Voltage Calculations in Secondary Distribution Networks via Physics-Inspired Neural Network Using Smart Meter Data

Liming Liu, *Graduate Student Member, IEEE*, Naihao Shi, *Graduate Student Member, IEEE*, Dingwei Wang, *Graduate Student Member, IEEE*, Zixiao Ma, *Member, IEEE*, Zhaoyu Wang, *Senior Member, IEEE*, Matthew J. Reno, *Senior Member, IEEE*, and Joseph A. Azzolini, *Member, IEEE*

***Abstract*—The increasing penetration of distributed energy resources (DERs) leads to voltage issues across distribution networks, necessitating voltage calculations by utilities. Electric model-free voltage calculation offers an enticing solution. However, most researches mainly focus on primary distribution networks ignoring secondary distribution networks and commonly overlook extreme voltage case calculations, which require the model's extrapolation abilities. In addressing the gaps, this paper presents a customized physics-inspired neural network (PINN) model, the structure of which is inspired by the derived coupled power flow model of primary-secondary distribution networks. To ensure precision and rapid convergence, a crafted training framework for the PINN model is proposed. The PINN's "structure-mimetic" design enables superior extrapolation for unseen scenarios and enhances physical information awareness. We demonstrate this through two applications: hosting capacity analysis and customer-transformer connectivity. The effectiveness and advantages of the proposed PINN model are validated on two public testing systems and one utility distribution feeder model.**

***Index Terms*—Distribution network, voltage calculation, electric model-free, physics-inspired neural network, extrapolation.**

NOMENCLATURE

Abbreviations

DER	Distributed energy resource
EN	Euclidean norm
EV	Electric vehicle
HC	Hosting capacity
LR	Learning rate
MAPE	Mean absolute percentage error
MLP	Multi-layer perception
MN	Manhattan norm
MSE	Mean squared error
OLTCs	On-load tap changers
PDNet	Primary distribution network
PFlw	Power flow
PINN	Physics-inspired neural network
PV	Photovoltaic
SDNet	Secondary distribution network
SGD	Stochastic gradient descent
SM	Smart meter
ST	Service transformer
TC	Transformer-customer

Constants

$[\boldsymbol{a}_0^{J*}, \boldsymbol{A}^{J*}]$	Incidence matrix of the radial SDNet J
α_0	Initial learning rate
$\boldsymbol{E}$	Coefficient matrix of customer active power
$\boldsymbol{G}$	Minimum connection matrix
$\boldsymbol{H}$	Coefficient matrix of customer rective power
$\boldsymbol{D}_r$	Line resistance matrices of PDNet
$\boldsymbol{D}_x$	Line reactance matrices of PDNet
δ	Factor for scaling the $L^{\eta}_{\theta_\eta}$
$\left[\boldsymbol{A}_0, \boldsymbol{A}^T\right]$	Incidence matrix of the radial PDNet graph
k	ST number
N	Number of the buses (except slack bus) in the PDNet
N_b	Data batch size for training
N_c	Total number of load buses in the feeder

Indices and Sets

$\mathcal{N}^s$	Index set of PDNet buses connected with SDNets
$\mathcal{N}^{J*}$	Non-head abuse index set of SDNet ($n_0{}^{J*}$, ϕ_J)
Θ	Parameter set of PINN model
$\{0\}\bigcup\mathcal{N}^p$	Index set of buses in the PDNet
n_0^{J*}	PDNet bus connected with SDNet J on phase-ϕ_J

Manuscript received 14 September 2023; revised 1 February 2024; accepted
AQ1 25 April 2024. This work was supported in part by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office under Award 38426, and in part by the National Technology & Engineering Solutions of Sandia, LLC, with the U.S. Department of Energy (DOE) under Contract DE-NA0003525. Paper no. TSG-01488-2023. *(Corresponding author: Zhaoyu Wang.)*

AQ2 Liming Liu, Naihao Shi, Dingwei Wang, and Zhaoyu Wang are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA (e-mail: wzy@iastate.edu).

Zixiao Ma is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA, and also with the Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98195 USA.

Matthew J. Reno and Joseph A. Azzolini are with the Electric Power Systems Research, Sandia National Laboratories, Albuquerque, NM 87110 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2024.3396434.

Digital Object Identifier 10.1109/TSG.2024.3396434

Variables

$[\boldsymbol{v}, \boldsymbol{v}_0]$	Squared voltage magnitudes of the PDNet buses
$\boldsymbol{b}_c$	Bias vector of $\eta_{\mathrm{c}}^{\mathrm{s}}$ layer
$\boldsymbol{p}$	Nodal injection active power of the PDNet
$\boldsymbol{q}$	Nodal injection reactive power of the PDNet
$\boldsymbol{W}_a$	Weight matrix of $\eta_{\mathrm{a}}^{\mathrm{p}}$ layer
$\boldsymbol{W}_b$	Weight matrix of $\eta_{\mathrm{b}}^{\mathrm{q}}$ layer
$\boldsymbol{P}$	Line active power of the PDNet
$\boldsymbol{Q}$	Line reactive power of the PDNet
$\mathcal{J}(\Theta)$	Total loss of the PINN model
$\mathcal{R}_\Theta$	Regularization term
θ_η	Parameters emerging physical information
θ_ϕ	Parameters without physical information embedded
$L_{\theta_\eta}^{\eta}$	Prediction error of physics-inspired module
L_Θ	PINN model prediction error
p_c	Active power collected from SMs
q_c	Reactive power collected from SMs
v_0^{I*}	Head node squared voltage of SDNet I
v_c	Squared voltage magnitudes derived from SMs

I. Introduction

PROLIFERATION of distributed energy resources (DERs), such as residential photovoltaic (PV) systems and electric vehicles (EVs), is reshaping modern distribution power networks. Spurred by technological growth and ecological needs, these DERs are increasingly connected to the low-voltage secondary distribution networks (SDNets), upending traditional energy practices. However, the integration of DERs introduces numerous operational and reliability hurdles. A prevalent issue is the voltage rise due to distributed PV, making it harder to maintain voltages within the ANSI C84.1 tolerances [1], [2], given the reverse power flow (PFlw) in the case of excess power generation. Hence, it is of great importance for utilities or distribution power companies to perform voltage calculations, enabling the design and development of effective voltage control strategies for the safe and reliable operation of distribution networks [3].

Voltage calculations rely on distribution network models, but these models are typically absent in SDNets populated by residential PV and electric vehicles. Although some utilities may record SDNet information, including topology, line parameters, and customer connectivity from transformers, maintaining or updating these models can be time-consuming and costly. As a result, these recorded models are mostly outdated or contain errors [4], critically impacting the accuracy of voltage calculations and model-based hosting capacity results [5].

As an alternative, electric model-free voltage calculation methods have gained increasing attraction with the rise of machine-learning technologies and the mass adoption of smart meters (SMs), presenting a promising solution to the outlined challenges. Rather than using electric power models for PFlw analysis, these methods leverage regression techniques to analyze historical SM data (i.e., P, Q, and V) and identify the correlation between load data and the voltage data from SMs. With this well-established mapping relationship, the voltage at customer nodes can be calculated in various scenarios by specifying the customers' active and reactive power (i.e., P and Q) at a given moment.

In recent years, there has been a significant upswing in scholarly interest in data-driven or model-free voltage calculation methodologies. These can be bifurcated into two primary categories: linear and nonlinear regression-based methods.

Class I - Linear regression-based methods: These methods mainly focus on the linearization of the PFlw model [6]. In the pioneering work by [7], a data-driven linearization approach of PFlw models was proposed, employing partial least squares-based and Bayesian linear regression-based algorithms to address collinearity and avoid overfitting of real operation data. Similarly, a robust data-driven linearization model utilizing linear support vector regression is presented in [8]. The ultimate goal of these methods is to estimate the parameters of the linearized PFlw model, then conduct voltage calculations based on these PFlw models. Further pushing the boundaries, a novel two-step regressor combining multiple techniques was proposed in [9]. This innovative methodology integrates linear and nonlinear regressors into a unified model, resulting in enhanced predictive capabilities, as evidenced by a substantial reduction in error across simulation scenarios.

Class II - Nonlinear regression-based methods: These methods leverage nonlinear regression, with a particular emphasis on neural network-related approaches, owing to their adeptness in capturing the inherent nonlinearities present in PFlw problems [10], [11], [12], [13], [14], [15], [16]. Specifically, authors in [10] put forth a deep belief network-based PFlw calculation method that, in addition to active/reactive power data, incorporated topology information to account for variability due to system topology changes. A deep neural network-based approach is proposed to depict the high-dimensional load-to-solution mapping and directly solved the optimal PFlw problem [17]. In [18], the authors introduced two voltage change prediction models leveraging deep neural networks, validated using three datasets. While the model's extrapolation capability was evaluated, the paper did not discuss methods for its enhancement.

Despite the valuable findings obtained from numerous studies focusing on developing model-free voltage calculation methods, several intricate challenges still necessitate further deliberation and exploration.

First, most existing studies focus only on primary distribution networks (PDNets), overlooking SDNets where SMs are usually installed. This oversight often results in the use of unconventional measurements, such as distribution transformer readings, making such methods incompatible with residential SM data. In response, neural networks are adopted in [13], [14] to model the relationships among historical SM data in the corresponding SDNet. However, the model's performance may falter when transformer-customer connectivity is inaccurate. This inaccurate connectivity information may also inflate calculation errors.

Second, many of these methods perform poorly for high-impact, low-probability extreme voltage scenarios (e.g., voltages are less than 0.95 p.u. or greater than 1.05 p.u.) due to insufficient extreme voltage scenario data [16]. However, the prediction performance for extreme voltage scenarios is crucial since those scenarios necessitate voltage control [9]. These scenarios require the neural network model to have extrapolation capabilities, given that target voltage values often reach the boundaries (e.g., 0.95 pu and 1.05 pu). Extrapolation refers to a model's ability to make accurate predictions for input data outside the range of its training data. While the model in [13] claimed enhanced extrapolation capabilities by adding aggregated active and reactive power of customers as input and forming multi-outputs, the core component of the model is still a multi-layer perception-based (MLP) model. Such a model has been shown to struggle with extrapolating most nonlinear tasks due to their linear extrapolation. The existing literature rarely discusses the reasons for the model's extrapolation ability they claimed [19].

Third, these previous model-free voltage calculation methods are typically black-box, lacking physics-informed interpretability. Unlike deep neural networks, PINNs offer enhanced interpretability and reliability in machine learning applications [20]. PINNs come in various paradigms, with the most prevalent one employing a physics-informed loss function to steer model training. For instance, Power-GNN, proposed in [21], addresses the state and parameter estimation challenges by constructing a loss function rooted in PFlw equation residuals. Reference [12] introduced a physics-guided neural network for PFlw problems, utilizing an MLP as encoder and a Kirchhoff's laws-based bi-linear neural network decoder. The model employs a tailored loss function to minimize voltage prediction errors and power mismatches, enhancing convergence and accuracy through the integration of physical laws. However, its adaptability to unbalanced primary-secondary integrated distribution networks remains uncertain. Beyond loss function modifications, another notable approach involves the physics-informed design of architecture. This strategy uses physical principles to guide the neural network's architecture, either by infusing physical significance into hidden layer outputs or by directly altering the network's connections. Reference [15] introduces a deep neural network with a skip-connection structure, inspired by the cyclic nature of the prox-linear solver, to facilitate efficient training. Reference [22] balances computational efficiency and PFlw analysis accuracy using an encoder-decoder framework and message propagation among nodes but is limited by its strong physical assumptions and dependence on the Newton-Raphson solver. Reference [13] proposes a model-free voltage calculation model incorporating total loads to address upstream voltage fluctuations but it lacks physical interpretability. Overall, prior studies rarely consider using customized and physical rule-inspired neural networks that are suitable for distribution networks to improve the performance and extrapolation ability of voltage calculation models [10], [14], [16], and how to combine the different paradigms can be further explored as well.

In light of these challenges, this study proposes a model-free voltage calculation method for distribution networks based on a customized PINN. The main contributions of this work are summarized as follows:

- This study presents a coupled distribution PFlw model for integrated primary-secondary networks, laying the foundation for the physics-inspired structure design of a customized neural network.
- This paper proposes a model-free voltage calculation method via a PINN tailored to the needs of diverse operational and planning scenarios. The proposed model's physics-inspired structure greatly enhances extrapolation capabilities beyond existing methods, supported by test results on the distribution models and the successful application in PV hosting capacity (HC) calculations.
- The proposed PINN model exploits its physics-inspired structure to capture the PDNet-SDNets' physical information, relying solely on SM data. Based on the extracted physical information, we develop a transformer-customer (TC) connectivity identification method, illustrating the PINN model's application in distribution power network information awareness tasks.

The rest of the paper is organized as follows. Section II presents the coupled linearization of the distribution power flow model for primary and secondary networks. The physics-inspired model free voltage calculation model is formulated in Section III. Section IV presents PINN voltage calculation model applications, including model-free locational PV hosting capacity calculation and transformer-customer connectivity identification. Numerical results on the proposed model are given in Section V, and the paper is concluded in Section VI.

II. PDNet-SDNets Coupled Power Flow Model

In this section, we develop a coupled distribution PFlw model for integrated primary-secondary networks to assist in designing the structure of the PINN model. Our focus is on a residential distribution feeder that comprises both the medium-voltage PDNet and the low-voltage SDNets. The SDNets consist of single-phase connections[1] that link to the PDNet through single-phase service transformers (STs). We operate under the assumption that all customers are connected to the feeder via SDNets and that the SM data for all these customers is readily accessible.

A. Linearization of Power Flow Model for PDNet and SDNets

Consider an unbalanced three-phase radial PDNet containing $N+1$ buses, whose index set can be represented as $\{0\}\bigcup\mathcal{N}^p$, where 0 denotes the slack bus and set $\mathcal{N}^p = \{1, 2, \ldots, N\}$ is the index set of all other buses in the PDNet. The indices of nodes that are connected with SDNets are denoted as $\mathcal{N}^s = \{n_0^{1*}, n_0^{2*}, \ldots, n_0^{I*}\}$, where $\mathcal{N}^s \subseteq \mathcal{N}^p$. Let

[1]Despite using a split-phase triplex cable in reality, our model approximates it as a single-phase 240V connection via Kron reduction and balanced current assumptions.

vectors $\boldsymbol{v}$, $\boldsymbol{p}$ and $\boldsymbol{q}$ collect the squared bus voltage magnitudes, nodal net active and reactive power consumption of the primary network. Based on the assumption that the line losses are small and that the voltages are nearly balanced [23], the PFlw relationship of the primary distribution system can be represented with the LinDistFlow model, and compactly expressed in a graph-based form [24]:

$$\begin{aligned} \boldsymbol{A}\boldsymbol{P} &= -\boldsymbol{p}, \\ \boldsymbol{A}\boldsymbol{Q} &= -\boldsymbol{q}, \\ \begin{bmatrix} \boldsymbol{A}_0, & \boldsymbol{A}^T \end{bmatrix} \begin{bmatrix} v_0 \\ \boldsymbol{v} \end{bmatrix} &= 2(\ \boldsymbol{D}_r\boldsymbol{P} + \boldsymbol{D}_x\boldsymbol{Q}\ ), \end{aligned} \tag{1}$$

where $\boldsymbol{A}_0 \in \mathbb{R}^{3N\times 3}$ represents the three-phase connection between bus$_0$ and each of the other buses; $\boldsymbol{A} \in \mathbb{R}^{3N\times 3N}$ is the incidence matrix for the PDNet that represents the three-phase connection among all non-head buses. $\boldsymbol{D}_r$ and $\boldsymbol{D}_x$ are block diagonal matrices that collect the line impedance matrices. The LinDistFlow model in (1) establishes a linear mapping from the PDNet's nodal power injections to the squared voltage magnitudes, and the linear relationship is determined by the system topology information.

Following the linearization of PDNet PFlw, we investigate the corresponding SDNet model, as residential customers are commonly connected to low-voltage SDNets. For convenience, we denote the selected network as SDNet(n_0^{J*}, ϕ_J), which signifies that the SDNet is electrically connected to bus n_0^{J*} of the PDNet through a phase-ϕ_J lateral line. For clarity in notation, we define $J* = \{n_0^{J*}, \phi_J\}$. Any variable with the superscript $J*$ denoted as $(\cdot)^{J*}$ signifies it belongs to the specific SDNet(n_0^{J*}, ϕ_J) connected to the PDNet.

By referring to the impedance of the ST's primary and secondary winding to the same voltage level, SDNet(n_0^{J*}, ϕ_J) can be considered a single-phase radial network. In this representation, the primary winding of the ST, identified by $n_0^{J*} \in \mathcal{N}^s$, acts as the head bus, and its squared voltage magnitude is denoted as v_0^{J*}, being an element of vector $\boldsymbol{v}_0$. Let $\mathcal{N}^{J*} = \{1, \ldots, n_s^{J*}\}$ be the index set of non-head buses in SDNet(n_0^{J*}, ϕ_J). Then, we collect the net bus consumption of active and reactive power, as well as squared nodal voltage magnitudes of the SDNet, into vectors $\boldsymbol{p}^{J*}$, $\boldsymbol{q}^{J*}$, and $\boldsymbol{v}^{J*}$. Similarly, assuming negligible line and transformer losses, the PFlw in the single-phase SDNet(n_0^{J*}, ϕ_J) can be approximately expressed by using the linearized DistFlow equations, which can be concisely represented in a graph-based compact form:

$$\begin{aligned} \boldsymbol{v}^{J*} = &-2\left[\boldsymbol{A}^{J*}\right]^{-T}\boldsymbol{R}^{J*}\left[\boldsymbol{A}^{J*}\right]^{-1}\boldsymbol{p}^{J*} \\ &-2\left[\boldsymbol{A}^{J*}\right]^{-T}\boldsymbol{X}^{J*}\left[\boldsymbol{A}^{J*}\right]^{-1}\boldsymbol{q}^{J*} - v_0^{J*}\left[\boldsymbol{A}^{J*}\right]^{-T}\boldsymbol{a}_0^{J*}, \end{aligned} \tag{2}$$

where $[\boldsymbol{a}_0^{J*}, [\boldsymbol{A}^{J*}]^T]^T \in \mathbb{R}^{(n_s^{J*}+1)\times n_s^{J*}}$ is the incidence matrix of the radial topology graph, $\boldsymbol{R}^{J*}$ and $\boldsymbol{X}^{J*}$ are diagonal matrices whose entries are the line resistance and reactance in the SDNet, respectively. Considering that $\boldsymbol{A}^{J*}$, $\boldsymbol{a}_0^{J*}$, $\boldsymbol{R}^{J*}$ and $\boldsymbol{X}^{J*}$ arise from the topology information of SDNet(n_0^{J*}, ϕ_J), (2) can be written in a more compact format as (3):

$$\boldsymbol{v}^{J*} = -\boldsymbol{B}^{J*}\boldsymbol{p}^{J*} - \boldsymbol{C}^{J*}\boldsymbol{q}^{J*} - v_0^{J*}\boldsymbol{m}^{J*}, \tag{3}$$

where

$$\begin{aligned} \boldsymbol{B}^{J*} &= 2\left[\boldsymbol{A}^{J*}\right]^{-T}\boldsymbol{R}^{J*}\left[\boldsymbol{A}^{J*}\right]^{-1} \in \mathbb{R}^{n_s^{J*}\times n_s^{J*}}, \\ \boldsymbol{C}^{J*} &= 2\left[\boldsymbol{A}^{J*}\right]^{-T}\boldsymbol{X}^{J*}\left[\boldsymbol{A}^{J*}\right]^{-1} \in \mathbb{R}^{n_s^{J*}\times n_s^{J*}}, \\ \boldsymbol{m}^{J*} &= \left[\boldsymbol{A}^{J*}\right]^{-T}\boldsymbol{a}_0^{J*} \in \mathbb{R}^{n_s^{J*}\times 1}. \end{aligned}$$

In this transformation, the complex coefficients are encapsulated within the newly introduced matrices $\boldsymbol{B}^{J*}$, $\boldsymbol{C}^{J*}$, and $\boldsymbol{m}^{J*}$. Notably, due to the inherent properties of the coefficient terms, both $\boldsymbol{B}^{J*}$ and $\boldsymbol{C}^{J*}$ manifest as symmetric matrices. In practical SDNets, not every bus is connected with load. Given the majority if measurements in the SDNet are obtained from SMs installed on the customer side, our attention is specifically directed toward buses serving customer loads. Buses of the SDNet without customer connections are excluded from consideration, as they do not yield measurable data. As a result, (3) can be further modified to represent the relationship among the SM measurements. Let $\mathcal{N}_c^{J*}$ denote the set of buses with loads in SDNet(n_0^{J*}, ϕ_J), where $\mathcal{N}_c^{J*} \subseteq \mathcal{N}^{J*}$. We define vectors $\boldsymbol{v}_c^{J*}$, $\boldsymbol{p}_c^{J*}$, and $\boldsymbol{q}_c^{J*}$ of size $c^{J*}\times 1$ to collect the squared voltage magnitudes, net active and reactive power consumption for all buses with loads. Here, c^{J*} represents the number of load buses in the network J, and (3) can be further reduced to (4):

$$\boldsymbol{v}_c^{J*} = -\boldsymbol{B}_c^{J*}\boldsymbol{p}_c^{J*} - \boldsymbol{C}_c^{J*}\boldsymbol{q}_c^{J*} - v_0^{J*}\boldsymbol{m}_c^{J*}, \tag{4}$$

where

$$\begin{aligned} \boldsymbol{B}_c^{J*} &= \left[\boldsymbol{B}_c^{J*}(x,y)\right]_{x\in\mathcal{N}_c^{J*}, y\in\mathcal{N}_c^{J*}}, \boldsymbol{B}_c^{J*} \in \mathbb{R}^{c^{J*}\times c^{J*}}, \\ \boldsymbol{C}_c^{J*} &= \left[\boldsymbol{C}_c^{J*}(x,y)\right]_{x\in\mathcal{N}_c^{J*}, y\in\mathcal{N}_c^{J*}}, \boldsymbol{C}_c^{J*} \in \mathbb{R}^{c^{J*}\times c^{J*}}, \\ \boldsymbol{m}_c^{J*} &= \left[\boldsymbol{m}_c^{J*}(x)\right]_{x\in\mathcal{N}_c^{J*}}, \boldsymbol{m}_c^{J*} \in \mathbb{R}^{c^{J*}\times 1}. \end{aligned}$$

The matrices $\boldsymbol{B}_c^{J*}$, $\boldsymbol{C}_c^{J*}$, and $\boldsymbol{m}_c^{J*}$ are derived from $\boldsymbol{B}^{J*}$, $\boldsymbol{C}^{J*}$, and $\boldsymbol{m}^{J*}$ by removing the entries associated with buses without connected loads. Fig. 1 depicts the architecture of the integrated primary-secondary distribution networks. Within this illustration, two SDNets connected to distinct phases are highlighted in blue and red, respectively, to provide a detailed representation of the network structure.

B. Primary-Secondary Distribution Network Combination

For a distribution network, the PDNet and SDNets are interconnected through STs. The aggregated power of the STs plays a crucial role in shaping the PFlw within the PDNet. Consequently, any changes in the PDNet's PFlw directly impact the sub-SDNets, specifically by altering the primary side voltage of the STs connecting them. When constructing the primary-secondary combined PFlw model, it is essential to consider this interdependence. The core of this coupling lies in the voltage of the ST's primary windings, which act as pivotal points. These voltage levels serve to connect the two-level PFlws, seamlessly integrating the PDNet and all SDNets into a unified framework. To comprehensively formulate the coupled PFlw model, we consolidate all SDNets into a compact expression, focusing on the role of $\boldsymbol{v}$ in connecting all components.

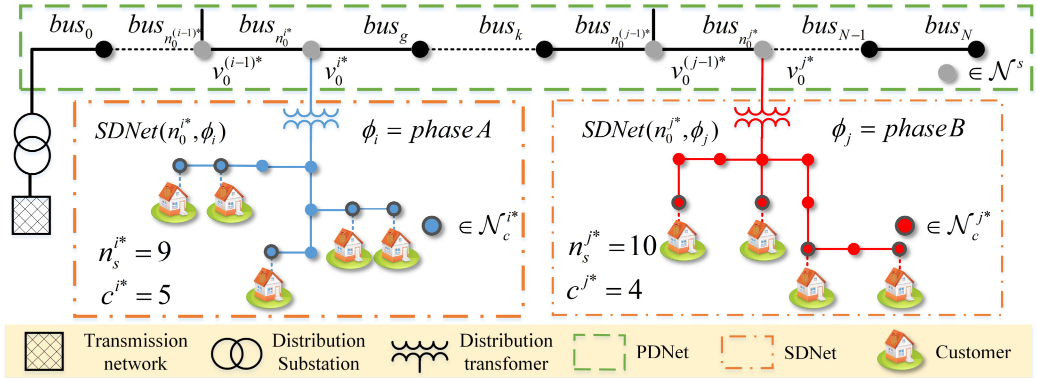


Fig. 1. The structure of integrated primary-secondary distribution networks.

Let I represent the number of SDNets in the distribution feeder. The measurements of all load buses in the SDNets can be compactly denoted by column vectors of size $N_c \times 1$, where $N_c = \sum_{J=1}^{J=I} n_c^{J*}$ represents the total number of load buses in the feeder, equaling the customer number. The last term in (4) can also be collected in a column vector as:

$$\boldsymbol{\mu}_c = \left[\left[\boldsymbol{\mu}_1\right]^T, \left[\boldsymbol{\mu}_2\right]^T, \ldots, \left[\boldsymbol{\mu}_I\right]^T \right]^T,$$

$$\boldsymbol{\mu}_J \in \left\{\boldsymbol{v}_c^{J*}, \boldsymbol{p}_c^{J*}, \boldsymbol{q}_c^{J*}\right\} \quad J \in \{1, 2, \ldots, I\},$$

$$\boldsymbol{m}_c = \left[\left[v_0^{1*}\boldsymbol{m}_c^{1*}\right]^T, \left[v_0^{2*}\boldsymbol{m}_c^{2*}\right]^T, \ldots, \left[v_0^{I*}\boldsymbol{m}_c^{I*}\right]^T \right]^T,$$

where $\boldsymbol{\mu}_c$ is a substitutable variable representing $\boldsymbol{p}_c$, $\boldsymbol{q}_c$ or $\boldsymbol{v}_c$, which denote the loading data and squared voltage data from the customer side. Then (4) can be expanded to reflect the relationship between voltage and power consumption of all load buses in the feeder:

$$\boldsymbol{v}_c = -\boldsymbol{B}_c\boldsymbol{p}_c - \boldsymbol{C}_c\boldsymbol{q}_c - \boldsymbol{m}_c, \tag{5}$$

where

$$\boldsymbol{B}_c = \mathrm{diag}\left(\boldsymbol{B}_c^{1*}, \boldsymbol{B}_c^{2*}, \ldots, \boldsymbol{B}_c^{I*}\right),$$

$$\boldsymbol{C}_c = \mathrm{diag}\left(\boldsymbol{C}_c^{1*}, \boldsymbol{C}_c^{2*}, \ldots, \boldsymbol{C}_c^{I*}\right).$$

In the context of the combined model represented by (5), several important points exist to be considered. Firstly, matrices $\boldsymbol{B}_c$ and $\boldsymbol{C}_c$ are derived from the topology information of all the SDNets. As per the equations, it is evident that these matrices are symmetrical and sparse. Secondly, the vector $\boldsymbol{m}_c$ is influenced by both the SDNet topology and the head node voltage $[v_0^{1*}, \ldots, v_0^{I*}]$ of the SDNets. The head nodes represent the primary side of the STs, directly connected to the primary network buses. This implies that the voltage of the primary network buses can impact the voltage of the load buses. At each time instance t, if we hold the constant components and separate the varying components of the voltage, the head node voltage can be expressed as $[v_s^{1*} + \Delta v^1(t), \ldots, v_s^{I*} + \Delta v^I(t)]$, where v_s^{I*} represents a constant voltage value, and $\Delta v^I(t)$ represents the voltage fluctuation at time t. As a result, the vector $\boldsymbol{m}_c$ can be decomposed into $\boldsymbol{m}_c^s$ and $\boldsymbol{m}_c^\Delta$, where $\boldsymbol{m}_c^s$ represents the constant component, and $\boldsymbol{m}_c^\Delta$ represents the fluctuating component. Notably, $\boldsymbol{m}_c^\Delta$ exhibits intricate relationships with customer load, voltage regulators, PDNet topologies, and other factors that can influence changes in PDNet's PFlw, making explicit calculation challenging. Thus, assuming fixed topologies and mainly considering customer loads and voltage regulators, the voltage variance item $\boldsymbol{m}_c^\Delta$ can be represented as $\boldsymbol{\Psi}(\boldsymbol{p_c}, \boldsymbol{q_c}, \boldsymbol{r})$, with $\boldsymbol{\Psi}(\cdot)$ representing the voltage variance relationship, and $\boldsymbol{r}$ denoting the actions of voltage regulators in the PDNet. Thirdly, for enhanced performance, accounting for linearization errors is crucial, especially when considering SDNets, which exhibit greater losses than PDNet due to their lower voltage. These errors are related to the squared line power and squared voltage terms, indicating their association with the quadratic terms of customers' net active/reactive power consumption and the squared voltage [24]. Due to its complexity, we represent the error term implicitly as $\boldsymbol{\chi}(\boldsymbol{p_c}, \boldsymbol{q_c}, \boldsymbol{v_c})$, where $\boldsymbol{\chi}(\cdot)$ represents the complex relationship. In summary, taking into account the above discussions, the final expression of the PDNet-SDNets coupled PFlw model can be written as follows:

$$\boldsymbol{v}_c = \boldsymbol{E}\boldsymbol{p}_c + \boldsymbol{H}\boldsymbol{q}_c - \boldsymbol{m}_c^s + \boldsymbol{\Psi}(\boldsymbol{p_c}, \boldsymbol{q_c}, \boldsymbol{r}) + \boldsymbol{\chi}(\boldsymbol{p_c}, \boldsymbol{q_c}, \boldsymbol{v_c}),$$

$$\boldsymbol{E} = -\boldsymbol{B}_c \; \boldsymbol{H} = -\boldsymbol{C}_c,$$

$$\boldsymbol{m}_c^s = \left[\left[v_s^{1*}\boldsymbol{m}_c^{1*}\right]^T, \left[v_s^{2*}\boldsymbol{m}_c^{2*}\right]^T, \ldots, \left[v_s^{I*}\boldsymbol{m}_c^{I*}\right]^T\right]^T. \tag{6}$$

In the next section, we will design the PINN model based on the format of the combined PFlw model mentioned above and the characteristics of the coefficient matrices.

III. Physics-Inspired Model-Free Voltage Calculation Method

A. Model-Free Voltage Calculation Problem Restatement

The essential thinking of our model-free voltage calculation method is to learn and model complex multi-dimensional underlying relationships between the input loads ($\mathcal{P}$, $\mathcal{Q}$) and corresponding voltages ($\mathcal{V}$). The relationship can be simply modeled as $\mathcal{V} = F(\mathcal{P}, \mathcal{Q})$. In our problem, load data $\mathcal{P}$, $\mathcal{Q}$, and $\mathcal{V}$ represent the system measurements in a period of time collected by customer-side SMs. Based on model training, the function $F(\cdot)$ can be obtained by estimating the model parameter $\boldsymbol{\theta}$ learned from the SM data. Unlike previous works, our approach does not rely entirely

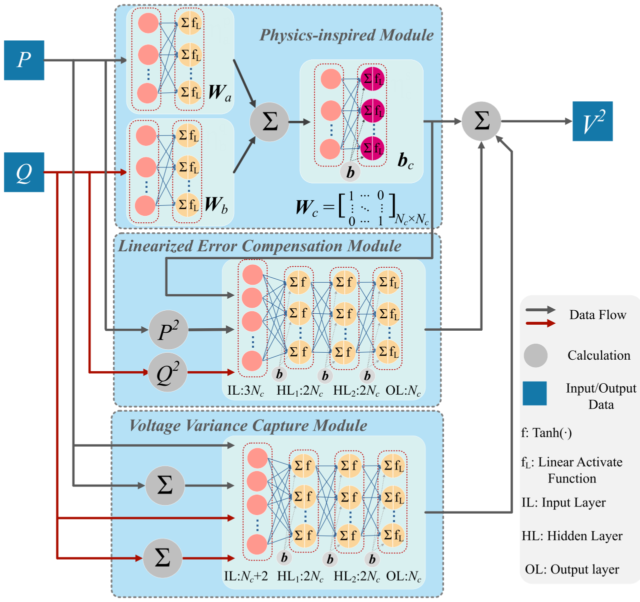


Fig. 2. The structure of the proposed PINN model.

on implicit PFlw relationship mapping. Instead, the PINN model is designed to implicitly learn the highly nonlinear components of the PFlw model that cannot be directly derived, while explicitly capturing the remainder and preserving the physical structure. With the integration of a physics-inspired structure, our goal is to enhance the model's extrapolation ability. The model's parameters thus consist of both physics-inspired and conventional components. Overall, this paper's focus can be summarized as $\boldsymbol{v}_c = F(\boldsymbol{p}_c, \boldsymbol{q}_c; \theta_\eta, \theta_\phi)$, where θ_η denotes the parameters emerging physical information; the other parameters are included into θ_ϕ. This paper will demonstrate how to design the $F(\cdot)$ model and present a customized framework to train the parameter $\Theta = \{\theta_\eta, \theta_\phi\}$ based on the dataset $\mathcal{D}_{tr} = \{\boldsymbol{v}_c, \boldsymbol{p}_c, \boldsymbol{q}_c\}$.

B. PINN Model Structure

In this section, we propose a customized neural network inspired by the PDNet-SDNets coupled PFlw model. The model structure is shown in Fig. 2. The PINN model comprises three modules: the physics-inspired module F_η, the linearized error compensation module F_e, and the voltage variance capture module F_v.

1) Physics-Inspired Module: As explained in Section II, the relationship between customer loads and voltages can be transformed into a linear relationship, combined with two complex implicit terms. The physics-inspired module, established on the linear feed-forward layers, is the core component that exhibits linear characteristics. It is important to note that the voltage terms are squared; hence, the data used in model training undergoes a similar squaring operation. This module incorporates three distinct neural layers - $\eta_{\mathrm{a}}^{\mathrm{p}}$, $\eta_{\mathrm{b}}^{\mathrm{q}}$, and $\eta_{\mathrm{c}}^{\mathrm{s}}$ - designed to simulate the linear part of (6). Apart from layer $\eta_{\mathrm{c}}^{\mathrm{s}}$, where the weight matrix $\boldsymbol{W_c}$ is an identity matrix, the parameters of the three layers make up the physical parameter set $\theta_\eta = \{\boldsymbol{W}_a, \boldsymbol{W}_b, \boldsymbol{b}_c\}$. In particular, $\boldsymbol{W}_a$ serves as the coefficient matrix for the active power matrix and primarily captures $\boldsymbol{E}$. On the other hand, $\boldsymbol{W}_b$ symbolizes the coefficient matrix for the reactive power matrix and is responsible for estimating $\boldsymbol{H}$. The model considers $-\boldsymbol{m}_c^s$ through $\boldsymbol{b}_c$. These coefficient matrices encapsulate the topology pattern and hold information about the line parameters. As a result, the parameters of the well-trained physics-inspired module can encapsulate network information, owing to its "structure-mimetic" design. The knowledge acquired by the coefficient matrices can provide the foundation for power network physics information awareness tasks, further explained in Section IV.

2) Linearized Error Compensation Module: This module serves the crucial role of mitigating the voltage calculation errors introduced by the linearized PFlw model, reflecting the $\boldsymbol{\chi}(\boldsymbol{p_c}, \boldsymbol{q_c}, \boldsymbol{v_c})$ term in (6). While previous works commonly neglect losses from power lines and STs, we recognize the significance of considering these losses to enhance the model's performance since the PDNet and SDNet combined network is considered. As discussed in Section II, voltage calculation errors are linked to the squared line power and squared voltage terms, which in turn, are associated with the quadratic expressions of customers' net active/reactive power consumption and the squared voltage. This association entails a complex relationship that is challenging to compute explicitly. Hence, our module employs fully connected MLPs with the tanh$(\cdot)$ activation function. The MLP module enables us to effectively model the intricate non-linear relationship between bus injection power and the error compensation for customer node voltages. Consequently, our model can more accurately compensate for voltage calculation errors. The inputs to this module consist of the squared customers' net active/reactive power consumption and squared voltage, while the outputs yield the error compensation for customer node voltages.

3) Voltage Variance Capture Module: This module serves to capture the voltage variance of the head bus voltage of each SDNet, represented as $\boldsymbol{\chi}(\boldsymbol{p_c}, \boldsymbol{q_c}, \boldsymbol{v_c})$ term. This voltage variance arises from changes in PFlw of the PDNet. The relationships between influencing factors (e.g., customer loads, voltage regulators) and voltage variance are complex, making explicit consideration challenging. To address this complexity, we employ the MLP model to effectively capture the nonlinear relationships. While our focus in this study is on fixed topologies, considering the topology modifications is potentially future work highlighted in Section V. Among the influential factors, customer loads and voltage regulators, notably on-load tap changers (OLTCs), are the key contributors. The actions of OLTCs closely align with the overall load conditions and the resultant PDNet voltage levels, which, in turn, depend on the load situations. To effectively capture this relationship, we utilize separate inputs for customer loads and total loads, representing the overall load conditions. The module outputs estimate the voltage variance at the head bus of each ST.

C. PINN Model Training Framework

To enhance the performance and accelerate the convergence of the PINN model, this paper employs customized training processes that account for the unique characteristics of the problem.

1) Data Normalization: Data normalization is an important pre-processing step when training deep neural networks, as it helps improve model convergence, reduce overfitting issues, and enhance generalization ability. Thus, we select the linear transformation method, specifically standardization, to accomplish this task [16].

2) Weight Initialization: Based on the designed neural network structure, two groups of weights need to be initialized that are $\theta_\eta = \{\boldsymbol{W}_a, \boldsymbol{W}_b, \boldsymbol{b}_c\}$ for the physics-inspired part and $\theta_\phi = \{\{\boldsymbol{W}_k^e\}_{k=1}^K, \{\boldsymbol{W}_l^v\}_{l=1}^L\}$ for other compensation parts, where $\boldsymbol{W}_k^e$ represents the kth layer in linearized error compensation module; the lth layer in voltage variance capture module is recorded as $\boldsymbol{W}_l^v$. According to the explanation in Section II, the $\boldsymbol{E}$ and $\boldsymbol{H}$ are non-positive symmetric matrices. To obtain better initial status and keep these properties, the $\boldsymbol{W}_a, \boldsymbol{W}_b$ are initialized as identity matrices I with the same size, that is, $\boldsymbol{W}_a^{init}, \boldsymbol{W}_b^{init} = -I_n \odot \boldsymbol{K}_{n\times n}$, where $\boldsymbol{K} \sim \mathcal{U}(0, 1)$. We initialize the $\boldsymbol{b}_c$ using random values yield to $\mathcal{U}(0, 1)$. To prevent the gradient from exploding or vanishing, we utilize the widely-used Xavier method for the initialization of other parameters θ_ϕ. Details of the methods can be found in [25].

3) Loss Function and Regularization: The loss function $\mathcal{L}(\cdot, \cdot)$ is a mathematical function that measures the difference between the predicted output of the neural network and the true output for a given input. In our problem, mean squared error (MSE) is used to measure the difference between predictive and actual voltage.

In addition to the typical error calculation components, regularization is another common element included in the loss function. Regularization is considered as penalty terms added to the loss function to impose soft constraints. In our problem, we employ the regularization method to encourage the network to retain physical information while updating to minimize loss. The designed loss function with regularization terms can be expressed as:

$$\mathcal{J}(\Theta) = L_\Theta + L_{\theta_\eta}^\eta + \mathcal{R}_\Theta, \tag{7}$$

$$L_\Theta = \frac{1}{N_b}\left(\sum_{i=1}^{N_b} \mathcal{L}\big(F\big(\boldsymbol{p}_i^{c_n}, \boldsymbol{q}_i^{c_n}; \Theta\big), \boldsymbol{v}_i^{c_n}\big)\right), \tag{8}$$

$$L_{\theta_\eta}^\eta = \frac{\delta}{N_b}\left(\sum_{i=1}^{N_b} \mathcal{L}\big(F_\eta\big(\boldsymbol{p}_i^{c_n}, \boldsymbol{q}_i^{c_n}; \theta_\eta\big), \boldsymbol{v}_i^{c_n}\big)\right), \tag{9}$$

$$\mathcal{R}_\Theta = \lambda \|\boldsymbol{W}_{\{a,b\}}\|_2 + \beta\left(\sum_i \sum_j \boldsymbol{W}_{\{a,b\}}^{i,j} - \|\boldsymbol{W}_{\{a,b\}}\|_1\right) + \gamma \|F_e^o\|_1, \tag{10}$$

where N_b is the batch size; δ is the scaling factor; $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the Manhattan Norm (MN) and Euclidean Norm (EN) respectively; λ, β and γ are the weighting factors for three regularization terms. The proposed loss function $\mathcal{J}(\Theta)$ incorporates three main components. F_e^o denotes the output of the linearization compensation module. First, L_Θ signifies the model prediction error calculated by MSE, forming the primary component of the loss function. Second, $L_{\theta_\eta}^\eta$ is a customized term that calculates the difference between the outputs of the physics-inspired module and actual voltages employing MSE. This term aims to reduce error compensation from other modules, thereby enhancing overall accuracy. Finally, the regularization term $\mathcal{R}_\Theta$ is included in the loss function. The EN of $\boldsymbol{W}_{\{a,b\}}$ is adopted in $\mathcal{R}_\Theta$ to make the weight matrices sparse, reflecting the characteristics of real $\boldsymbol{E}$ and $\boldsymbol{H}$. To maintain $\boldsymbol{W}_{\{a,b\}}$ as non-positive, we introduce the subtraction between the element summation of $\boldsymbol{W}_{\{a,b\}}$ and MN as soft constraints. Similar to $L_{\theta_\eta}^\eta$, we supplement the MN of F_e in the regularization terms to minimize error compensation, as the actual linearized error cannot be large.

Algorithm 1 PINN Model Training Algorithm

Require: Training set $\mathcal{D}_{tr} = \{\boldsymbol{v}_c, \boldsymbol{p}_c, \boldsymbol{q}_c\}$, initial learning rate (LR) α_0, decay factor k, momentum ζ, mini-batch size N_b, number of epochs $\mathcal{T}$

Initialize the parameters of network $F_{\boldsymbol{\theta}}$ as $\Theta = \{\theta_\eta^0, \theta_\phi^0\}$ by designed rules; update initial LR as $\alpha \leftarrow \alpha_0$
for *epoch* $= 1$ to $\mathcal{T}$ **do**
  for $i = 1$ to $\lceil N/N_b \rceil$ **do**
    Select N_b example pairs from shuffled $\mathcal{D}_{tr}$ forming mini-batch $S_i = \{\boldsymbol{p}_b^{c_n}, \boldsymbol{q}_b^{c_n}, \boldsymbol{v}_b^{c_n}\}_{b=1}^{N_b}$
    Compute gradient of the loss function with respect to network parameters as
    $$\nabla_\theta \mathcal{J}(\Theta; S_i) = \{\nabla_{\theta_\eta}\mathcal{J}, \nabla_{\theta_\phi}\mathcal{J}\}$$
    Editing gradient of physics-inspired module based on weight symmetry averaging as
    $$\nabla_{\theta_\eta}\mathcal{J} \leftarrow \tfrac{1}{2}\big(\nabla_{\theta_\eta}\mathcal{J} + \nabla_{\theta_\eta}\mathcal{J}^T\big)$$
    Update the parameters using SGD update rule:
    $$\bar{\boldsymbol{v}} \leftarrow \zeta \boldsymbol{v} + (1-\zeta)\nabla_\theta \mathcal{J}(\Theta; S_i)$$
    $$\Theta \leftarrow \Theta - \alpha \bar{\boldsymbol{v}} \qquad \triangleright\ v \leftarrow \nabla_\theta \mathcal{J}(\Theta; S_{i-1}).$$
    if $\lceil \alpha/e \rceil == 0$ **then**
      $\alpha \leftarrow k\alpha$ $\triangleright$ decays LR α by k every e epochs
    end if
  end for
end for
return F_{θ_f}

4) Gradient Editing: Due to the properties of the $\boldsymbol{E}$ and $\boldsymbol{H}$, it is crucial to maintain the symmetry of the weight matrices $\boldsymbol{W}_{\{a,b\}}$ during network training to achieve better performance. Considering the symmetrical initialization weights, one straightforward approach is to enforce weight symmetry by replacing the gradient of the weight matrix $\boldsymbol{W}_{\{a,b\}}$ with the average of the gradient and its transpose during the backpropagation phase. This technique is known as the weight symmetry averaging. After considering all the steps outlined previously, we utilize Stochastic Gradient Descent (SGD), a widely used optimization technique, as the optimizer for updating the model parameters. The training procedure for the PINN model is provided in **Algorithm 1**.

IV. Applications of PINN-Based Voltage Calculation Model

A. Model-Free Locational PV Hosting Capacity Calculation

To ensure the seamless integration of new PV installations, it is essential to conduct the locational PV HC analysis [5], [26]. This analysis helps to determine the maximum PV capacity that can be accommodated within the grid without violating operational constraints at specific locations or necessitating grid upgrades. The HC analysis considers various impact criteria, such as system overvoltage, thermal stress, harmonics, etc. Its primary focus is to uphold good voltage quality, particularly for typical North American residential circuits [27]. Estimating PV HC based on voltage constraints requires accurate voltage estimation in new scenarios, such as reverse PFlw or large voltage fluctuations. This underlines the paramount significance of extrapolation capabilities. Our designed model demonstrates excellent potential extrapolation capabilities due to the special structure, making it suitable for calculating voltages in high-penetration PV scenarios. As a result, we conducted locational HC analysis to show the potential application of our proposed model.

B. PDNet-SDNets Physics Information Awareness

The lack of detailed SDNet models impedes effective decision-making and planning for operators. To tackle this challenge, earlier research efforts have delved into TC relationship identification [28], [29]. However, the predominant reliance on voltage correlation combined with manual parameter adjustments hinders existing methods from achieving consistent and stable performance. Our proposed model, featuring a well-designed physics-inspired module, offers novel perspectives on solving TC connectivity problems. To demonstrate the model's support for physics information awareness, we developed a method for identifying TC connectivity. This method leverages the abundant physical information contained in $\boldsymbol{W}_a$ and $\boldsymbol{W}_b$. The procedure for connectivity identification is detailed in **Algorithm 2**.

Initially, the algorithm transforms the $\boldsymbol{W}_a$ and $\boldsymbol{W}_b$ into the minimum connection matrix $\boldsymbol{G}$, adhering to the threshold τ, which has been proposed in **Algorithm 2** and proof to be the lower bound of non-zero[2] elements in $\boldsymbol{W}_a$ or $\boldsymbol{W}_b$. $\boldsymbol{G}$ only contains partial customer connection information; detailed below, if the element $\boldsymbol{G}^{i,j}$ is non-zero, customer i and j should be connected to the same ST, but the opposite is not true because only the minimum connection number is considered to generate $\boldsymbol{G}$. Hence, the algorithm then applies the "transitive relation" rule to augment $\boldsymbol{G}$. For instance, if customers i and j, and customers j and d, are respectively connected to the same ST, then customers i, j, and d are considered as linked to the same ST. Based on the modified $\boldsymbol{G}$, customers connected to the same ST form a cluster, and all such clusters constitute a cluster list $\mathcal{C}$. The algorithm subsequently and iteratively merges the clusters in $\mathcal{C}$, after discarding duplicate items, based on the correlation between two clusters until the number of clusters matches the ST counts k. The cluster relationship $RV = \rho(z_s, z_t)$, where z_s and z_t are two clusters from $\mathcal{C}$, can be calculated as $RV = \sum_{i=1}^{|z_s|}\sum_{j=1}^{|z_t|}(|\boldsymbol{W}_a^{z_s^i}| + |\boldsymbol{W}_b^{z_t^j}|)$. A higher RV value indicates that customers from the two clusters are likely to be connected to the same transformer, suggesting they should be merged. The final TC results are recorded in $\mathcal{C}$. Utilizing this straightforward method, we can extract TC information from the well-trained PINN model.

[2]Training errors may result in sparse elements in $\boldsymbol{W}_a$ and $\boldsymbol{W}_b$ being small but not exactly zero. We still refer to these elements as "zero elements" for convenience and the others as "non-zero elements." This approximation does not affect the final results.

Algorithm 2 TC Connectivity Identification

Require: $\boldsymbol{W}_a$, $\boldsymbol{W}_b$, Customer Num N_c, Transformer Num k

Calculate threshold index $\tau \leftarrow \lfloor \frac{N_c^2}{k} \rfloor$

Update $\boldsymbol{W}_a$, $\boldsymbol{W}_b$ as:

$$\boldsymbol{W}_a^{i,j} \geq \boldsymbol{W}_a^{[\tau]} \leftarrow 1;\ \boldsymbol{W}_a^{i,j} < \boldsymbol{W}_a^{[\tau]} \leftarrow 0;$$
$$\boldsymbol{W}_b^{i,j} \geq \boldsymbol{W}_b^{[\tau]} \leftarrow 1;\ \boldsymbol{W}_b^{i,j} < \boldsymbol{W}_b^{[\tau]} \leftarrow 0;$$
$$\boldsymbol{G}^{i,j} \leftarrow [\![(\boldsymbol{W}_a + \boldsymbol{W}_b)_{i,j} > 0]\!] i, j = 1, 2, \ldots, N_c;$$

$\boldsymbol{W}^{[\tau]}$ denotes the τ largest element of $\boldsymbol{W}$.

for $i = 1$ to N_c **do**

  Create initial set $R = \{j | G_{i,j} == 1, j \in 1, \ldots, N_c\}$ $CS \leftarrow R$, $FS \leftarrow R$

  For every item m from R, conduct update below until $|FS|$ equals to $|CS|$:

$$FS \leftarrow FS \cup \{j | G_{m,j} == 1\}$$
$$CS \leftarrow FS$$

  Add FS to the cluster list $\mathcal{C}$, and remove duplicates

end for

while $|\mathcal{C}| \geq k$ **do**

  Calculate $RV = \rho(z_s, z_t)$, $s, t \in \{1, \ldots, |\mathcal{C}|\}$

  Find minimum value RV_{z_s,z_t}, then merge z_s, z_t two sets and update $\mathcal{C}$

  Recalculate $RV = \rho(z_s, z_t)$, $s, t \in \{1, \ldots, |\mathcal{C}| - 1\}$

end while

return $\mathcal{C}$

Proposition 1. The lower bound for the number of non-zero elements in matrix $\boldsymbol{W}_a$ or $\boldsymbol{W}_b$ is greater than τ, where $\tau = \lfloor \frac{N_c^2}{k} \rfloor$; k and N_c denote ST number and total customer number, respectively.

Proof: When customer i and j share the same ST, the element $\boldsymbol{W}_a^{i,j}$ and $\boldsymbol{W}_b^{i,j}$ will be non-zero. We define $\boldsymbol{x} \in \mathbb{Z}^k$ as the number of customers connected to each ST. The problem of finding the minimum number of non-zero elements in matrices can be formulated as $\min y = \boldsymbol{x}^T\boldsymbol{x}$, subject to the constraint $\sum_{i=1}^{k} x_i = N_c$. To solve the problem, we relax $\boldsymbol{x}$ to $\bar{\boldsymbol{x}} \in \mathbb{R}^k$ and obtain the objective function as $\bar{y}$, yielding $\min\ \bar{y} \leq \min\ y$. Notably, the relaxed problem achieves its optimal solution when each ST has an equal number of customers. The optimal value of objective function $\bar{y}^*$ in this case is $\frac{N_c^2}{k}$. To satisfy the integer requirement, we can round this value down to $\lfloor \frac{N_c^2}{k} \rfloor$, which preserves the relationship that $\lfloor \bar{y}^* \rfloor \leq \bar{y}^* \leq y^*$, where y^* denotes the optimal value of original problem. The proposition is thus proven.

V. Case Studies

A. Test Circuits and SM Datasets

Three distribution feeder models are used for conducting the designed case studies, comprising two public testing circuits, namely, "EPRI12Bus" (small) and "EPRICk5" (complex) circuits, along with one real utility feeder. Each model integrates STs and SDNets. The small circuit serves 46 customers spread over 12 unique low-voltage SDNets, each boasting distinct topologies and conductor lengths [18]. The complex circuit is modeled after EPRI Ckt5 and includes 591 STs connected with 1379 customers [30]. The real feeder circuit, marked as "Real40Bus", originates from a distribution network in the Midwest U.S., powered by a 69 kV substation. In contrast to the small test circuit, the real utility feeder model features an extended three-phase feeder line with 40 STs connected with 52 customers. Moreover, each customer across the three test circuits was allocated a unique load profile with real and reactive power derived from actual utility smart meter data, with a data resolution of 30 minutes over two years. Utilizing authentic smart meter data, voltage values are produced through OpenDSS based on the corresponding distribution systems.

B. Voltage Calculation

1) Simulation Scenario Generation: We tested our proposed model through five scenarios, denoted as S1 to S5 in Table I. The PV load data are sourced from over 300 solar inverters with 4-10 kW capacities in the Middle U.S. The EV data, culled from various real datasets, had charging capacities of 3-10 kW. During scenario generation, annual PV curves and EV charging profiles are randomly sampled from these datasets and added to customer load curves. In S1, we fully trained and tested the model on historical data without additional PV or EV loads, assessing its performance under normal conditions. S2 introduced PV for 25% of customers in both training and testing data. This scenario tested the model's performance under increased voltage variations caused by fluctuating PV generation. In S3, S4, and S5, the datasets included various PV and EV penetration levels, while the training data remained historical data as in S1. These experiments evaluated the model's extrapolation capability under "unseen" scenarios. Given that our model incorporates SM data as inputs, the analysis of the effects of measurement noise and synchronization discrepancies is conducted to ensure model robustness. The deviations in SM data comprise two primary components. The first component, measurement noise, has been extensively investigated. Research indicates that it generally adheres to a Gaussian distribution with a zero mean and a specific standard deviation [31]. The second component stems from the asynchronous nature of smart meters, which also exhibits a normal distribution as suggested in [31]. Consequently, we can model the overall error as a composite of two normally distributed variables, which inherently results in a normal distribution. Aligning with prior studies [31], [32], and [33], we adopted a deviation level of 5, signifying that measurements are within ±5% of the actual values. To simulate a more realistic dataset, Gaussian noise masks were applied to the loading data. The deviation σ of the setting is given by $\sigma = dl^{*}|z^{\mathrm{m}}|/3$, where dl is the deviation level; z^{m} represents the loading data measured from the SMs. Therefore, we configured the dl to 5% with a mean of 0 for loading data, and to 1% with a mean of 0 for voltage data. This setup ensures that our proposed model undergoes testing with data that closely mimics real-world conditions.

TABLE I
SIMULATION SCENARIO GENERATION SETTING

Scenario	Training	Testing	PV Penetration(%) EPRI12Bus/Real40Bus/EPRICk5
S1	basic	basic	0%/ 0%/ 0%
S2	25%PV	25%PV	39%/ 56%/ 57%
S3	basic	25%PV	39%/ 56%/ 57%
S4	basic	50%PV	114%/ 108%/ 93%
S5	basic	50%PV + 20%EV	114%/ 108%/ 93%

2) Results Analysis: The voltage calculation tasks on the five scenarios are carried out by three models, including PINN, linear NN (LNN), and Deep neural network (DNN). The LNN model is the PINN model without the two compensation modules. DNN refers to the fully connected neural network. All the models are built on one-year SM data, among which 80% data for training and 20% data for validating, and then tested on one-year long data. The error of the voltage calculation results are shown in Fig. 3 and Fig. 4.

The bar chart in Fig. 3, illustrating the mean absolute error (MAE) values over all time points and customers, shows the PINN models exhibit lower MAE values than the DNN model across all scenarios. As PV and EV penetration levels increase, the MAE differences between the DNN and other models become more prominent. In scenarios S1 and S2, where no unseen cases are present in the test set, the DNN model shows excellent results, with accuracy roughly consistent with the PINN and even better than LNN in large systems. However, when data from new scenarios, e.g., high DER integration, are included in the testing set, the error of the DNN model significantly increases, reaching higher levels. In contrast, the PINN and LNN models continue to perform well, showcasing their strong extrapolation ability. Overall, the PINN and LNN models perform well across most scenarios. When the testing model is small (e.g., EPRI12Bus and Real40Bus), the accuracy of the two models is similar. However, in the EPRICk5 model, where the PFlw relationships of PDNet become more complex due to a larger number of buses, the LNN model struggles to capture these complexities, resulting in increased MAE errors. Conversely, including compensation modules in the PINN model enhances its performance, particularly in complex scenarios where PFlw relationships are intricate. We can also see from Fig. 5 that the differences between PINN and LNN usually occurred on the tip points where voltage regulators could act. The blue error line above each bar in Fig. 3 represents the MAE values obtained when the models are trained and tested with the noisy data. Notably, the proposed model maintains robust performance, even when accounting for potential variations stemming from measurement inaccuracies and synchronization discrepancies commonly present in real-world SM measurements. The boxplots in Fig. 4 further clarify these findings by illustrating the error distributions during

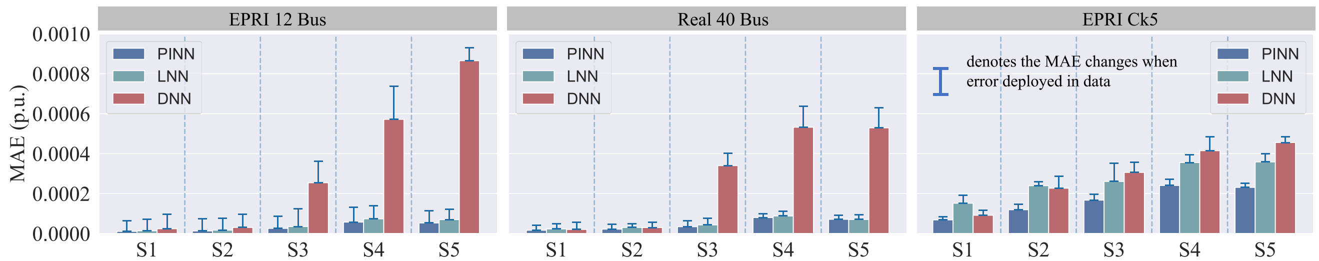


Fig. 3. The MAE of three models over different scenarios based on accurate data and noise-added data.

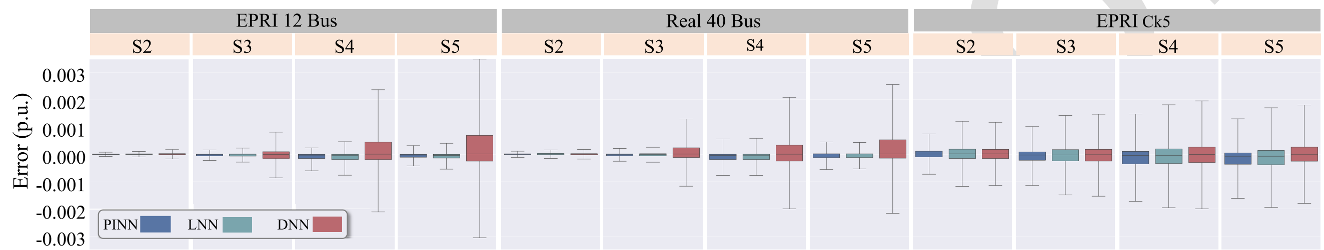


Fig. 4. The error (actual value minus predicted value) distribution of three models over different scenarios during daytime (6 a.m. to 6 p.m.) period.

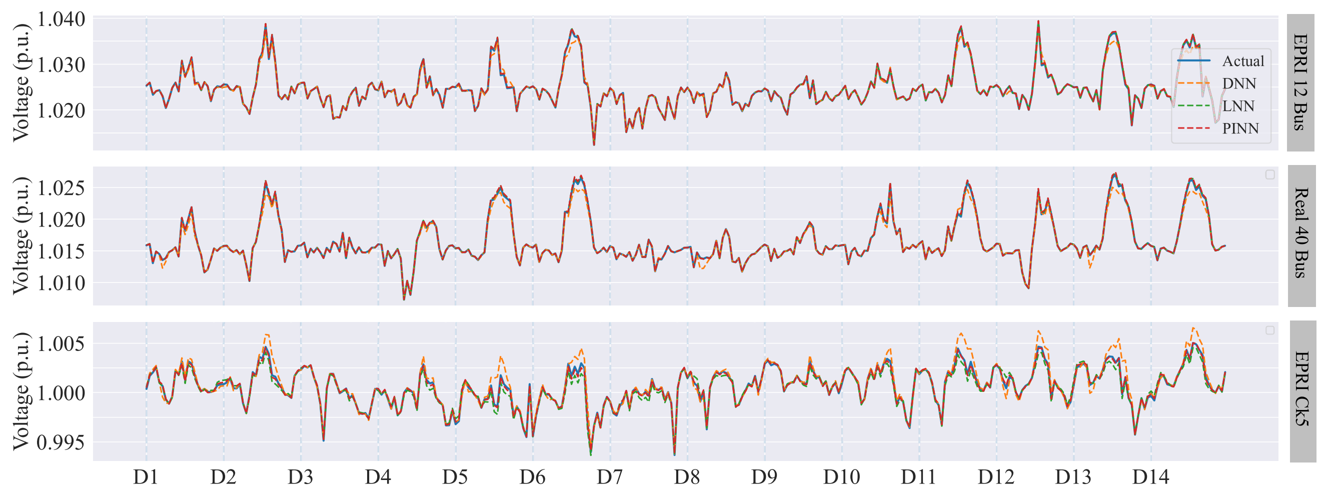


Fig. 5. Voltage estimation results for three customers from all circuits in S5.

the 6 a.m. to 6 p.m. daytime period, where PV generations have the largest impacts. These visualizations underline that, compared to the DNN model, the errors of PINN results are more concentrated, and such differences are notably prominent in the EPRI12Bus and Real40Bus because of the higher PV penetration level of these two models in the scenarios S4 and S5.

Fig. 6 shows the training results for $\boldsymbol{W}_a$ and $\boldsymbol{W}_b$ across all test circuits. Physical connections between customers are evident from the significant values (darker colors) in the plots, indicating links between customers. Customers connected to the same transformer exhibit pronounced voltage correlation, forming darker sub-squares in the plots. Notably, the structured connectivity in these plots is influenced by the customer order in the input data, which is inaccessible in real scenarios, resulting in more randomized matrices. Additionally, weak correlations between some customers and potential training errors may hinder extracting physical information. Hence, this paper proposes a TC identification algorithm to address these issues, with detailed testing results presented later.

3) Assessing the Impacts of Training Dataset Durations: In practical applications of the PINN model, the available training data volume may not be as extensive as in simulation scenarios. For instance, the addition of new customers to the system will result in limited smart meter data. Additionally, the smart meter data missing will also lead to the training dataset shrink. Consequently, understanding the minimal training dataset size required to maintain model efficacy is crucial under these circumstances. To explore this, several simulations are carried out, training the PINN model with datasets spanning one year, six months, three months, one month, and one week. We assessed the models' effectiveness using simulated data from "S5". Fig. 7 illustrates the average MAE in voltage estimations for models trained across these durations.

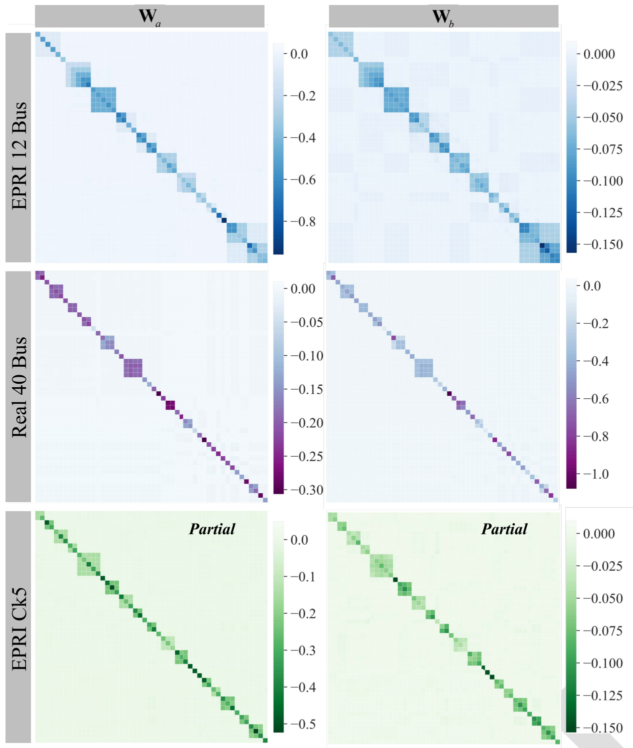


Fig. 6. The training results of $\boldsymbol{W}_a$ and $\boldsymbol{W}_b$ across all datasets.

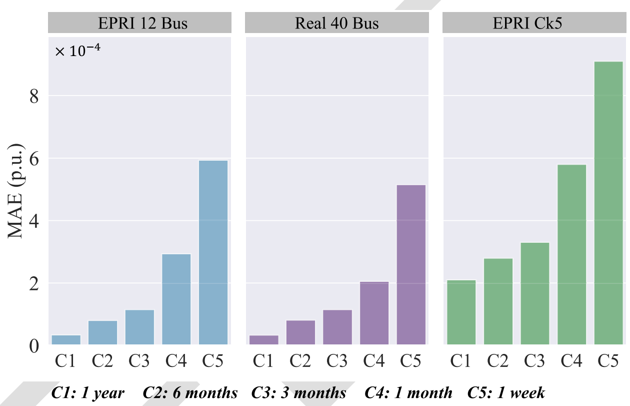


Fig. 7. Average MAE of voltage estimations across PINN models trained with datasets of varying time spans.

Fig. 7 clearly illustrates that the MAE of the models escalates as the duration of the training datasets diminishes, transitioning from a year to a week. Specifically, the ERPI12Bus and Real40Bus models exhibit a marginal rise in MAE when the dataset length is curtailed from one year to three months. Although training with one month's data leads to a notable error increase, they are still albeit within acceptable limits. However, the scenario changes drastically under the training of the one-week dataset, where the MAE surges significantly, indicating the model's diminished capacity to discern the underlying PFlw relationships. The scale of the challenge is more pronounced in the PINN model applied to the EPRICk5, due to its more extensive scale (requiring the training of more parameters). A pronounced jump in MAE is observed when the dataset is limited to one month, suggesting that larger distribution systems necessitate more extensive datasets. It's important to note that these observations are based on 30-minute interval smart meter data. Increasing the granularity of the data to 15-minute intervals could potentially reduce the minimum dataset size required for effective model training. Preliminary findings suggest that for smaller systems, a dataset spanning two weeks may suffice, while larger systems, akin to the EPRICk5, may necessitate a dataset ranging from two weeks to a month.

4) Analysis of Model Retraining Timings: When integrating the proposed model into actual utility systems, maintaining its updated state is crucial for precise voltage estimation. Consequently, model retraining becomes indispensable. This section delineates three triggers for initiating model retraining: error-oriented, event-oriented, and manual intervention. For the error-oriented trigger mechanism, the system operators establish specific error thresholds that are thoughtfully tailored to the unique demands of each distribution system. This customization is crucial to ensure that the model's sensitivity to errors is appropriately calibrated for each system's diverse conditions. When new smart meter data is fed into the model for validation, the model is flagged for retraining if the voltage calculation errors surpass these predetermined thresholds. From the moment these errors are detected, the newly collected smart meter data are gathered and employed as training data for the PINN. In the event-oriented approach, training can be proactively initiated even when the voltage estimation errors by the PINN remain within acceptable limits. This approach is triggered by specific events, which may not necessarily cause immediate errors but also need an update in the model. Key factors that activate this event-oriented trigger include the scheduled changes within the network or the onset of unique operational scenarios. For instance, the peak load scenarios caused by extreme weather conditions (e.g., extremely high/low temperatures) are typically underrepresented in historical datasets. Including data from these unique scenarios enhances model accuracy, as a more diverse training set improves the model's performance in different conditions. Additionally, the proposed model can incorporate the manual setting option for retraining, an essential feature to maintain its relevance and accuracy over time. This approach involves periodically (e.g., weekly, monthly) reviewing and updating the model, regardless of whether it has reached a specific error threshold or encountered a notable event. While this approach may entail a higher computational burden, it is crucial for keeping the model current. Moreover, once the system reaches the error- or event-oriented trigger, the model will be updated again, following the respective retraining protocols.

5) Discussion: Relying on the derived coupled PFlw model, the PINN aims to imbue each module with physical significance. The linear neural network portion replicates the linear component of (5) via parameter estimation, exemplifying an accurate modeling approach that strictly adheres to the physics rules of the system. This approach allows for precise estimation across various scenarios, regardless of whether the data exists in the historical dataset. The non-linear elements, on the other hand, are indirectly captured by the MLPs, leveraging their exceptional non-linear mapping capabilities. However, several potential issues warrant discussion. Firstly, cross-compensation of errors may exist among the three modules during training. Given the lack of model information and the absence of measurements from PDNet, SMs are the only viable data source. Through carefully designed regularization terms, we strive to prevent such compensation. While complete eradication may not be possible, our results demonstrate the effectiveness of the voltage variance capturing module, evidenced by comparing PINN and LNN results. Another challenge relates to topology modifications. Changes in the PDNet topology can impact the whole PFlw, thereby compromising calculation accuracy. Current strategies involve retraining the entire model using new data to tackle this problem. However, the proposed PINN model takes a more efficient approach. It retains unchanging parameters such as $\boldsymbol{W}_a$ and $\boldsymbol{W}_b$, and fine-tunes the remaining model components, thus reducing the need for extensive training data and computational capacity. This strategy can be regarded as genuine transfer learning, a machine learning technique where the model developed for a specific task is adapted for a second related task. Additionally, if certain system information is partially available, we can employ a masking mechanism to reduce the number of training parameters, thereby accelerating and enhancing model convergence. Future work will focus on exploring the integration of known system information and addressing topology modification.

In the real-world deployment of models within utility systems, navigating the issue of smart meter data missing is crucial. There are three predominant scenarios of missing the model could face. Firstly, when individual customers experience a short range of data missing, we could address the problem by removing the load data for all customers during those specific intervals, leveraging the fact that our model's input doesn't necessitate continuous data, thereby ensuring minor omissions would not significantly affect the model's accuracy. Secondly, a more challenging scenario arises when a significant portion of data is missing across many customers, leading to a limited dataset for training. Regarding this issue, our previous analysis shows that the model can still yield acceptable results with around one month of complete historical data, indicating a certain resilience to this data missing problem. The third scenario involves extensive data gaps concentrated among a few customers. In such cases, using advanced training methods like transfer learning on an existing, outdated model can help minimize the requirement for large volumes of training data and lessen the impact of these data gaps. These strategies, aimed at mitigating the effects of missing data, are pivotal areas of focus in our upcoming research, offering potential solutions to enhance model reliability in real-world applications.

C. Locational Hosting Capacity Estimation

To exhibit how the designed model performs in the calculation of locational HC, the Real40Bus model is selected to complete the test. Instead of analyzing just a handful of worst-case scenarios and obtaining one PV HC value, the proposed voltage calculation model calculates the maximum accessible PV power at every time point for each customer location. In this context, the locational HC can be regarded as the minimum value of the maximum accessible PV power across all time points. However, our discussion here is confined to the maximum accessible PV power. To generally exhibit the performance of our model, the MAE and the mean absolute percentage error (MAPE) of the estimation results over all the time points are discussed. The model-based algorithm that uses quasi-static time series simulations is adopted to be the benchmark to calculate the maximum accessible PV power, with more details provided in [5]. The estimation error obtained from the designed model is shown in Fig. 8. Each bar exhibits the average MAE of maximum accessible PV power for one customer over one-year time points, and the green curve presents the MAPE of corresponding estimation results. It can be seen that the average MAE error for each customer remains in a small range, with the maximum error below 3.5 kW. The average error over all the customers is just 0.87kW, and the MAPE averages below 2.5%. Compared to previous locational HC work, the performance of the proposed model is competitive [30].

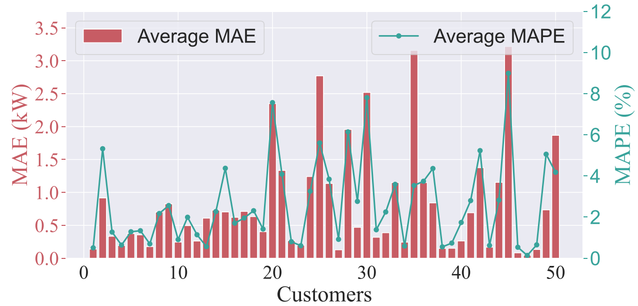


Fig. 8. Average MAE and MAPE of maximum accessible PV power for all customers.

D. Power Network Physics Information Awareness

Previous research primarily focuses on TC identification based on voltage correlation among customers [29], considering only voltage information. On the contrary, our proposed method leverages the knowledge learned by the physics-inspired module that is well-trained using $\mathcal{D}_{tr}$ and integrates both load and voltage data as input. By incorporating additional information, our method offers a higher capability for TC identification. We tested the designed algorithm on three distribution models, and the results are presented in Table II, where the accuracy metric equals the ratio of the accurately identified ST number to the total ST number. As shown, our method achieves excellent results with 100% accuracy, whether in the designed system with diverse secondary topologies (i.e., EPRI 12 bus) or in a real utility model. This indicates its effectiveness in handling complex SDNet patterns and real-world conditions. Furthermore, the favorable test results in large distribution networks (i.e., EPRI Ck5) validate the scalability of our approach.

VI. Conclusion

This paper introduced an electric model-free voltage calculation methodology designed to accommodate the operational

TABLE II
TC CONNECTIVITY IDENTIFICATION RESULTS

Model	EPRI 12 Bus	Real 40 Bus	EPRI Ck5
Transformer Number	12	40	591
Customer Number	46	50	1379
Correctly Identified	12	40	575
Accuracy Rate	100%	100%	97.3%

and planning needs of distribution networks without the necessity for accurate electrical models. Leveraging the structure inspired by the PDNet-SDNets coupled PFlw, the PINN model displays potential for extrapolation and exhibits the ability to capture the physical characteristics of the electrical network. Supported by a customized training framework, the model ensures convergence and robust performance. Evaluations using two public testing systems and a real utility feeder model affirmed the effectiveness of the model in voltage calculation. The testing results also corroborated the proposed model's extrapolation and physical awareness capabilities in locational HC and TC identification applications. Future work will explore integrating known system information and assess the model's adaptability to topology modification. Additionally, efforts will be directed toward enhancing the model to support both three-phase and two-phase loads, thereby bolstering the applicability and accuracy of the PINN.

ACKNOWLEDGMENT

The employee owns all right, title and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan https://www.energy.gov/downloads/doe-public-access-plan.

REFERENCES

[1] *ANSI C84.1-2020: American National Standard for Electric Power Systems and Equipment—Voltage Ratings (60 Hz)*, Am. Nat. Stand. Inst., Washington, DC, USA, 2020.

[2] Z. Ma, Q. Zhang, and Z. Wang, "Safe and stable secondary voltage control of microgrids based on explicit neural networks," *IEEE Trans. Smart Grid*, vol. 14, no. 5, pp. 3375–3387, Sep. 2023.

[3] R. Cheng, N. Shi, S. Maharjan, and Z. Wang, "Automatic self-adaptive local voltage control under limited reactive power," *IEEE Trans. Smart Grid*, vol. 14, no. 4, pp. 2851–2862, Jul. 2023.

[4] L. Blakely, M. J. Reno, and J. Peppanen, "Identifying common errors in distribution system models," in *Proc. IEEE 46th Photovoltaic Specialists Conf. (PVSC)*, 2019, pp. 3132–3139.

[5] J. A. Azzolini et al., "Improving behind-the-Meter PV impact studies with data-driven modeling and analysis," in *Proc. IEEE Photovoltaic Specialists Conf. (PVSC)*, 2022, pp. 204–204.

[6] S. Balduin, T. Westermann, and E. Puiutta, "Evaluating different machine learning techniques as surrogate for low voltage grids," *Energy Inform.*, vol. 3, no. 1, pp. 1–12, 2020.

[7] Y. Liu, N. Zhang, Y. Wang, J. Yang, and C. Kang, "Data-driven power flow linearization: A regression approach," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2569–2580, May 2019.

[8] J. Chen, W. Li, W. Wu, T. Zhu, Z. Wang, and C. Zhao, "Robust data-driven linearization for distribution three-phase power flow," in *Proc. IEEE 4th Conf. Energy Internet Syst. Integr. (EI2)*, 2020, pp. 1527–1532.

[9] 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. Gener. Meet. (PESGM)*, 2020, pp. 1–5.

[10] Z. Yu and Y.-Q. Bao, "Data-driven power flow calculation based on deep learning," in *Proc. IEEE 5th Conf. Energy Internet Syst. Integr. (EI2)*, 2021, pp. 3230–3234.

[11] Y. Chen, Y. Shi, and B. Zhang, "Data-driven optimal voltage regulation using input convex neural networks," *Electr. Power Syst. Res.*, vol. 189, Dec. 2020, Art. no. 106741.

[12] X. Hu, H. Hu, S. Verma, and Z.-L. Zhang, "Physics-guided deep neural networks for power flow analysis," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2082–2092, May 2021.

[13] V. Bassi, L. F. Ochoa, T. Alpcan, and C. Leckie, "Electrical model-free voltage calculations using neural networks and smart meter data," *IEEE Trans. Smart Grid*, vol. 14, no. 4, pp. 3271–3282, Jul. 2023.

[14] A. Simonovska, V. Bassi, A. G. Givisiez, L. F. Ochoa, and T. Alpcan, "An electrical model-free optimal power flow for PV-rich low voltage distribution networks," in *Proc. IEEE PES Innovat. Smart Grid Technol. Conf. Europe (ISGT-Europe)*, 2022, pp. 1–5.

[15] L. Zhang, G. Wang, and G. B. Giannakis, "Real-time power system state estimation and forecasting via deep unrolled neural networks," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 4069–4077, Aug. 2019.

[16] Y. Yang, Z. Yang, J. Yu, B. Zhang, Y. Zhang, and H. Yu, "Fast calculation of probabilistic power flow: A model-based deep learning approach," *IEEE Trans Smart Grid*, vol. 11, no. 3, pp. 2235–2244, May 2020.

[17] X. Pan, "DeepOPF: Deep neural networks for optimal power flow," in *Proc. 8th ACM Int. Conf. Syst. Energy Efficient Build., Cities, Transp.*, 2021, pp. 250–251.

[18] J. Yusuf, J. A. Azzolini, and M. J. Reno, "Predicting voltage changes in low-voltage secondary networks using deep neural networks," in *Proc. IEEE Power Energy Conf. Illinois (PECI)*, 2023, pp. 1–8.

[19] K. Xu, M. Zhang, J. Li, S. S. Du, K.-i. Kawarabayashi, and S. Jegelka, "How neural networks extrapolate: From feedforward to graph neural networks," 2020, *arXiv:2009.11848*.

[20] 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, Jan. 2023.

[21] L. Pagnier and M. Chertkov, "Embedding power flow into machine learning for parameter and state estimation," 2021, *arXiv:2103.14251*.

[22] B. Donon, B. Donnot, I. Guyon, and A. Marot, "Graph neural solver for power systems," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2019, pp. 1–8.

[23] L. Gan and S. H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial networks," in *Proc. Power Syst. Comput. Conf.*, 2014, pp. 1–9.

[24] H. Zhu and H. J. Liu, "Fast local voltage control under limited reactive power: Optimality and stability analysis," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3794–3803, Sep. 2016.

[25] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.

[26] S. Talkington, S. Grijalva, M. J. Reno, J. A. Azzolini, and D. Pinney, "A measurement-based approach to voltage-constrained hosting capacity analysis with controllable reactive power behind-the-meter," *Electr. Power Syst. Res.*, vol. 221, Aug. 2023, Art. no. 109395.

[27] A. Dubey, S. Santoso, and A. Maitra, "Understanding photovoltaic hosting capacity of distribution circuits," in *Proc. IEEE Power Energy Soc. General Meet.*, 2015, pp. 1–5.

[28] J. Zhao, M. Xu, X. Wang, J. Zhu, Y. Xuan, and Z. Sun, "Data-driven based low-voltage distribution system transformer-customer relationship identification," *IEEE Trans. Power Deliv.*, vol. 37, no. 4, pp. 2966–2977, Aug. 2022.

[29] W. Luan, J. Peng, M. Maras, J. Lo, and B. Harapnuk, "Smart meter data analytics for distribution network connectivity verification," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1964–1971, Jul. 2015.

[30] J. A. Azzolini, M. J. Reno, J. Yusuf, S. Talkington, and S. Grijalva, "Calculating PV hosting capacity in low-voltage secondary networks using only smart meter data," in *Proc. IEEE Power Energy Soc. Innovat. Smart Grid Technol. Conf. (ISGT)*, 2023, pp. 1–5.

[31] A. Alimardani, F. Therrien, D. Atanackovic, J. Jatskevich, and E. Vaahedi, “Distribution system state estimation based on nonsynchronized smart meters,” *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2919–2928, Nov. 2015.

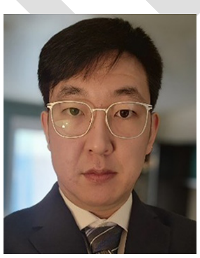
[32] J. A. Massignan, J. B. London, M. Bessani, C. D. Maciel, R. Z. Fannucchi, and V. Miranda, “Bayesian inference approach for information fusion in distribution system state estimation,” *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 526–540, Jan. 2022.

[33] Q. Chen, D. Kaleshi, and Z. Fan, “Inferring low voltage transformer state using only smart metering data,” in *Proc. IEEE PES ISGT Eur.*, 2013, pp. 1–5.

Liming Liu (Graduate Student Member, IEEE) received the B.S. and M.S. degrees in electrical engineering from North China Electric Power University in 2016 and 2019, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Iowa State University. His research interests encompass power distribution systems, distribution system modeling, and application of optimization and machine learning techniques to power systems.

Naihao Shi (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from North China Electric Power University in 2017 and the M.S. degree in electrical engineering from George Washington University in 2020. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA. His research interests include distribution system modeling, voltage/var control, and applications of optimization in power systems.

Dingwei Wang (Graduate Student Member, IEEE) received the B.S. degree in electrical and computer engineering from Marquette University, Milwaukee, WI, in 2018 and the M.S. degree in electrical and computer engineering from George Washington University, Washington, DC, in 2020. He is currently pursuing the Ph.D. degree with Iowa State University. His research interests include distribution system resilience, data analytics, and applications of machine learning techniques to power systems.



Zixiao Ma (Member, IEEE) received the B.S. degree in automation and the M.S. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 2014 and 2017, respectively, and the Ph.D. degree in electrical and computer engineering from Iowa State University, Ames, IA, USA, in 2023. He is currently a Distinguished Postdoctoral Fellow with the Clean Energy Institute, Department of Electrical and Computer Engineering, University of Washington, Seattle, WA, USA. His research interests focus on control theory and machine learning with their applications to inverter-based resources, microgrids, and load modeling. He was the recipient of the Outstanding Reviewer Award from IEEE Transactions on Power Systems, the Research Excellence Award from Iowa State University, the Chinese Government Award for Outstanding Self-Financed Students Abroad, the Distinguished Postdoctoral Fellowship from the University of Washington, and the Rising Stars Award in Cyber-Physical Systems from the University of Virginia.

Zhaoyu Wang (Senior Member, IEEE) received the B.S. and first M.S. degrees in electrical engineering from Shanghai Jiao Tong University, and the second M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology. He is the Northrop Grumman Endowed Associate Professor with Iowa State University. He is the Principal Investigator for a multitude of projects funded by the National Science Foundation, the Department of Energy, National Laboratories, PSERC, and Iowa Economic Development Authority. His research interests include optimization and data analytics in power distribution systems and microgrids. He was the recipient of the National Science Foundation CAREER Award, the Society-Level Outstanding Young Engineer Award from IEEE Power and Energy Society, the Northrop Grumman Endowment, College of Engineering's Early Achievement in Research Award, and the Harpole-Pentair Young Faculty Award Endowment. He is the Technical Committee Program Chair of IEEE Power System Operation, Planning and Economics Committee, the Chair of IEEE PSOPE Award Subcommittee, the Vice Chair of IEEE Distribution System Operation and Planning Subcommittee, and the Vice Chair of IEEE Task Force on Advances in Natural Disaster Mitigation Methods. He is an Associate Editor of the IEEE Transactions on Sustainable Energy, IEEE Open Access Journal of Power and Energy, IEEE Power Engineering Letters, and *IET Smart Grid*. He was an Associate Editor for the IEEE Transactions on Power Systems and IEEE Transactions on Smart Grid.

Matthew J. Reno (Senior Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology. He is currently a Distinguished Member of the Technical Staff with the Electric Power Systems Research Department, Sandia National Laboratories. His research interests include distribution system modeling and analysis with high penetration PV, including advanced software tools for automated analysis of hosting capacity, PV interconnection studies, and rapid quasi-static time series simulations. He is involved with the IEEE Power System Relaying Committee for developing guides and standards for the protection of microgrids and systems with high penetrations of inverter-based resources.

Joseph A. Azzolini (Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering with a focus on power systems from Arizona State University. He is currently a Senior Member of Technical Staff with the Electric Power Systems Research Department, Sandia National Laboratories, Albuquerque, NM, USA. His research interests include distribution system planning and analysis, power systems protection, distributed energy resource integration, and data-driven analysis.