifier, where node embeddings are leveraged to generate edge embeddings. Finally, these edge embeddings are employed to forecast the status of switches and the tap position of voltage regulators. By fixing some decision variables predicted by the proposed physics-informed graph convolutional model, we can simplify the original optimization problem as sub-MIP problems which are much easier to solve.

A. Graph Representation and Feature Definition

In this subsection, we represent the unbalanced distribution system as a graph, denoted by $\mathbb{G} = (\mathcal{V}, \mathbf{E}, \mathcal{A})$. Here \mathcal{V} is the set of nodes that corresponds to the buses in the distribution system, $|\mathcal{V}| = N$. Similarly, \mathbf{E} is the set of edges that corresponds to the lines within the distribution system, $|\mathbf{E}| = |\mathcal{E}|$. $\mathcal{A} \in \mathbb{R}^{n \times n}$ is the adjacency matrix that indicates the connectivity of line segments, \mathcal{A}_{ij} equals 1 if line (i, j) are connected and 0 otherwise.

In order to capture the spatial and temporal pattern of the unbalanced distribution system, we define the node feature as $x_{i,t} = [\overline{\mathbf{P}}\mathbf{g}_{i,t}, \mathbf{P}\mathbf{d}_{i,t}, \mathbf{Q}\mathbf{d}_{i,t}, \mathbf{V}_i, \overline{\mathbf{V}}_i]$, $x \in \mathbb{R}^{N \times 11 \times T}$. Here $\overline{\mathbf{P}}\mathbf{g}_{i,t}$ is a vector of the three-phase DER maximum active power generation on bus i at time t . $\mathbf{P}\mathbf{d}_{i,t}$ and $\mathbf{Q}\mathbf{d}_{i,t}$ are the three-phase active and reactive load demand on bus i at time t . In order to make the graph homogeneous, for buses that do not have DER and/or load, we set $\overline{\mathbf{P}}\mathbf{g}_{i,t}$ and/or $\mathbf{P}\mathbf{d}_{i,t}, \mathbf{Q}\mathbf{d}_{i,t}$ to zero vectors. For buses that are not three-phase, we use the value of 0 for the missing phases.

Additionally, the normalized graph Laplacian matrix is defined as $L = I_n - D^{-1/2}AD^{-1/2}$, where I_n is the identity matrix, D is the diagonal degree matrix. If there are two branches connected to bus i , then $D_{ii} = 2, D_{ij} = 0, \forall i \neq j$.

B. Extracting Spatial Features using Chebyshev Convolution

We adopt Chebyshev convolution [34] to extract the spatial features of the graph. This is because Chebyshev convolution enables localized information aggregation by allowing the model to focus on a limited number of neighboring nodes, which is particularly useful for graph-structured data where nodes have varying degrees of connectivity. Besides, Chebyshev convolution is computationally efficient, making it scalable to large graphs with many nodes and edges such as that of power distribution systems.

For each time period t , the node embedding is calculated using the following graph convolution (36). Here $*_{\mathbb{G}}$ is the spatial convolution operator, which represents the multiplication of a signal x_t with a spatial kernel Θ . By setting different

respective fields K_s , the information of K_s -hop neighbors will be propagated to its center node.

$$x'_t = \Theta *_{\mathbb{G}} x_t = \Theta(L)x_t \approx \sum_{k=0}^{K_s-1} \theta_k T_k(\tilde{L})x_t, \quad (36)$$

where Θ is the convolution kernel. $*_{\mathbb{G}}$ is the graph convolution operator. $\tilde{L} = \frac{2L}{\lambda_{\max}} - I$ is the scaled Laplacian matrix. λ_{\max} is the largest eigenvalue of L . $T_k(\tilde{L}) \in \mathbb{R}^{n \times n}$ is the k -th order Chebyshev polynomial, with $T_k(\tilde{L}) = 2\tilde{L}T_{k-1}(\tilde{L}) - T_{k-2}(\tilde{L}), T_0(\tilde{L}) = I_n, T_1(\tilde{L}) = \tilde{L}$. θ_k is the weight matrix.

C. Extracting Temporal Features using Causal Convolution

We use Causal convolution with a K_t -width kernel to extract the temporal features, which can maintain the temporal causality of the input sequence. The Causal convolution is followed by a gated linear unit (GLU) to model non-linearity.

For each node i , the node embedding is calculated using the graph convolution (37). By varying the value of K_t , we explore the K_t neighbors of elements within the input sequence x_i . Similarly, we introduce $*_{\mathcal{T}}$ as the temporal convolution operator and Γ is the temporal kernel.

$$x'_i = \Gamma *_{\mathcal{T}} x_i = (\mathbf{W}_1 * x_i) \odot \sigma(\mathbf{W}_2 * x_i), \quad (37)$$

where $\mathbf{W}_1, \mathbf{W}_2 \in \mathbb{R}^{k_t \times c_i \times c_o}$ are casual convolution kernels. c_i, c_o are the numbers of input channel and output channel of kernels $\mathbf{W}_1, \mathbf{W}_2$. $*$ is the convolution operation. $\sigma(\cdot)$ is the sigmoid function.

D. Spatiotemporal Graph Convolution Block

After setting up the Chebyshev and Casual convolution layer, we use them to construct a spatiotemporal convolution block to jointly process the spatiotemporal features of the physics-informed graph. As suggested by [35], we apply a ‘‘sandwich’’ structure, with a spatial layer positioned between two temporal layers. The output of block x' is calculated as:

$$x' = \Gamma_1 *_{\mathcal{T}} \max(0, \Theta *_{\mathbb{G}} (\Gamma_0 *_{\mathcal{T}} x)), \quad (38)$$

where Γ_0, Γ_1 are the upper and lower temporal kernels. Note that since x is a 3-D variable, the temporal kernels Γ_0, Γ_1 are applied on every node of x while the spatial kernel is applied on every time step of x .

E. Link Classifier for Switches and Voltage Regulators

After processing nodal features with the spatiotemporal graph convolution module, we now propose a link classifier to predict the connectivity of switches and the tap position of voltage regulators. For every time step, we first obtain the embedding for each device as:

$$l_{ij} = \prod_{k=1}^{K_f} W_k(x'_i \odot x'_j), \forall (i, j) \in \mathcal{E}^s \cup \mathcal{E}^r, \quad (39)$$

where W_k is a weight matrix of fully-connected layer k . $\mathcal{E}^s, \mathcal{E}^r$ are the set of switches and the set of voltage regulators.

Then we calculate the probability that a switch is closed as:

$$\rho_{ij,t}^s = \sigma \left(\sum_{d=1}^D l_{ij,d,t} \right), \forall (i, j) \in \mathcal{E}^s, \quad (40)$$

where D is the output dimension of the final fully-connected layer K . Here we set D to be the maximum number of tap positions of all voltage regulators.

For voltage regulators, we adopt an ordinal representation [36] to encode the natural order between the discrete actions of setting the tap position. Specifically, for each tap position $m \leq M_{ij}$, the embedding for the voltage regulator is first pre-processed as:

$$l'_{ij,m,t} = \sum_{d=1}^m \ln l_{ij,d,t} + \sum_{d=m+1}^{M_{ij}} \ln (1 - l_{ij,d,t}), \quad (41)$$

where $l'_{ij,m,t}$ is the ordinal encoding of tap position m of the voltage regulator (i, j) at time t . Then, the probability distribution for tap positions can be calculated as:

$$\rho_{ij,m,t}^r = \frac{\exp(l'_{ij,m,t})}{\sum_d^{M_{ij}} \exp(l'_{ij,d,t})}, \forall m \in M_{ij} \quad (42)$$

The overall loss function for the neural network model has three components:

$$\mathcal{L}(\Xi, \mathbb{W}) = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3, \quad (43)$$

where Ξ is the set of all convolution kernels. \mathbb{W} is the set of weight matrices associated with fully-connected layers.

The first component is the logistic loss function for the switches, designed to enhance the accuracy of predicting switch statuses.

$$\mathcal{L}_1 = - \sum_{t=1}^T \sum_{ij \in \mathcal{E}^s} (\alpha_{ij,t} \log(\rho_{ij,t}^s) + (1 - \alpha_{ij,t}) \log(1 - \rho_{ij,t}^s)) \quad (44)$$

The second component of the loss function incorporates the physical network's radiality constraint, ensuring the topology of the power distribution grid remains radial.

$$\mathcal{L}_2 = \sum_{t=1}^T \left(\sum_{ij \in \mathcal{E}^s} \lfloor \rho_{ij,t}^s + 0.5 \rfloor - (N - N_s) \right)^2 \quad (45)$$

The third component of the loss function corresponds to the cross entropy loss for voltage regulators, which is used to improve the prediction accuracy of tap position.

$$\mathcal{L}_3 = - \sum_{t=1}^T \sum_{m=1}^{M_{ij}} \sum_{ij \in \mathcal{E}^r} z_{ij,m,t} \log(\rho_{ij,m,t}^r), \quad (46)$$

where $z_{ij,m,t}$ is a binary indicator used in SOS1 that is set to 1 if the tap position of voltage regulator (i, j) is at position m at time t , and 0 otherwise.

F. Sub-MIP Solving with Variable Predictions

Given the predictions of the decision variables, we could fix those with high prediction performance and solve the simplified sub-MIP problems. We establish two evaluation metrics for prediction performance associated with switch and VR variables. For all time steps t within the set T , we define the prediction accuracy (Acc) and prediction discrepancy (Dcc) as follows:

$$Acc(\hat{\alpha}_{ij,t}) = \frac{1}{N_g} \sum_{n=1}^{N_g} \mathbf{1}(y_{ij,t}^s(\mathbb{G}^n) = \alpha_{n,ij,t}) \times 100\%, \quad (47)$$

$$Dcc(\hat{\tau}_{ij,t}) = \frac{1}{N_g} \sum_{n=1}^{N_g} |y_{ij,t}^r(\mathbb{G}^n) - \tau_{n,ij,t}| \times 100\%, \quad (48)$$

where $\hat{\alpha}_{ij,t}$, $\hat{\tau}_{ij,t}$ are the switch and VR variables that we aim to predict. N_g is the number of graphs in the dataset. $y_{ij,t}^s = \lfloor \rho_{ij,t}^s + 0.5 \rfloor$ is the predicted status for the switch $(i, j) \in \mathcal{E}^s$. $y_{ij,t}^r = \operatorname{argmax}_{m \in M_{ij}} \{\rho_{ij,m,t}^r\}$ is the predicted tap position for the voltage regulator $(i, j) \in \mathcal{E}^r$.

The detailed process of sub-MIP solving is given in Algorithm 1. The main steps of the algorithm include formulating the MIP instance and constructing the graph based on system parameters. It then iterates over time steps and edges, calculating prediction accuracies for switch variables (Acc) and prediction discrepancies for VR variables (Dcc) using the validation dataset. Based on these metrics and threshold comparisons, the algorithm selectively fixes variables by adding constraints to the MIP instance. Finally, it solves the simplified sub-MIP to obtain the desired outputs, including switch statuses, tap positions of VRs, and DER outputs.

Algorithm 1 Sub-MIP Solving

Input: System topology, load, and DER data, learned probability distribution ρ_{θ}^s , ρ_{θ}^r , switch accuracy and VR discrepancy thresholds ϵ^s , ϵ^r .

Output: switch statuses, tap positions of VRs, and DER output.

- 1: Formulate MIP instance I and graph \mathbb{G} based on system topology, load and DER data.
 - 2: **for** $t \in T$ **do**
 - 3: **for** $(i, j) \in \mathcal{E}^s$ **do**
 - 4: Calculate the prediction Acc of variable $\hat{\alpha}_{ij,t}$ on validation dataset according to (47).
 - 5: **if** $Acc(\hat{\alpha}_{ij,t}) \geq \epsilon^s$ **then**
 - 6: Add constraint $\alpha_{ij,t} = y_{ij,t}^s(\mathbb{G})$ to I .
 - 7: **end if**
 - 8: **end for**
 - 9: **for** $(i, j) \in \mathcal{E}^r$ **do**
 - 10: Calculate the prediction Dcc of variable $\hat{\tau}_{ij,t}$ on validation dataset according to (48).
 - 11: **if** $Dcc(\hat{\tau}_{ij,t}) \leq \epsilon^r$ **then**
 - 12: Add constraint $\tau_{ij,t} = y_{ij,t}^r(\mathbb{G})$ to I .
 - 13: **end if**
 - 14: **end for**
 - 15: **end for**
 - 16: Solve the simplified sub-MIP I' and obtain the output.
-

IV. CASE STUDIES

In this section, we begin by presenting the numerical setup. Subsequently, we explain the training process for both the baseline neural network models and our proposed STGCN model. Following that, we provide a comprehensive comparison of the performance of the proposed and baseline models, along with analyses of the voltage and solar PV profiles of testing instances. Lastly, we examine the sensitivity of DER curtailment and voltage deviation to the ratio of their cost coefficients c_g/c_d .

A. Numerical Study Setup

In this subsection, we present the numerical setup for our study. We selected the modified IEEE 123-bus feeder system [37] as the test system. As depicted in Fig. 3, the system exhibits significant unbalance owing to its multi-phase distribution line configurations. Note that in visualizing the feeder, we excluded voltage regulators, shunt capacitors, and transformers. To study distribution systems with significant DERs, we integrated 14 solar PV stations in the feeder. In addition, 4 voltage regulators, 6 tie switches, and 3 sectionalizers are incorporated into the feeder as shown in Fig. 3. The key parameters of the MIQCP are given in Table I. Note that for all solar PV stations, their capacity are set to 50 KVA.

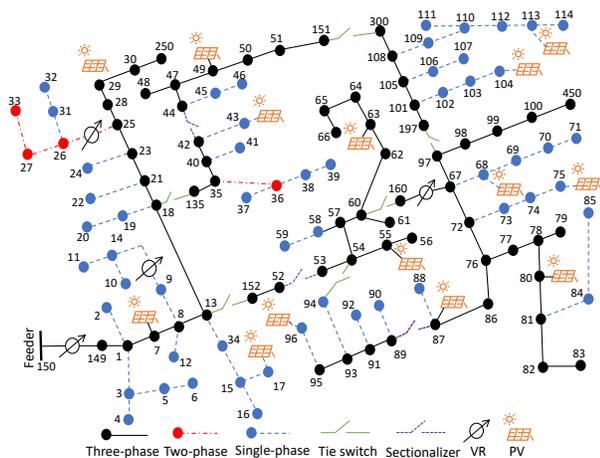


Fig. 3. 123-bus unbalanced distribution system.

TABLE I
KEY PARAMETERS OF THE MIQCP PROBLEM

Parameter	Value	Parameter	Value
c_g	10 k\$/KWh	δ	0.1
c_d	1 k\$/(KV) ²	M_{ij}	11
V_i	0.90 p.u.	$Sg_{i,t}$	50 KVA
\bar{V}_i	1.10 p.u.	\bar{S}_{ij}	2000 KVA
V_i^r	0.95 p.u.	γ_{\max}	10
\bar{V}_i^r	1.05 p.u.	λ_{\max}	3

We adopt the time series load and solar PV data [38] created by the National Renewable Energy Laboratory (NREL) as the

dataset, which contains one year's worth of data. We allocate 300 days of data for training, 30 days of data for model validation, and 30 days of data for testing. To assess the effectiveness of the proposed STGCN, we compare it with various benchmark neural networks, which include the fully-connected neural network (FNN), graph convolution network (GCN) [39], and the inductive graph convolution network (SAGE) [40]. The hyperparameters of the proposed STGCN and link classifier are given in Table II. The benchmark neural networks share the same hyperparameters with STGCN. All models were trained on a server with a 32-core AMD Ryzen Threadripper 3970X 3.7GHz CPU and three 10 GB NVIDIA GeForce RTX 2080 Ti. We adopt Gurobi [41] as our MIP solver. We set the MIP gap tolerance to 0.01% and impose a computation time limit of 20 minutes for the solver.

TABLE II
HYPERPARAMETERS OF THE PROPOSED STGCN AND LINK CLASSIFIER

Model	STGCN	Link classifier
Number of layers	2	3
Input dimension	11	64
Hidden dimension	64	64
Output dimension	64	11
Learning rate	0.005	
Batch size	32	
Training epochs	200	

B. Regulation Effect on Voltage Imbalance

In this subsection, we study the effect of regulation on voltage imbalance. We begin by defining the voltage imbalance factor (VIF) as follows:

$$VIF_{i,t}(\%) = \frac{V_{i,t}^{\max} - V_{i,t}^{\min}}{V_{i,t}^{avg}} \times 100, \forall i \in \mathcal{N}_3, t \in T, \quad (49)$$

where $V_{i,t}^{\max}$, $V_{i,t}^{\min}$, and $V_{i,t}^{avg}$ are the maximum, minimum, and average voltage magnitude of bus i at time t across three phases.

Then, we analyze the VIF of testing instances for two MIQCP formulations: with and without the voltage imbalance constraint (14). Table III presents the total VIF of all three-phase buses over 7 testing days for both formulations. For simplicity, we denote MIQCP without voltage imbalance constraint (14) as MIQCP w/o VIC. It is evident that with the voltage imbalance constraint (14), the solutions of the MIQCP model consistently achieve a lower VIF across all testing days.

TABLE III
TOTAL VIF OF TESTING INSTANCES FOR TWO FORMULATIONS (%)

Day	1	2	3	4	5	6	7
MIQCP	64.2	65.6	65.5	62.2	64.5	62.0	65.0
MIQCP w/o VIC	70.9	78.6	77.2	75.9	73.8	80.0	79.2

To further illustrate the regulation effect on voltage imbalance, we present the VIF of different buses on day 6 in Fig. 4. The results clearly show that constraint (14) effectively reduces the voltage imbalance to a lower level.

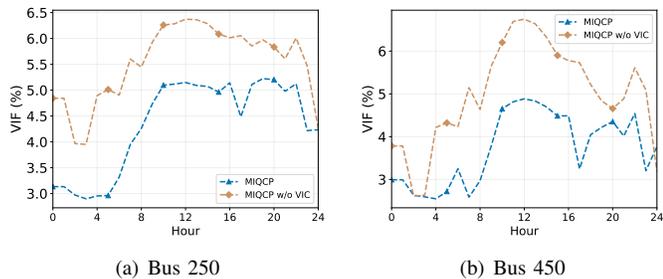


Fig. 4. VIF of different buses on day 6.

C. Training Process Evaluations

In this subsection, we train the baseline and proposed models using the collected dataset. Fig. 5 shows the validation loss of the proposed and three baseline models. The dark-colored curves are the average values across 10 experiments with different random seeds and the light-colored regions indicate the error bounds. The training times for the three baseline models and STGCN are approximately 10 minutes and 20 minutes, respectively. The extended training time for STGCN is due to the temporal convolution operation. As the training progresses, our proposed STGCN model consistently shows a lower loss compared to the baseline models.

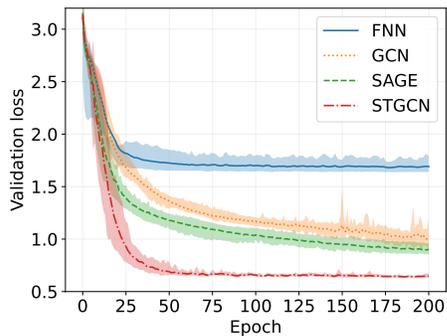


Fig. 5. Losses on validation dataset.

To further compare the performance of the baseline models and the proposed model, we show the average Acc of all switch variables and the average Dcc of all VR variables on the validation dataset. Fig. 6 shows the average Acc of switch variables and average Dcc of VR variables of four models, respectively. We can see that our proposed GNN model can achieve a higher prediction accuracy and a lower prediction discrepancy than the baseline models.

D. Testing Performance Comparisons

Once the offline training process concludes, we proceed to evaluate the performance of the proposed and baseline models by applying the trained models to the test dataset. Initially, we employ four trained models on the test dataset to assess their prediction capabilities. Table IV presents the average count of switch variables within various accuracy intervals across 10 runs with different random seeds, considering that the total number of switch variables for prediction is 216. Similarly,

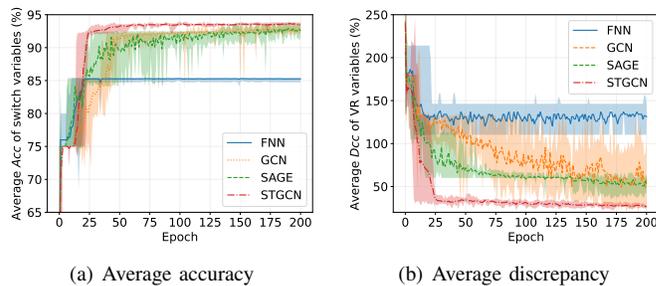


Fig. 6. Average prediction accuracy and discrepancy of four models on the validation dataset.

Table V illustrates the average count of VR variables within different discrepancy intervals across 10 runs with different random seeds, with a total of 96 VR variables for prediction. As shown in Tables IV and V, the proposed STGCN consistently outperforms the baseline models by yielding the highest prediction accuracy and the lowest prediction discrepancy.

TABLE IV
AVERAGE NUMBER OF SWITCH VARIABLES IN DIFFERENT ACCURACY INTERVALS

Algorithm	Acc (%) ≥ 80	≥ 85	≥ 90	≥ 95	$= 100$
FNN	133.7	131.0	121.2	116.2	107.2
GCN	143.2	141.2	134.0	129.0	119.0
SAGE	145.4	142.0	134.0	129.0	119.0
STGCN	154.0	149.0	138.0	133.0	123.0

TABLE V
AVERAGE NUMBER OF VR VARIABLES IN DIFFERENT DISCREPANCY INTERVALS

Algorithm	Dcc (%) ≤ 10	≤ 20	≤ 30	≤ 40	≤ 50
FNN	31.4	42.6	46.9	48.0	48.0
GCN	45.9	62.1	70.5	76.7	78.8
SAGE	47.3	62.5	69.3	71.9	72.4
STGCN	50.2	66.7	75.4	79.0	80.7

Then, we apply four trained models to the testing instances according to Algorithm 1. Here we set the thresholds $\epsilon^s = 1.0$, $\epsilon^r = 0.1$. We randomly select 7 days from the pool of 30 testing days for evaluation. The objective values and solution times of the four models are presented in Table VI and Table VII. A dash (“-”) in the table indicates instances where the algorithm did not find a feasible solution within the 20-minute time limit. Additionally, to demonstrate the effectiveness of our proposed model, we include the objective value and solution time of the baseline MIQCP approach.

From Table VI we can see that our proposed model combined with MIQCP can achieve the lowest objective values for all testing days. This suggests that incorporating the STGCN model enhances the optimization performance the most compared to other approaches. From Table VII we can see that the MIQCP + STGCN approach exhibits the lowest average solving time among all models, indicating its superior efficiency in solving testing instances with the shortest computation time.

TABLE VI
OBJECTIVE VALUE OF TESTING INSTANCES

Day	MIQCP	MIQCP +FNN	MIQCP +GCN	MIQCP +SAGE	MIQCP +STGCN
1	—	2.92	6.41	—	0.47
2	1.58	4.30	1.38	1.78	1.23
3	1.96	8.08	0.74	0.79	0.42
4	1.07	3.90	16.18	2.90	0.88
5	1.27	14.24	0.67	2.34	0.42
6	0.42	3.38	1.41	0.79	0.39
7	0.97	4.02	1.71	1.55	0.38
Avg obj	1.21	5.83	4.07	1.69	0.59

TABLE VII
SOLVING TIME OF TESTING INSTANCES (s)

Day	MIQCP	MIQCP +FNN	MIQCP +GCN	MIQCP +SAGE	MIQCP +STGCN
1	1200.0	146.6	102.2	—	43.7
2	1200.0	146.2	58.25	29.8	54.3
3	1200.0	92.6	55.5	52.0	42.9
4	1200.0	132.7	41.6	55.9	30.6
5	1200.0	54.9	55.6	29.9	41.9
6	1200.0	35.6	53.1	70.2	53.0
7	1200.0	124.1	21.6	37.3	48.6
Avg time	1200.0	104.7	55.4	45.9	45.0

E. Voltage and Solar PV Output Profiles

In this subsection, we present the voltage and solar PV output profiles of the testing instances. Following the resolution of the testing instances, we first extract the voltage profiles for day 3 to show the effectiveness of our proposed model. Fig. 7 illustrates the voltage profiles of phase A on bus 114 obtained using the proposed and four baseline approaches. Notably, our proposed method, which combines MIQCP with STGCN consistently upholds the voltage of bus 114 above 0.95 p.u. across all time periods. It is worth mentioning that the MIQCP with the FNN model exhibits the highest voltage deviation due to its low prediction performance for switch statuses and VRs' tap positions.

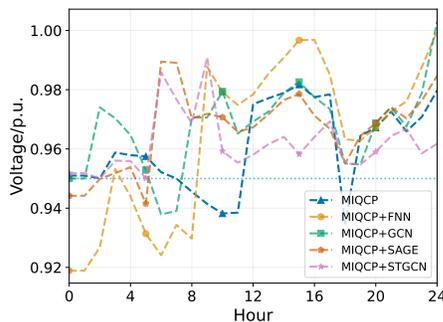


Fig. 7. Voltage profiles of bus 114 on phase A.

Similarly, Fig. 8 shows the voltage profiles of phase B on bus 160. Once again, our proposed STGCN model ensures that the MIQCP maintains the voltage of bus 160 below 1.05 p.u., except for hour 9. In contrast, other approaches exceed the preferred voltage upper bound for a greater number of hours. Fig. 9 shows the voltage profiles of phase C on bus 95 and

we can see that the MIQCP with STGCN provides the most stable voltage profile compared to other methods.

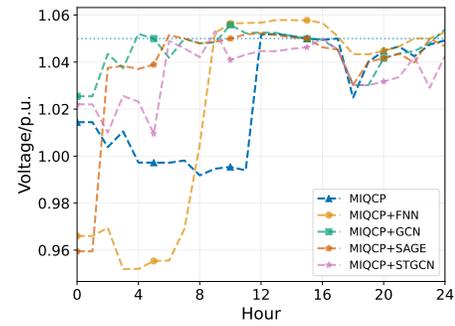


Fig. 8. Voltage profiles of bus 160 on phase B.

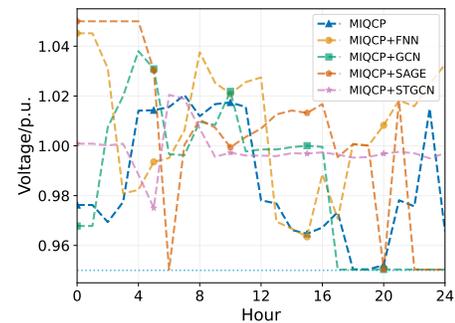


Fig. 9. Voltage profiles of bus 95 on phase C.

Next, we extract the solar PV output profiles of testing day 3 and compute the generation curtailments. Fig. 10 illustrates the curtailment of solar PV generations obtained by five different approaches. Due to space constraints, we present only the curtailments of solar PV output at buses 63, 68, 75, 80, 87, 101, and 113. It is evident that the proposed model, which combines MIQCP with the STGCN model leads to zero curtailment, whereas other baseline approaches all yield generation curtailment to some extent.

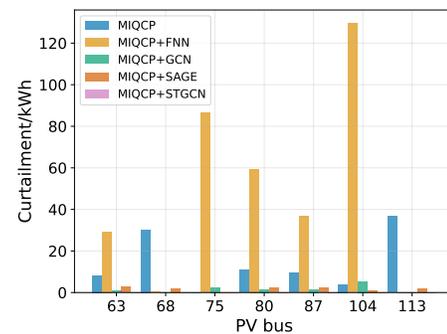


Fig. 10. Overall solar PV generation curtailment at different buses of day 3.

F. Sensitivity to the Ratio of Cost Coefficients c_g/c_d

In this section, we investigate the sensitivity of solar PV curtailment and voltage deviation to the ratio of their cost coefficients c_g and c_d , where c_g and c_d regulate the weighting of

DER curtailment and voltage deviation in the objective function. In real-world operations, distribution system operators may wish to manage DER curtailment and voltage deviation in accordance with their preferences or system conditions.

For every setting of c_g/c_d , we gather 300 training samples and 30 validation samples to train four neural network models. Subsequently, we combine these trained models with MIQCP, to perform joint network reconfiguration and voltage regulation on 7 random testing days. Across each c_g/c_d setting, we select identical testing days and present the average values of solar PV curtailment and voltage deviation in Table VIII.

TABLE VIII
SOLAR PV CURTAILMENT AND VOLTAGE DEVIATION UNDER DIFFERENT RATIOS OF c_g/c_d

c_g/c_d ((KV) ² /KWh)	Solar PV curtailment (KWh)				
	MIQCP	MIQCP +FNN	MIQCP +GCN	MIQCP +SAGE	MIQCP +STGCN
0.1	1069.7	738.8	713.3	353.5	96.4
0.5	877.3	735.0	656.2	190.6	93.8
1	863.3	638.0	322.0	162.6	23.9
5	163.5	527.3	123.1	82.8	21.5
10	55.8	329.0	101.4	42.1	13.2
c_g/c_d ((KV) ² /KWh)	Voltage deviation (KV)				
	MIQCP	MIQCP +FNN	MIQCP +GCN	MIQCP +SAGE	MIQCP +STGCN
0.1	0.89	4.37	3.19	2.02	0.69
0.5	1.24	4.96	3.96	2.07	0.74
1	1.25	6.59	4.85	2.17	0.75
5	1.41	7.30	5.65	2.49	1.06
10	1.56	12.58	6.24	3.91	1.20

As the ratio c_g/c_d increases, the solar PV curtailment decreases across all five approaches, while the voltage deviation increases. We can also observe that solar PV curtailment demonstrates greater sensitivity to variations in the c_g/c_d ratio. Finally, among all the approaches, MIQCP + STGCN consistently achieves the lowest levels of solar PV curtailment and voltage deviation.

V. CONCLUSION

This paper presents a novel approach to perform joint dynamic reconfiguration and voltage regulation in unbalanced three-phase distribution systems. We proposed an approximated MIQCP model for the problem and introduced a novel formulation for VR tap-settings based on SOS1. To improve computation efficiency, we proposed a physics-informed STGCN, which effectively captures the spatial and temporal load, DER, and voltage information in producing accurate predictions for tie switches/sectionalizers and VR tap positions. Leveraging these probabilistic predictions, we devise an efficient algorithm for sub-MIP solving after fixing certain binary variables. Numerical studies demonstrate the superiority of our proposed approach over baseline models, offering improved solutions with reduced DER curtailment and voltage deviation with shorter computation time. The proposed model and numerical study results showcase the potential of physics-informed machine learning in improving the operation efficiency of unbalanced distribution systems.

VI. ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support provided by the University of California Office of the President, under award number L22CR4556.

REFERENCES

- [1] S. Babaei, R. Jiang, and C. Zhao, "Distributionally robust distribution network configuration under random contingency," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 3332–3341, 2020.
- [2] S. Lei, C. Chen, Y. Song, and Y. Hou, "Radiality constraints for resilient reconfiguration of distribution systems: Formulation and application to microgrid formation," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 3944–3956, 2020.
- [3] Q. Shi, F. Li, M. Olama, J. Dong, Y. Xue, M. Starke, C. Winstead, and T. Kuruganti, "Network reconfiguration and distributed energy resource scheduling for improved distribution system resilience," *Int. J. of Elect. Power & Energy Syst.*, vol. 124, p. 106355, 2021.
- [4] M. R. Dorostkar-Ghamsari, M. Fotuhi-Firuzabad, M. Lehtonen, and A. Safdarian, "Value of distribution network reconfiguration in presence of renewable energy resources," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 1879–1888, 2015.
- [5] A. Parchure, S. J. Tyler, M. A. Peskin, K. Rahimi, R. P. Broadwater, and M. Dilek, "Investigating PV generation induced voltage volatility for customers sharing a distribution service transformer," *IEEE Trans. Ind. Appl.*, vol. 53, no. 1, pp. 71–79, 2016.
- [6] K. Jhala, B. Natarajan, and A. Pahwa, "Probabilistic voltage sensitivity analysis (PVSA)—A novel approach to quantify impact of active consumers," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2518–2527, 2017.
- [7] M. Mahdavi, H. H. Alhelou, N. D. Hatzargyriou, and F. Jurado, "Reconfiguration of electric power distribution systems: Comprehensive review and classification," *IEEE Access*, vol. 9, pp. 118 502–118 527, 2021.
- [8] Y. Liu, J. Li, and L. Wu, "Coordinated optimal network reconfiguration and voltage regulator/DER control for unbalanced distribution systems," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2912–2922, 2018.
- [9] F. Ding and K. A. Loparo, "Hierarchical decentralized network reconfiguration for smart distribution systems—Part I: Problem formulation and algorithm development," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 734–743, 2014.
- [10] Y. Ch, S. Goswami, and D. Chatterjee, "Effect of network reconfiguration on power quality of distribution system," *Int. J. of Elect. Power & Energy Syst.*, vol. 83, pp. 87–95, 2016.
- [11] Y.-K. Wu, C.-Y. Lee, L.-C. Liu, and S.-H. Tsai, "Study of reconfiguration for the distribution system with distributed generators," *IEEE Trans. Power Del.*, vol. 25, no. 3, pp. 1678–1685, 2010.
- [12] R. S. Rao, K. Ravindra, K. Satish, and S. Narasimham, "Power loss minimization in distribution system using network reconfiguration in the presence of distributed generation," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 317–325, 2012.
- [13] S. Chen, W. Hu, and Z. Chen, "Comprehensive cost minimization in distribution networks using segmented-time feeder reconfiguration and reactive power control of distributed generators," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 983–993, 2015.
- [14] R. V. A. Monteiro, J. P. Bonaldo, R. F. da Silva, and A. S. Bretas, "Electric distribution network reconfiguration optimized for PV distributed generation and energy storage," *Elect. Power Syst. Res.*, vol. 184, p. 106319, 2020.
- [15] C. Wang, S. Lei, P. Ju, C. Chen, C. Peng, and Y. Hou, "MDP-based distribution network reconfiguration with renewable distributed generation: Approximate dynamic programming approach," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3620–3631, 2020.
- [16] J. Zhan, W. Liu, C. Chung, and J. Yang, "Switch opening and exchange method for stochastic distribution network reconfiguration," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 2995–3007, 2020.
- [17] W. Zheng, W. Huang, D. J. Hill, and Y. Hou, "An adaptive distributionally robust model for three-phase distribution network reconfiguration," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1224–1237, 2020.
- [18] H. Zhai, M. Yang, B. Chen, and N. Kang, "Dynamic reconfiguration of three-phase unbalanced distribution networks," *Int. J. of Elect. Power & Energy Syst.*, vol. 99, pp. 1–10, 2018.
- [19] B. A. Robbins and A. D. Domínguez-García, "Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2903–2913, 2015.

- [20] K. M. Muttaqi, A. D. Le, M. Negnevitsky, and G. Ledwich, "A coordinated voltage control approach for coordination of OLTC, voltage regulator, and DG to regulate voltage in a distribution feeder," *IEEE Trans. Ind. Appl.*, vol. 51, no. 2, pp. 1239–1248, 2015.
- [21] M. Ahmed, R. Bhattacharai, S. J. Hossain, S. Abdelrazek, and S. Kamasalasan, "Coordinated voltage control strategy for voltage regulators and voltage source converters integrated distribution system," *IEEE Trans. Ind. Appl.*, vol. 55, no. 4, pp. 4235–4246, 2019.
- [22] S. Mahdavi, H. Panamtash, A. Dimitrovski, and Q. Zhou, "Predictive coordinated and cooperative voltage control for systems with high penetration of PV," *IEEE Trans. Ind. Appl.*, vol. 57, no. 3, pp. 2212–2222, 2021.
- [23] W.-H. Liu, A. D. Papalexopoulos, and W. F. Tinney, "Discrete shunt controls in a Newton optimal power flow," *IEEE Trans. Power Syst.*, vol. 7, no. 4, pp. 1509–1518, 1992.
- [24] E. Acha, H. Ambriz-Perez, and C. Fuerte-Esquivel, "Advanced transformer control modeling in an optimal power flow using Newton's method," *IEEE Trans. Power Syst.*, vol. 15, no. 1, pp. 290–298, 2000.
- [25] M. Adibi, R. Polyak, I. Griva, L. Mili, and S. Ammari, "Optimal transformer tap selection using modified barrier-augmented Lagrangian method," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 251–257, 2003.
- [26] B. A. Robbins, H. Zhu, and A. D. Domínguez-García, "Optimal tap setting of voltage regulation transformers in unbalanced distribution systems," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 256–267, 2015.
- [27] A. Savasci, A. Inaolaji, and S. Paudyal, "Optimal tuning of local voltage control rule of load tap changers for dynamic operation of unbalanced distribution networks," *IEEE Trans. Ind. Appl.*, vol. 60, no. 1, pp. 1322–1323, 2024.
- [28] B. Huang and J. Wang, "Applications of physics-informed neural networks in power systems—a review," *IEEE Trans. Power Syst.*, vol. 38, no. 1, pp. 572–588, 2022.
- [29] Y. Du, Q. Huang, R. Huang, T. Yin, J. Tan, W. Yu, and X. Li, "Physics-informed evolutionary strategy based control for mitigating delayed voltage recovery," *IEEE Trans. Power Syst.*, vol. 37, no. 5, pp. 3516–3527, 2021.
- [30] K. Mahapatra, X. Fan, X. Li, Y. Huang, and Q. Huang, "Physics informed reinforcement learning for power grid control using augmented random search," Pacific Northwest National Lab (PNNL), Richland, WA (United States), Tech. Rep., 2022.
- [31] R. R. Hossain, K. Mahapatra, Q. Huang, and R. Huang, "Physics-informed deep reinforcement learning-based adaptive generator out-of-step protection for power systems," in *IEEE Power & Energy Soc. General Meeting*, 2023, pp. 1–5.
- [32] J. Authier, R. Haider, A. Annaswamy, and F. Dorfler, "Physics-informed graph neural network for dynamic reconfiguration of power systems," *arXiv preprint arXiv:2310.00728*, 2023.
- [33] J. Qin and N. Yu, "Reconfigure distribution network with physics-informed graph neural network," in *Proc. IEEE PES Innovative Smart Grid Technol., Europe*, 2023, pp. 1–6.
- [34] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," *Proc. NeurIPS*, vol. 29, 2016.
- [35] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, 2018, pp. 3634–3640.
- [36] Y. Tang and S. Agrawal, "Discretizing continuous action space for on-policy optimization," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 04, 2020, pp. 5981–5988.
- [37] W. H. Kersting, "Radial distribution test feeders," in *Proc. IEEE Power Eng. Soc. Winter Meeting*, vol. 2, 2001, pp. 908–912.
- [38] B. K. Elgindy, Tarek, "Sample IEEE123 Bus system for OEDI SI." [Online]. Available: <https://data.openei.org/submissions/5773>
- [39] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Int. Conf. Learn. Representat.*, 2016.
- [40] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *Proc. NeurIPS*, vol. 30, 2017.
- [41] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," 2024. [Online]. Available: <https://www.gurobi.com>



Jingtao Qin (S'21) He received his B.S. and M.S. degrees in Electrical Engineering from Shandong University, Jinan, China, in 2018 and 2020. He is now pursuing his Ph.D. degree in Electrical Engineering from the University of California, Riverside. His research interests include machine learning, optimization and their applications in power systems. Specifically, he focuses on applications such as unit commitment, network reconfiguration, and voltage regulation in power distribution systems.



Rui Yang (Member, IEEE) received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2009, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2014. She is the manager of the Sensing and Predictive Analytics group in the Power Systems Engineering Center, National Renewable Energy Laboratory, Golden, CO, USA. Her research interests include advanced data analytics, machine learning, and optimization for power systems applications.



Nanpeng Yu (Senior Member, IEEE) received the B.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 2006, and the M.S. and Ph.D. degrees in electrical engineering from Iowa State University, Ames, IA, USA, in 2007 and 2010, respectively. He is a Full Professor and Vice Chair with the Department of Electrical and Computer Engineering and Director of Energy, Economics, and Environment Research Center at University of California, Riverside, CA, USA. His current research interests include physics-informed machine learning

in smart grids, electricity market design and optimization, transportation electrification, and decarbonization planning.