Physics-Informed Fully Convolutional Network-Based Power Flow Analysis for Multi-Terminal MVDC Distribution Systems

Sun, Pingyang; Wu, Rongcheng; Wang, Hongyi; Li, Gen; Khalid, Muhammad; Konstantinou, Georgios

Published in:
I E E E Transactions on Power Systems

DOI (link to publication from Publisher):
10.1109/TPWRS.2024.3382266

Publication date:
2024

Document Version
Accepted author manuscript, peer reviewed version

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

Physics-informed Fully Convolutional Network-based Power Flow Analysis for Multi-terminal MVDC Distribution Systems

Pingyang Sun, *Graduate Student Member, IEEE,* Rongcheng Wu, Hongyi Wang, *Student Member, IEEE,* Gen Li, Muhammad Khalid, and Georgios Konstantinou, *Senior Member, IEEE*

***Abstract*—Numerical methods in power flow (PF) studies for medium-voltage direct current (MVDC) distribution systems require repetitive computations, particularly in scenarios with time-variable facilities that often alter the system operation points. Conventional neural networks (NNs), though efficient in rapid PF calculations, face accuracy challenges with untrained data distributions and varied topology structures. This highlights the need for more robust approaches to improve reliability in diverse scenarios. This paper proposes a physics-informed fully convolutional network (PI-FCN) to address this issue. The architecture of the PI-FCN is enhanced with the inclusion of two additional layers: i) a channel combination layer and ii) a physics operation layer. The former channel combination layer strengthens the model feature extraction capability by converting all input channels constituted by initial PF data matrix into dc voltage, current and line conductance matrix channels. The latter physics operation layer reformulates the combined input channels by physical connections in MVDC systems. The new layers enhance the prediction accuracy and allow generalization of the model. Five multi-terminal MVDC (MT-MVDC) distribution networks with different dc voltage levels and network layouts are used to verify the superiority of proposed PI-FCN compared to other NNs in fixed and varied topology structures.**

***Index Terms*—Multi-terminal medium-voltage direct current (MT-MVDC) system, power flow (PF), neural network (NN), fully-convolutional network (FCN).**

I. Introduction

MEDIUM-voltage direct current (MVDC) power systems have the potential to outperform equivalent ac systems in many applications, especially in distribution grids [1]. Multiple ac/dc and dc/dc converters can be interconnected in MVDC distribution systems to form multi-terminal MVDC (MT-MVDC) systems [2], [3]. To enhance operational security and future planning for MT-MVDC networks, it is crucial to examine the power flow (PF) analysis in response to different system operation scenarios such as line disconnection, converter outage [4]. Various numerical methods (also known as model-driven methods) can be employed for PF analysis in MT-MVDC networks [5]–[7]. Methods represented by the Newton-Raphson (NR) are effective for PF solutions in an MVDC distribution system that utilize various converters with diverse control schemes [8]. However, numerical methods come with certain disadvantages such as convergence for ill-conditioned systems, and computational complexity for large systems [8], [9].

With advances in computer processing capabilities, machine learning-based data-driven methods are increasingly employed to handle PF problems, such as regression techniques and neural networks (NNs) [10], [11]. Data-driven methods have the benefit of directly learning from historical data, bypassing the need for complex models and assumptions often required by model-driven approaches. This results in a more efficient, real-time optimization of PF, significantly enhancing decision-making speeds to three orders of magnitude [12] in practical examples, particularly in scenarios involving changes in load/generation and grid expansions [13]. An MT-MVDC distribution system often experiences shifts in steady-state operation points due to the presence of time-variable loads, distributed generation (DG), and energy storage systems (ESSs) with power sharing achieved through connected converters utilizing various control schemes [14]. This underscores the necessity for employing data-driven methods that possess excellent generalization capabilities, adapting to varying operational scenarios and ensuring stable system performance.

Regression techniques: Least squares regression (LSR) can be used to determine the variable relationships in a linearized PF model [10]. For scenarios of data multicollinearity in PF models, partial least square (PLS) [15] and ridge regression (RR) [10] can approach issues through extracting new combined variables (latent variables) and introducing penalty terms (regularization terms), respectively. The combination of PLS and Bayesian linear regression (BLR) effectively addresses both collinearity and overfitting issues, facilitating accurate linearization calculations [16]. Locally weighted RR (LWRR) incorporates locally weighted fitting into RR to identify time-varying sensitivities in grids with limited data

The work of Pingyang Sun and Georgios Konstantinou was supported under Australian Research Council's Discovery Project (DP210102294). The work of Rongcheng Wu was funded by the Digital Finance CRC supported by the Cooperative Research Centres program, an Australian Government initiative. The work of Muhammad Khalid was supported from IRC for Sustainable Energy Systems at KFUPM under project # INSE2415.

P. Sun, G. Konstantinou are with the School of Electrical Engineering and Telecommunications, UNSW Sydney, Australia (email: pingyang.sun@student.unsw.edu.au; g.konstantinou@unsw.edu.au).

R. Wu is with the Data Science Institute, University of Technology Sydney, Australia (email: rongcheng.wu@student.uts.edu.au).

H. Wang is with the Department of AAU energy, Alborg University, Denmark (email: howa@et.aau.dk).

M. Khalid is with the Electrical Engineering Department & Interdisciplinary Research Center for Sustainable Energy Systems, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia (email: mkhalid@kfupm.edu.sa).

G. Li is with the Electric Energy Section, Department of Engineering Technology, Technical University of Denmark (DTU), 2750 Ballerup, Denmark (email: genli@dtu.dk).

availability [17]. Support vector machine (SVM), is capable in handling data outliers and collinearity, achieving enhanced performance when integrated with RR [18]. Furthermore, PF mapping can be recovered from historical data in matrix form using support matrix regression (SMR) [19]. Nevertheless, linear regression is limited to capturing linear relationships in PF models, demonstrating reduced adaptability when dealing with high-dimensional data that exhibit complex nonlinear interactions.

Artificial NNs (ANNs): Conventional ANNs with fully-connected (FC) layers have found significant success in PF predictions. Multilayer perceptrons (MLPs) [11] and radial basis function networks (RBFNs) [20] are among the early methods employed in research for conducting PF analysis, particularly in ac transmission systems and microgrids. ANNs can also support PF analysis even in the absence of comprehensive initial system data [21]. In addition, ANNs have been widely used in optimal PF to enhance the efficiency of traditional methods [22] or to directly produce optimal solutions [23]–[25]. Although ANNs are effective in fixed topology structures, they encounter challenges when the system topologies shift. The number of neurons in the input layer of conventional ANNs must be specified to match the size or dimensionality of the input data, necessitating a fixed-size input. While inputs of variable size can be preprocessed through padding, truncating, or resampling, these operations may result in information loss and introduce invalid data [26].

Convolutional NNs (CNNs): Compared to conventional ANNs, CNNs inherently have the capability to handle input with variable sizes due to spatial hierarchies and local connectivity patterns, which allow adoption of different spatial dimensions without requiring a fixed-size input vector [27]. CNNs have been extensively employed in PF studies under uncertain contingency scenarios [28], demonstrating efficiencies up to 100 times [29] and 350 times [30], [31] that of traditional numerical methods for optimal PF solutions. In addition, the Latin-hypercube sampling method has been incorporated into CNNs to enhance training efficiency in probabilistic PF analyses [32]. The integration of CNNs with long short-term memory (LSTM) networks facilitates the inclusion of time-varying weather conditions in PF determinations [33].

Nevertheless, the traditional approach to CNN channel construction fails to fully capture the complex relationships between node and line parameters such as node voltage, power and line impedance that constitute the inherent physical connections in power systems. This lack of representation leads to diminished performance when dealing with unseen data, as it overlooks the interactions between these key elements that fundamentally drive the electrical network behaviors [34]. Moreover, the requirement to flatten the output of convolutional layers for integration with multiple FC layers in a conventional CNN constraints the flexibility of input topology structures [35].

This paper proposes a novel physics-informed fully-convolutional network (PI-FCN), specifically designed for PF studies in MT-MVDC distribution systems. It is also the first time to apply FCN in PF studies considering different converter control schemes of MVDC systems, including MVDC terminal voltage control, power/current control and droop control. In this model, the input channels are reformulated to account for the relationships between voltage, power, current at MVDC terminals and line parameters, while substituting the FC layer with a convolutional layer assuring flexible topology structure input.

The contributions in this paper are:

1) The proposed PI-FCN is suitable for PF studies of MT-MVDC distribution systems taking account of diverse control schemes. The inherent weight sharing and sparse connectivity characteristics in the PI-FCN efficiently reduce the training parameters/model capacity and alleviate the overfitting issue. By removing the FC layers, the proposed NN achieves fully variable-length input accepting the PF prediction of any MT-MVDC network structures after training. In addition, all input channels determined by the number of different MVDC bus types are consolidated into three channels constructed by dc voltage, current and line conductance matrices. This enhancement is realized through a specially designed channel combination layer, which effectively merges the data from the original power, current, power/voltage, current/voltage channels into a singular current channel. It simplifies the computation of the proposed network and reinforces feature extraction by integrating channels that may individually have limited data.
2) The proposed PI-FCN improves prediction accuracy through a physics-informed channel reformulation method. After the input channel combination, the voltage, current and line conductance channels are further reformulated based on the physical relation between the three channels. The voltage and current channels mutually transform through the line conductance channel, reflecting the actual voltage and current interactions. Furthermore, a more accurate representation of the physical connection is captured through the execution of multiple reformulation calculations. This operation enhances both the prediction accuracy and generalization capability, as it allows the network to comprehend the complex interaction between voltage, current, and line conductance.

To validate the superior performance of the proposed PI-FCN, it is tested across five MT-MVDC distribution networks featuring various dc voltage levels and connection methods, comparing it against other NNs in scenarios with both fixed and variable topologies.

II. Numerical Power Flow Approach for MT-MVDC Networks

In MT-MVDC distribution systems, an ac/dc or dc/dc converter can regulate PF by performing different control functions, including *i*) dc bus voltage control, *ii*) dc power or current control, and *iii*) power/voltage (P/V) or current/voltage (I/V) droop control [4]. Droop control is widely employed to automatically facilitate power sharing among converters connected to the ac grid, DGs, and ESSs following system disturbances such as line disconnections and converter outages [14].

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

According to the corresponding control modes of different converters [4], five MVDC bus types can be defined to compute the PF based on NR method in an MT-MVDC system with n buses. They are *i*) one MVDC voltage bus (slack bus), *ii*) m dc power buses, *iii*) l dc current buses, *iv*) k P/V droop-controlled buses, *v*) $(n-m-l-k-1)$ I/V droop-controlled buses. The PF equations are summarized in (1) to (5) for dc voltage, dc power, dc current, P/V droop-controlled and I/V droop-controlled buses, respectively.

$$V_{dcrefm,i} - V_{dcm,i} = 0, \qquad (i = 1) \tag{1}$$

$$P_{dcrefm,i} - P_{dcm,i} = 0, \qquad (i = 2, ..., m+1) \tag{2}$$

$$I_{dcrefm,i} - I_{dcm,i} = 0, \qquad (i = m+2, ..., l+m+1) \tag{3}$$

$$\begin{aligned}&(P_{dcrefm,i} - P_{dcm,i}) + K^{PV}_{droop,i}(V_{dcrefm,i} - V_{dcm,i}) = 0,\\ &(i = l+m+2, ..., k+l+m+1)\end{aligned} \tag{4}$$

$$\begin{aligned}&(I_{dcrefm,i} - I_{dcm,i}) + K^{IV}_{droop,i}(V_{dcrefm,i} - V_{dcm,i}) = 0,\\ &(i = k+l+m+2, ..., n)\end{aligned} \tag{5}$$

where V_{dcrefm} is usually set as rated voltage V^{rated}_{dcm}, $K^{PV}_{droop,i}$ and $K^{IV}_{droop,i}$ are the droop constants for converters with P/V and I/V droop, respectively. $K^{PV}_{droop,i} = (V^{rated}_{dcm}\delta_{droop})^{-1}P^{rated}_{dc}$, $K^{IV}_{droop,i} = (V^{rated}_{dcm} \cdot V^{rated}_{dcm}\delta_{droop})^{-1}P^{rated}_{dc}$ and δ_{droop} is the maximum allowable dc voltage deviation ratio [36].

Moreover, the dc power and current injected to any MVDC buses is expressed as:

$$\begin{cases} P_{dcm,i} = V_{dcm,i}\sum_{j=1}^{n} G_{dcm,ij}V_{dcm,j} & \text{(6a)}\\ I_{dcm,i} = \sum_{j=1}^{n} G_{dcm,ij}V_{dcm,j} & \text{(6b)}\end{cases}$$

Therefore, mismatch equations (7)-(10) can be obtained by substituting (6) into (2)-(5).

$$\Delta P^{P\text{-}bus}_{dcm,i} = P_{dcrefm,i} - V_{dcm,i}\sum_{j=1}^{n} G_{dcm,ij}V_{dcm,j}, \tag{7}$$

$$\Delta I^{I\text{-}bus}_{dcm,i} = I_{dcrefm,i} - \sum_{j=1}^{n} G_{dcm,ij}V_{dcm,j}, \tag{8}$$

$$\begin{aligned}\Delta P^{P/V\text{-}bus}_{dcm,i} =&(P_{dcrefm,i} + K^{PV}_{droop,i}V_{dcrefm,i}) - K^{PV}_{droop,i}V_{dcm,i}\\ &- V_{dcm,i}\sum_{j=1}^{n} G_{dcm,ij}V_{dcm,j},\end{aligned} \tag{9}$$

$$\begin{aligned}\Delta I^{I/V\text{-}bus}_{dcm,i} =&(I_{dcrefm,i} + K^{IV}_{droop,i}V_{dcrefm,i}) - K^{IV}_{droop,i}V_{dcm,i}\\ &- \sum_{j=1}^{n} G_{dcm,ij}V_{dcm,j},\end{aligned} \tag{10}$$

where $G_{dcm,ij}$ is the line conductance in an MT-MVDC system. Eqs. (7) to (10) can also be expressed in a matrix format as $\boldsymbol{\Delta F_{dcm}} = \boldsymbol{J_{dcm}}\boldsymbol{\Delta V_{dcm}}$, where $\boldsymbol{\Delta F_{dcm}} = [\boldsymbol{\Delta P^{P\text{-}bus}_{dcm}}, \boldsymbol{\Delta I^{I\text{-}bus}_{dcm}}, \boldsymbol{\Delta P^{P/V\text{-}bus}_{dcm}}, \boldsymbol{\Delta I^{I/V\text{-}bus}_{dcm}}]^T$, $\boldsymbol{\Delta V_{dcm}} = [\boldsymbol{\Delta V^{P\text{-}bus}_{dcm}}, \boldsymbol{\Delta V^{I\text{-}bus}_{dcm}}, \boldsymbol{\Delta V^{P/V\text{-}bus}_{dcm}}, \boldsymbol{\Delta V^{I/V\text{-}bus}_{dcm}}]^T$, $\boldsymbol{J_{dcm}}$ represent the derived Jacobian matrix. The detailed calculations for $\boldsymbol{J_{dcm}}$ can be found in [37].

Although a numerical approach is capable of accurately determining the PF in MT-MVDC distribution systems, its dependence on iterative calculations significantly increases the computational cost. This high demand for computational resources renders it impractical for rapid PF estimation in MT-MVDC systems, especially those characterized by time-variable loads, DGs, and ESSs, where processing efficiency are of paramount importance.

III. Proposed PI-FCN for MT-MVDC Distribution Networks

A. Conventional CNN Description

A CNN comprises multiple convolutional layers that extract various feature maps from inputs (F) using specified convolution kernels (K). An activation function $\sigma(\cdot)$ offers nonlinearity to CNN. The first convolution layer connected to the input map in a CNN should follow the forward propagation rule [27] as:

$$\begin{cases} C_f = \sigma\left(F * K_{1,f} + B_f\right), \\ C_f(u,v) = \sigma\left(\sum_{g=-r}^{r}\sum_{h=-r}^{r} F(u-g, v-h)\cdot K_{1,f}(g,h) + B_f\right), \end{cases} \tag{11}$$

where f refers to the f^{th} feature map, $*$ denotes the convolution operator, B is bias, r is the length of convolution range, and u, v are row and column indices of feature map C_f. Pooling layer can also be used in CNNs, while not necessary, to achieve dimensionality reduction and perform downsampling. Average pooling (D^{ave}) and maximum pooling (D^{max}) are two typical methods expressed as:

$$D^{ave}_f(u,v) = \frac{1}{w^2}\sum_{g=0}^{w-1}\sum_{h=0}^{w-1} D_f(s\cdot u - g, s\cdot v - h), \text{ or} \tag{12}$$

$$D^{max}_f(u,v) = Max^{w-1}_{g=0}Max^{w-1}_{h=0} D_f(s\cdot u - g, s\cdot v - h), \tag{13}$$

where w represents the size of the pooling window, and s denotes the stride with which the window is moved. Moreover, the size of convolution and pooling layers is respectively calculated as $\lfloor(\text{size}(F) - \text{size}(K) + 2p)/s + 1\rfloor$ and $\lceil(\text{size}(C) - \text{size}(K))/s + 1\rceil$, where p refers to the total number of padding. The FC layer can be used as the final layer in a conventional CNN, although it is not required for an FCN.

$$O_{fc} = \sigma(Wh + B), \tag{14}$$

where $h = H(\{D_f\})$ represents the vectorization of each D_f and concatenation of all vectors [38].

B. Network Description of Proposed PI-FCN

The proposed PI-FCN, as shown in Fig. 1, uses a FCN with residual connection and incorporates a physics-based convolutional channel redesign. The new NN offers two key advancements: 1) accommodating variable input sizes by adaptable kernel sliding across input matrices, and 2) improving prediction accuracy with the integration of restructured input channels. Following the PF determination discussed in Section II, there are six input channels including voltage, power, current, P/V droop, I/V droop and line conductance channels. However, both forward and backward propagations require more computations with the increase in the number of

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

input channels. The potential for information extraction could be constrained when dealing with a channel that contains less effective information (fewer number of buses for a specific bus type). Beyond the two factors, a critical concern is the absence of inter-channel data exchange, considering the relationships among terminal voltage, power, current, and line conductance in an MT-MVDC system, which results in reduced network generalization. To overcome these issues, the six channels are reformulated to three channels (voltage, current and conductance) while embedding physics information. A detailed discussion of the channel reformulation process will be provided in Section IV.

In the proposed PI-FCN, there is one channel combination layer (depicted by the green box in Fig. 1), one physics operation layer (red box), and six encoder blocks, each comprising two convolutional layers (orange boxes). The number of input features $N_{\text{feature,in}}$ is equivalent to the number of channels N_{channel} in the initial encoder block, and it progressively expands to $8N_{\text{feature,in}}$ by the final encoder block. The data output (dc voltage at different MVDC buses), calculated as (15), emanates from an additional convolutional layer illustrated as a grey box. Moreover, residual connection is further added as (16) (purple boxes) for mitigating vanishing/exploding gradient and improving model stability [39].

$$O_{\text{PI-FCN}} = \sigma\left(C_f^{(e6)} * K_f^{(\text{final})} + B_f^{(\text{final})}\right) \tag{15}$$

$$\begin{cases} C_f^{(e3)} = \sigma\left(\left(C_f^{(e2)} + C_f^{(r1)}\right) * K_f^{(e3)} + B_f^{(e3)}\right) \\ C_f^{(e5)} = \sigma\left(\left(C_f^{(e4)} + C_f^{(r2)}\right) * K_f^{(e5)} + B_f^{(e5)}\right) \end{cases} \tag{16}$$

The network input does not specify the height and width of the input data to allow for flexibility in accommodating varying sizes. In addition, the proposed PI-FCN omits the pooling layer to prevent the potential loss of information that could disrupt the intrinsic physical relation among different input channels.

IV. Physics-informed Convolutional Channel Reformulation

Physics-informed learning techniques can be embedded into a NN through architecture reformulation to strengthen the NN performance [40]. In the proposed PI-FCN, the network architecture is reformulated through a physics-guided convolutional channel operation, which offers 1) simplified computation complexity and improved feature extraction capability by the transformation of initial input channels; and 2) enhanced prediction accuracy/generalization capability by the reconstruction of merged channels, using the physical relations between line conductance, voltage and current in MVDC systems. A channel reformulation example provided in this section illustrates the process of capturing physical information in detail.

A. Channel Reformulation Approach

1) Input Channel Combination: Based on MVDC PF equations (1) to (5), the matrices of initial six input channels are constructed as (17) to (22) for voltage, power, current, P/V droop, I/V droop and line conductance, respectively.

$$\boldsymbol{\mathcal{CH}_V} = \text{diag}(V_{dcrefm,1}, V_{dcrefm,1}, ..., V_{dcrefm,1})^{n\times n} \tag{17}$$

$$\boldsymbol{\mathcal{CH}_P} = \text{diag}(0, P_{dcrefm,2}, ..., P_{dcrefm,m+1}, ..., 0)^{n\times n} \tag{18}$$

$$\boldsymbol{\mathcal{CH}_I} = \text{diag}(0, ..., I_{dcrefm,m+2}, .., I_{dcrefm,l+m+1}, ..., 0)^{n\times n} \tag{19}$$

$$\begin{aligned}\boldsymbol{\mathcal{CH}_{P/V}} =&\text{diag}(0, ..., P_{dcrefm,l+m+2} + K^{PV}_{droop,l+m+2}V_{dcrefm,l+m+2}, \\ &.., P_{dcrefm,k+l+m+1} + K^{PV}_{droop,k+l+m+1}V_{dcrefm,k+l+m+1}, \\ &..., 0)^{n\times n}\end{aligned} \tag{20}$$

$$\begin{aligned}\boldsymbol{\mathcal{CH}_{I/V}} =&\text{diag}(0, ..., I_{dcrefm,k+l+m+2} + K^{IV}_{droop,k+l+m+2}\times \\ &V_{dcrefm,k+l+m+2}, .., I_{dcrefm,n} + K^{IV}_{droop,n}V_{dcrefm,n})^{n\times n}\end{aligned} \tag{21}$$

$$\boldsymbol{\mathcal{CH}_G} = \boldsymbol{G_{dcm}} \tag{22}$$

However, fewer input channels are required for simplifying computations in forward/backward information propagations. The six input channels operate independently, without any data exchange based on physical knowledge, resulting in limited adaptability to changing topological structures.

A physics-informed channel combination scheme is presented to address the issue. Initially, the matrices (18), (20), and (21) related to power, P/V droop, and I/V droop channels are transformed into the format of matrix (19) by following transformation:

$$\begin{cases} \boldsymbol{\mathcal{CH}_P} \rightarrow \boldsymbol{\mathcal{CH}_I^P}, \quad \text{by } P_{dcrefm,i}/V^{rated}_{dcm} & \text{(23a)} \\ \boldsymbol{\mathcal{CH}_{P/V}} \rightarrow \boldsymbol{\mathcal{CH}_I^{P/V}}, \quad \text{by } P_{dcrefm,i}/V_{dcrefm,i} & \text{(23b)} \\ \boldsymbol{\mathcal{CH}_{I/V}} \rightarrow \boldsymbol{\mathcal{CH}_I^{I/V}}, \quad \text{Only keep } I_{dcrefm,i} & \text{(23c)} \end{cases}$$

They are merged with the current matrix (19) to create an updated current matrix as (24), hence the six input channels are converted into three channels (voltage, current, conductance).

$$\boldsymbol{\mathcal{CH}_{I,\text{total}}} = \boldsymbol{\mathcal{CH}_I^P} + \boldsymbol{\mathcal{CH}_I} + \boldsymbol{\mathcal{CH}_I^{P/V}} + \boldsymbol{\mathcal{CH}_I^{I/V}} \tag{24}$$

2) Inter-channel Transformation: Further conversion is needed to enhance the NN adaptability to diverse input network structures. Eq. (6b) shows the terminal current has a linear correlation with the terminal voltage. Hence, (25) can be established for any input power or current vectors derived from (6b) and (23).

$$\mathcal{I}_{dcm,i} = \sum_{j=1}^{n} G_{dcm,ij} \cdot V_{dcm,j}, \tag{25}$$

where $\mathcal{I}_{dcm,i}$ includes accurate and estimated dc current values. In addition, (26) can also be established by splitting self-conductance and mutual conductance of (25).

$$\mathcal{I}_{dcm,i} = G_{dcm,ii}V_{dcm,i} + \sum_{j\neq i}^{n} G_{dcm,ij}V_{dcm,j}, \tag{26}$$

which can also be written as (27) to reflect how one node voltage is influenced by its neighbouring node voltage values.

$$V_{dcm,i} = \frac{\mathcal{I}_{dcm,i} - \sum\limits_{j\neq i}^{n} G_{dcm,ij} \cdot V_{dcm,j}}{G_{dcm,ii}} \tag{27}$$

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

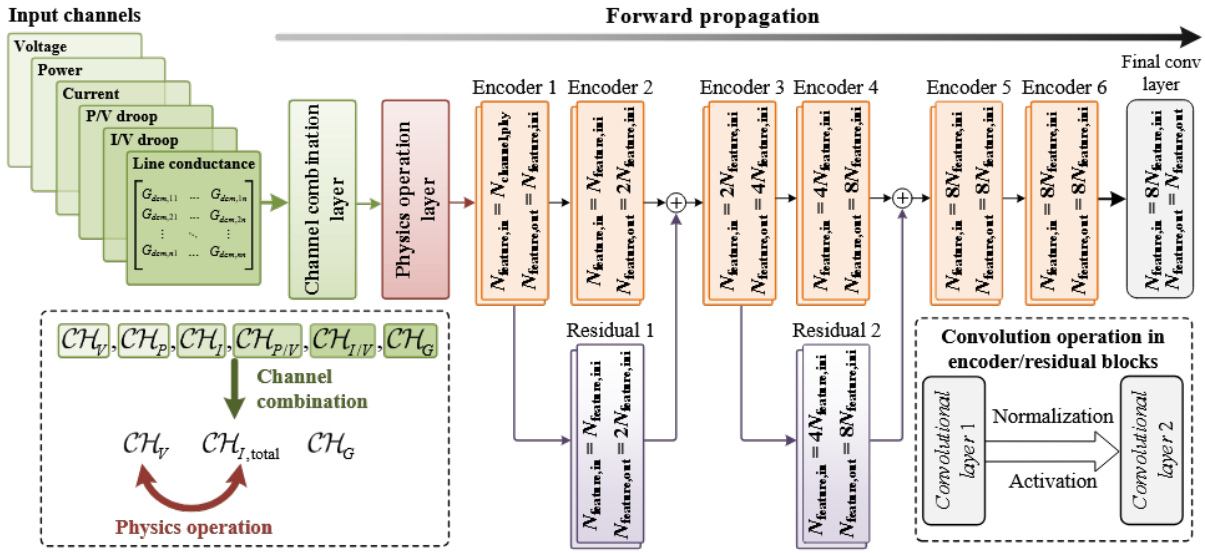


Fig. 1. Network structure of proposed PI-FCN for MT-MVDC distribution systems.

The input voltage and current matrices are further reconstructed by (28) and (29), respectively, using matrix $G_{dcm,ij}$.

$$\underbrace{\boldsymbol{\mathcal{I}}_{\boldsymbol{dcm}}^{\mathbf{new}}}_{\mathcal{CH}_{I,\text{total}}^{\text{new}}} = \underbrace{\boldsymbol{G}_{\boldsymbol{dcm}}}_{\mathcal{CH}_G} \cdot \underbrace{\text{diag}(\boldsymbol{V}_{\boldsymbol{dcm}})}_{\mathcal{CH}_V} \tag{28}$$

$$\underbrace{\boldsymbol{V}_{\boldsymbol{dcm}}^{\mathbf{new}}}_{\mathcal{CH}_V^{\text{new}}} = \frac{\underbrace{\text{diag}(\boldsymbol{\mathcal{I}}_{\boldsymbol{dcm}})}_{\mathcal{CH}_{I,\text{total}}} - \underbrace{\text{nondiag}(\boldsymbol{G}_{\boldsymbol{dcm}})}_{\mathcal{CH}_G^{\text{nondiag}}} \cdot \underbrace{\text{diag}(\boldsymbol{V}_{\boldsymbol{dcm}})}_{\mathcal{CH}_V}}{\underbrace{\text{diag}(\boldsymbol{G}_{\boldsymbol{dcm}})}_{\mathcal{CH}_G^{\text{diag}}}} \tag{29}$$

The reconstructed voltage and current matrices consist of N dc voltage and current components:

$$\boldsymbol{\mathcal{X}}_{\boldsymbol{dcm}}^{\mathbf{new}} = \begin{bmatrix} \mathcal{X}_{dcm,11} & \mathcal{X}_{dcm,12} & \cdots & \mathcal{X}_{dcm,1N} \\ \mathcal{X}_{dcm,21} & \mathcal{X}_{dcm,21} & \cdots & \mathcal{X}_{dcm,2N} \\ \vdots & \vdots & \ddots & \vdots \\ \underbrace{\mathcal{X}_{dcm,n1}}_{\text{component 1}} & \underbrace{\mathcal{X}_{dcm,n2}}_{\text{component 2}} & \cdots & \underbrace{\mathcal{X}_{dcm,nN}}_{\text{component N}} \end{bmatrix}, \tag{30}$$

where $\mathcal{X}$ refers to V or $\mathcal{I}$ in (28) and (29), respectively. The sum of all components in each row of (30) is the actual dc voltage or current values as:

$$\begin{cases} \sum_{row} V_{dcm,i}^{\text{new}} = \sum_{N=1}^{N_{\text{column}}} V_{dcm,iN}^{\text{new}}, & \text{(31a)} \\ \sum_{row} \mathcal{I}_{dcm,i}^{\text{new}} = \sum_{N=1}^{N_{\text{column}}} \mathcal{I}_{dcm,iN}^{\text{new}}. & \text{(31b)} \end{cases}$$

This operation achieves mutual conversion of the current channel and voltage channel as shown in Fig. 2. In voltage channel $\boldsymbol{\mathcal{CH}_V}$, the 2^{nd} to n^{th} diagonal voltage elements are substituted by corresponding elements in $V_{dcm,i}^{\text{new}}$.

$$\mathcal{CH}_{Vi} = \sum_{row} V_{dcm,i}^{\text{new}} \quad (i = 2, 3, ..., n). \tag{32}$$

In current channel $\boldsymbol{\mathcal{CH}_{I,\text{total}}}$, dc current elements in addition to the elements related to the current buses should be all updated based on the following rule:

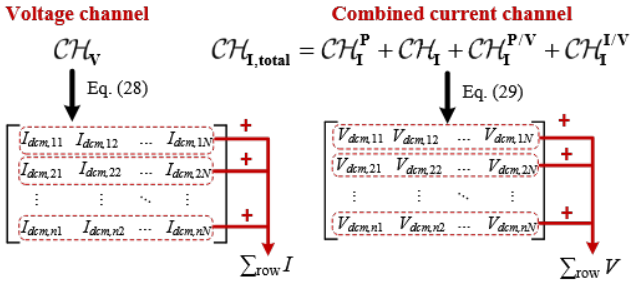


Fig. 2. Input channel conversion process.

- Voltage bus-related element:

$$\mathcal{CH}_{Ii,\text{total}} = \sum_{row} \mathcal{I}_{dcm,i}^{\text{new}} \quad (i = 1) \tag{33}$$

- Power bus-related elements:

$$\mathcal{CH}_{Ii,\text{total}} = P_{dcrefm,i} / \sum_{row} V_{dcm,i}^{\text{new}} \quad (i = 2, ..., m+1) \tag{34}$$

- P/V bus-related elements:

$$\mathcal{CH}_{Ii,\text{total}} = \frac{P_{dcrefm,i} + K_{droop,i}^{PV}(V_{dcrefm,i} - \sum_{row} V_{dcm,i}^{\text{new}})}{\sum_{row} V_{dcm,i}^{\text{new}}}, \quad (i = l+m+2, ..., k+l+m+1) \tag{35}$$

- I/V bus-related elements:

$$\mathcal{CH}_{Ii,\text{total}} = P_{dcrefm,i} + K_{droop,i}^{IV}(V_{dcrefm,i} - \sum_{row} V_{dcm,i}^{\text{new}}), \quad (i = k+l+m+2, ..., n) \tag{36}$$

By such conversion method, the model generalization and prediction accuracy can be increased due to the embedded physics principle. Utilizing a single channel conversion calculation by (28) and (29) yields limited physics information. By employing multiple conversion calculations or iterations, one can garner adequate physics insights. The complete flowchart illustrating the physics information embedding process is

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

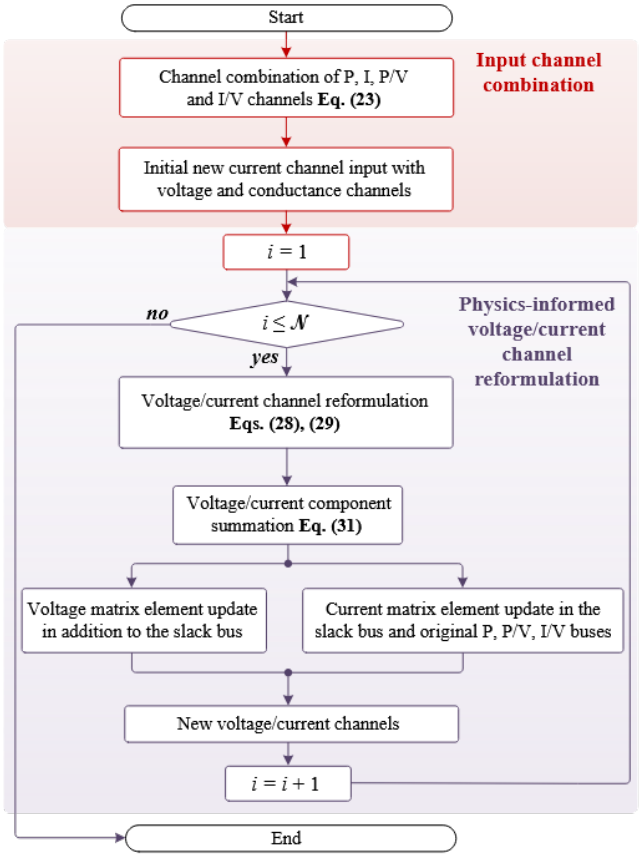


Fig. 3. Physics information embedding process in the proposed PI-FCN for MT-MVDC distribution systems.

shown in Fig. 3. However, it is important to recognize that while the training time increases with the number of iterations, obtaining sufficient physics information can expedite network convergence.

B. Matrix Element Update Explanation

A six-terminal MVDC (6T-MVDC) distribution network is used to explain the channel reformulation process. The network structure is shown in Fig. 5(b) and related parameters are provided in Table I. Initial conditions for PF determination are:

- Grid-connected converter 1 (voltage control) - voltage reference (12 kV);
- Grid-connected converter 2 (P/V droop control) - voltage reference, power reference, P/V droop constant (12 kV, 8 MW, 8.33 MW/kV);
- Load-connected converter 3 (current control) - voltage reference, current reference (12 kV, −0.6 kA);
- Load-connected converter 4 (power control) - voltage reference, power reference (12 kV, −8 MW);
- DG-connected converter 5 (I/V droop control) - voltage reference, current reference, I/V droop constant (12 kV, 0.1 kA, 0.3472 kA/kV);
- ESS-connected converter 6 (I/V droop control) - voltage reference, current reference, I/V droop constant (12 kV, 0 kA, 0.3472 kA/kV).

Following the channel reformulation process in Section IV-A, the initial six input matrices are converted into three by (23). The reformulated three channels are then acquired:

$$\mathcal{CH}_{I,\text{total}} = \text{diag}(0, \underbrace{\mathbf{0.67}}_{\text{(23b)}}, -0.6, \underbrace{\mathbf{-0.67}}_{\text{(23a)}}, \underbrace{\mathbf{0.1}}_{\text{(23c)}}, \underbrace{\mathbf{0}}_{\text{(23c)}}) \tag{37}$$

$$\mathcal{CH}_V = \text{diag}(12, 12, 12, 12, 12, 12) \tag{38}$$

$$\mathcal{CH}_G = \begin{bmatrix} 2.25 & 0 & -1.25 & 0 & -1 & 0 \\ 0 & 2.25 & 0 & -1.25 & 0 & -1 \\ -1.25 & 0 & 1.25 & 0 & 0 & 0 \\ 0 & -1.25 & 0 & 2.92 & -1.67 & 0 \\ -1 & 0 & 0 & -1.67 & 2.67 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{39}$$

Furthermore, new current and voltage matrices are obtained after executing one calculation:

$$\mathcal{I}_{dcm}^{\text{new}(1)} = \begin{bmatrix} \underbrace{27}_{\mathcal{I}_{dcm,1}^{c.1}} & \underbrace{0}_{\mathcal{I}_{dcm,1}^{c.2}} & \underbrace{-15}_{\mathcal{I}_{dcm,1}^{c.3}} & \underbrace{0}_{\mathcal{I}_{dcm,1}^{c.4}} & \underbrace{-12}_{\mathcal{I}_{dcm,1}^{c.5}} & \underbrace{0}_{\mathcal{I}_{dcm,1}^{c.6}} \\ \underbrace{0}_{\mathcal{I}_{dcm,2}^{c.1}} & \underbrace{27}_{\mathcal{I}_{dcm,2}^{c.2}} & \underbrace{0}_{\mathcal{I}_{dcm,2}^{c.3}} & \underbrace{-15}_{\mathcal{I}_{dcm,2}^{c.4}} & \underbrace{0}_{\mathcal{I}_{dcm,2}^{c.5}} & \underbrace{-12}_{\mathcal{I}_{dcm,2}^{c.6}} \\ \underbrace{-15}_{\mathcal{I}_{dcm,3}^{c.1}} & \underbrace{0}_{\mathcal{I}_{dcm,3}^{c.2}} & \underbrace{15}_{\mathcal{I}_{dcm,3}^{c.3}} & \underbrace{0}_{\mathcal{I}_{dcm,3}^{c.4}} & \underbrace{0}_{\mathcal{I}_{dcm,3}^{c.5}} & \underbrace{0}_{\mathcal{I}_{dcm,3}^{c.6}} \\ \underbrace{0}_{\mathcal{I}_{dcm,4}^{c.1}} & \underbrace{-15}_{\mathcal{I}_{dcm,4}^{c.2}} & \underbrace{0}_{\mathcal{I}_{dcm,4}^{c.3}} & \underbrace{35}_{\mathcal{I}_{dcm,4}^{c.4}} & \underbrace{-20}_{\mathcal{I}_{dcm,4}^{c.5}} & \underbrace{0}_{\mathcal{I}_{dcm,4}^{c.6}} \\ \underbrace{-12}_{\mathcal{I}_{dcm,5}^{c.1}} & \underbrace{0}_{\mathcal{I}_{dcm,5}^{c.2}} & \underbrace{0}_{\mathcal{I}_{dcm,5}^{c.3}} & \underbrace{-20}_{\mathcal{I}_{dcm,5}^{c.4}} & \underbrace{32}_{\mathcal{I}_{dcm,5}^{c.5}} & \underbrace{0}_{\mathcal{I}_{dcm,5}^{c.6}} \\ \underbrace{0}_{\mathcal{I}_{dcm,6}^{c.1}} & \underbrace{-12}_{\mathcal{I}_{dcm,6}^{c.2}} & \underbrace{0}_{\mathcal{I}_{dcm,6}^{c.3}} & \underbrace{0}_{\mathcal{I}_{dcm,6}^{c.4}} & \underbrace{0}_{\mathcal{I}_{dcm,6}^{c.5}} & \underbrace{12}_{\mathcal{I}_{dcm,6}^{c.6}} \end{bmatrix}, \tag{40}$$

$$\boldsymbol{V}_{dcm}^{\text{new}(1)} = \begin{bmatrix} \underbrace{0}_{V_{dcm,1}^{c.1}} & \underbrace{0}_{V_{dcm,1}^{c.2}} & \underbrace{6.6667}_{V_{dcm,1}^{c.3}} & \underbrace{0}_{V_{dcm,1}^{c.4}} & \underbrace{5.3333}_{V_{dcm,1}^{c.5}} & \underbrace{0}_{V_{dcm,1}^{c.6}} \\ \underbrace{0}_{V_{dcm,2}^{c.1}} & \underbrace{0.2978}_{V_{dcm,2}^{c.2}} & \underbrace{0}_{V_{dcm,2}^{c.3}} & \underbrace{6.6667}_{V_{dcm,2}^{c.4}} & \underbrace{0}_{V_{dcm,2}^{c.5}} & \underbrace{5.3333}_{V_{dcm,2}^{c.6}} \\ \underbrace{12}_{V_{dcm,3}^{c.1}} & \underbrace{0}_{V_{dcm,3}^{c.2}} & \underbrace{-0.48}_{V_{dcm,3}^{c.3}} & \underbrace{0}_{V_{dcm,3}^{c.4}} & \underbrace{0}_{V_{dcm,3}^{c.5}} & \underbrace{0}_{V_{dcm,3}^{c.6}} \\ \underbrace{0}_{V_{dcm,4}^{c.1}} & \underbrace{5.1429}_{V_{dcm,4}^{c.2}} & \underbrace{0}_{V_{dcm,4}^{c.3}} & \underbrace{-0.2297}_{V_{dcm,4}^{c.4}} & \underbrace{6.8571}_{V_{dcm,4}^{c.5}} & \underbrace{0}_{V_{dcm,4}^{c.6}} \\ \underbrace{4.5}_{V_{dcm,5}^{c.1}} & \underbrace{0}_{V_{dcm,5}^{c.2}} & \underbrace{0}_{V_{dcm,5}^{c.3}} & \underbrace{7.5}_{V_{dcm,5}^{c.4}} & \underbrace{0.0375}_{V_{dcm,5}^{c.5}} & \underbrace{0}_{V_{dcm,5}^{c.6}} \\ \underbrace{0}_{V_{dcm,6}^{c.1}} & \underbrace{12}_{V_{dcm,6}^{c.2}} & \underbrace{0}_{V_{dcm,6}^{c.3}} & \underbrace{0}_{V_{dcm,6}^{c.4}} & \underbrace{0}_{V_{dcm,6}^{c.5}} & \underbrace{0}_{V_{dcm,6}^{c.6}} \end{bmatrix}. \tag{41}$$

By summing all components in each row of (40) and (41), the actual dc voltage and current values are:

$$\begin{cases} \sum_{row} \mathcal{I}_{dcm}^{\text{new}(1)} = [0, 0, 0, 0, 0, 0]^T \text{ kA} \\ \sum_{row} \boldsymbol{V}_{dcm}^{\text{new}(1)} = [12, 12.2978, 11.52, 11.7703, 12.0375, 12]^T \text{ kV} \end{cases} \tag{42}$$

Following the element update rule (32) to (36), the updated current and voltage channels are expressed as:

$$\begin{cases} \mathcal{CH}_{I,\text{total}}^{\text{update}(1)} = \text{diag}(0, 0.4487, -0.6, -0.6797, 0.0870, 0) \\ \mathcal{CH}_{V}^{\text{update}(1)} = \text{diag}(12, 12.2978, 11.52, 11.7703, 12.0375, 12) \end{cases} \tag{43}$$

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

Efficient physics information would be captured by executing multiple iterations ($i = \mathcal{N}$). The new derived current and voltage matrices after ten iteration calculations are:

$$\boldsymbol{\mathcal{I}}_{\boldsymbol{dcm}}^{\textbf{new(10)}} = \begin{bmatrix} 27 & 0 & -14.4 & 0 & -12.0489 & 0 \\ 0 & 28.2929 & 0 & -14.5191 & 0 & -11.6775 \\ -15 & 0 & 14.4 & 0 & 0 & 0 \\ 0 & -15.7183 & 0 & 33.8780 & -20.0814 & 0 \\ -12 & 0 & 0 & -19.3588 & 32.1303 & 0 \\ 0 & -12.5746 & 0 & 0 & 0 & 11.6775 \end{bmatrix}, \tag{44}$$

$$\boldsymbol{V}_{\boldsymbol{dcm}}^{\textbf{new(10)}} = \begin{bmatrix} 0.3438 & 0 & 6.4 & 0 & 5.3551 & 0 \\ 0 & 0.1135 & 0 & 6.4529 & 0 & 5.19 \\ 12 & 0 & -0.48 & 0 & 0 & 0 \\ 0 & 5.3891 & 0 & -0.2361 & 6.8851 & 0 \\ 4.5 & 0 & 0 & 7.2596 & 0.0311 & 0 \\ 0 & 12.5746 & 0 & 0 & 0 & 0.112 \end{bmatrix}. \tag{45}$$

The actual dc voltage and current values are calculated as:

$$\begin{cases} \sum_{row} \boldsymbol{\mathcal{I}}_{\boldsymbol{dcm}}^{\textbf{new(10)}} = [0.5511, 2.0963, -0.6, -1.9217, \\ 0.7715, -0.8971]^T \text{ kA} \\ \sum_{row} \boldsymbol{V}_{\boldsymbol{dcm}}^{\textbf{new(10)}} = [12.0989, 12.7565, 11.52, 12.0381, \\ 11.7907, 12.6866]^T \text{ kV} \end{cases} \tag{46}$$

by summing all components in each row of (44) and (45). Furthermore, the updated current and voltage channels are expressed as:

$$\begin{cases} \boldsymbol{\mathcal{CH}}_{\boldsymbol{I,\text{total}}}^{\textbf{update(10)}} = \text{diag}(0.5511, 0.8531, -0.6, -0.6646, \\ 0.1727, -0.2384) \\ \boldsymbol{\mathcal{CH}}_{\boldsymbol{V}}^{\textbf{update(10)}} = \text{diag}(12, 11.7565, 11.52, 12.0381, \\ 11.7907, 12.6866) \end{cases} \tag{47}$$

This procedure demonstrates how the proposed method for input channel reformulation captures physical information across all terminals through multiple iterations. The provided pseudo-code *Algorithm* 1, written in Python using the Pytorch framework [41], represents the proposed iterative calculation for the 6T-MVDC distribution network after 10 iterations.

Algorithm 1 Iteration calculation for voltage/current channel matrices in 4T-MVDC distribution network.

$\mathcal{I}_{\text{dcm}}^{\text{2D-Input}} \leftarrow \text{diag}([0, 0.67, -0.6, -0.67, 0.1, 0])$
$V_{\text{dcm}}^{\text{2D-Input}} \leftarrow \text{diag}([12, 12, 12, 12, 12, 12])$
$\mathcal{N} \leftarrow 10$
for i in range($\mathcal{N}$) **do**
  $\mathcal{I}_{\text{dcm}}^{\text{new}} \leftarrow G_{\text{dcm}} \times V_{\text{dcm}}^{\text{2D-Input}}$
  $V_{\text{dcm}}^{\text{new}} \leftarrow \text{inv}(\text{diag}G_{\text{dcm}}) \times (\mathcal{I}_{\text{dcm}}^{\text{2D-Input}} - \text{ndiag}G_{\text{dcm}} \times V_{\text{dcm}}^{\text{2D-Input}})$
  $\sum_{row} \mathcal{I}_{dcm}^{\text{new}} \leftarrow \text{sum}(G_{\text{dcm}} \times V_{\text{dcm}}^{\text{2D-Input}}, \text{axis} = 1)$
  $\sum_{row} V_{dcm}^{\text{new}} \leftarrow \text{sum}(\text{inv}(\text{diag}G_{\text{dcm}}) \times (\mathcal{I}_{\text{dcm}}^{\text{2D-Input}} - \text{ndiag}G_{\text{dcm}} \times V_{\text{dcm}}^{\text{2D-Input}}), \text{axis} = 1)$
  $V_{\text{dcm}}^{\text{2D-Input}}[1:] \leftarrow \sum_{row} V_{dcm}^{\text{new}}[1:]$
  $\mathcal{I}_{\text{dcm}}^{\text{2D-Input}}[\text{type V}] \leftarrow \sum_{row} \mathcal{I}_{dcm}^{\text{new}}[\text{type V}]$
  $\mathcal{I}_{\text{dcm}}^{\text{2D-Input}}[\text{type P}] \leftarrow P_{\text{dcrefm}}[\text{type P}]/V_{\text{dcm}}^{\text{2D-Input}}[\text{type P}]$
  $\mathcal{I}_{\text{dcm}}^{\text{2D-Input}}[\text{type P/V}] \leftarrow (P_{\text{dcrefm}}[\text{type P/V}] + K_{\text{droop}}^{\text{P/V}} \times (V_{\text{dcrefm}}[\text{type P/V}] - V_{\text{dcm}}^{\text{2D-Input}}[\text{type P/V}]))/V_{\text{dcm}}^{\text{2D-Input}}[\text{type P/V}]$
  $\mathcal{I}_{\text{dcm}}^{\text{2D-Input}}[\text{type I/V}] \leftarrow I_{\text{dcrefm}}[\text{type I/V}] + K_{\text{droop}}^{\text{I/V}} \times (V_{\text{dcrefm}}[\text{type I/V}] - V_{\text{dcm}}^{\text{2D-Input}}[\text{type P/V}])$
end for

C. Training Process

The general training process of the proposed PI-FCN is shown in Fig. 4. Data is acquired using the conventional NR method implemented in Python. Detailed information on data acquisition will be discussed in Section V-A. Data sequentially flow through all layers, with the number of features incrementally increasing from the initial number of channels to eightfold, thereby enhancing the learning capability of the proposed PI-FCN. The initial six input channels are first converted into three channels through a predefined channel combination layer based on (23) and (24). Moreover, the subsequent physics operation layer provides physical insights to the network by the mutual conversion of these combined three channels according to (28)-(36).

The encoder and residual blocks perform convolutional operations on the output channels of the physics operation layer. Each encoder/residual block contains two convolutional layers with different kernel sizes, stride and padding. Following each convolutional operation, normalization technique such as Instance Normalization is utilized to ensure data stability, while an activation function like the Rectified linear unit (ReLU) is applied to introduce nonlinearity into the network. The predicted MVDC bus voltage values from the proposed PI-FCN are finally compared with the target values by the computation of loss function gradient during the backward propagation to update the weights.

Hyperparameters are fine-tuned through both Bayesian optimization and heuristic methods. The evaluation metric is the accuracy on the validation set. Bayesian optimization specifically determines the number of encoder/residual blocks and the convolutional layers in each block, while heuristic tuning is used to set the values of other parameters. Early stopping is applied on the validation set to mitigate overfitting, which also assists the adjustment of hyperparameters by interrupting the training process at optimal epochs.

V. Experiment Verification

The proposed PI-FCN, along with other neural network models, are developed using PyTorch. Three distinct cases are considered to validate the superior performance of the proposed PI-FCN in MT-MVDC distribution networks compared to conventional FCN and MLP. All evaluations were conducted on a PC equipped with an Intel(R) Core(TM) i7-13700K CPU@3.4GHz, 64GB RAM, and a NVIDIA GeForce RTX 4090 GPU. The three cases include: 1) fixed network topology that five designed MT-MVDC networks as shown in Fig. 5(a)-(e) are trained using random sampling of line resistance, as well as reference values for grid-connected, DG-connected, ESS-connected, and load-connected converters; 2) varying network topology that random line disconnections are considered; 3) varying network topology that random converter outages and corresponding line disconnections are included.

A. Training Data Derivation

The initial PF data is obtained from the MT-MVDC systems through a conventional NR algorithm programmed in Python.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

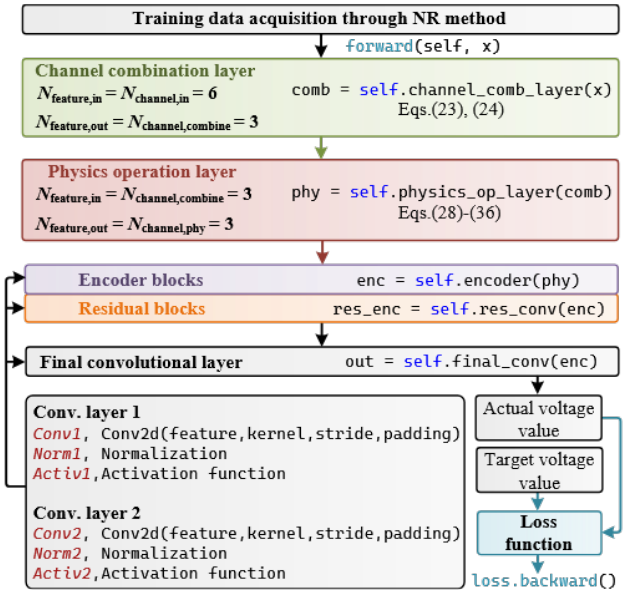


Fig. 4. General training process of proposed PI-FCN.

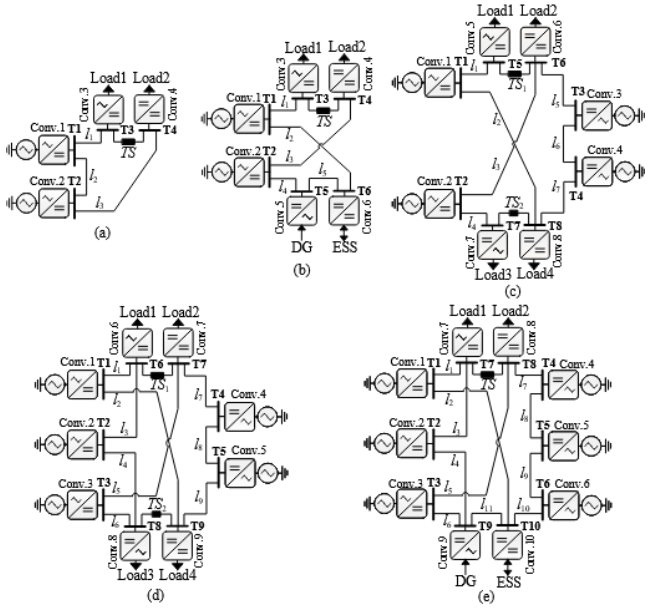


Fig. 5. MT-MVDC network structures: (a) four-terminal MVDC system with load, (b) six-terminal MVDC system with load/DG/ESS, (c) eight-terminal MVDC system with load, (d) nine-terminal MVDC system with load, and (e) ten-terminal MVDC system with load/DG/ESS.

The conventional PF derivation approach has been discussed in Section II considering different control methods used with power electronic converters [37]. This PF algorithm incorporates different control modes in the bus type definition (1) to (5), making it broadly applicable for acquiring training data in MVDC systems with converters featuring different control schemes.

The initial parameters of converters and lines for the five networks are listed in Table I. To obtain enough training data, random sampling for the three cases are also considered in the employed NR algorithm, following:

- Voltage reference values (p.u.) of all converters are uniformly varied from 0.9 to 1.1.
- For load-connected converters, the power/current reference values (p.u.) uniformly span from 0.8 to 1.2.
- For grid-connected, DG-connected, and ESS-connected converters, the power/current reference values (p.u.) are uniformly sampled between 0.9 and 1.1.
- Random sampling of line resistance values ranging from 0.95 p.u. to 1.05 p.u. has also been taken into account to enhance the robustness of the model.
- Any single lines/double lines are disconnected at random considering $N-1$ and $N-2$ contingency, respectively. At the same time, the tie switch is closed to maintain an uninterrupted power flow to the loads, which is integral to the subsequent monitoring of converter outages. Furthermore, as part of this process, each terminal is automatically checked during the sampling process to ensure that no independent terminals are present.
- Single or double grid-connected converter outages are included and associated lines are disconnected. To guarantee the NR algorithm operates correctly after the converter isolation, the converter-linked terminals are renumbered, ensuring sequential continuity of the terminal indices.

B. Neural Network Modelling

Unlike the conventional FCN, six original input channels are converted to three channels (voltage, current, and conductance) which are subsequently reformulated in the proposed PI-FCN. The output features are set as 64 in the first FCN encoder block, which are converted to 128, 256, 512 in the following blocks as shown in Fig. 1. In the encoder and residual blocks, the kernel sizes are set to 3×3 and 1×1, respectively. The output is produced using a 1×1 kernel, with both stride s and padding p configured to 1. ReLU activation function is adopted in each convolutional layer. Since MLP only accepts the input of one-dimensional data, the inputs of an MLP are n dc voltage, n dc current and $n \times n$ conductance values. The used MLP includes four hidden layers with in total 1000 neurons in all hidden layers.

A total of 20,000 samples within the specified data distribution are generated for each case, with 12,000 allocated for model training, 4000 reserved for validation, and another 4000 dedicated to testing the performance of all NNs. All NNs are trained for 1000 epochs using the Adam optimizer, with a learning rate of 0.001 and a batch size of 32. The network loss function employs MSE for each sample (n terminal voltage values are included), hence the calculation of total loss in all sample data follow:

$$Loss_{MSE}^{total} = \sum_{p=1}^{N_{data}} \left(\frac{1}{n} \sum_{i=1}^{n} \left(\frac{V_{dcm,i}^{act,p} - V_{dcm,i}^{tar,p}}{V_{dcm}^{rated}} \right)^2 \right), \quad (48)$$

where N_{data} is the number of sample data, V_{dcm}^{act} is the predicted MVDC voltage output from networks, and V_{dcm}^{tar} is the corresponding target values. Loss reduction is also used as monitor metric to trigger early stopping by a predefined patience setting of 20 epochs. The prediction accuracy A_{test}

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

TABLE I
SYSTEM PARAMETERS OF MT-MVDC NETWORKS.

Converter Parameters		**Converter 1**	**Converter 2**	**Converter 3**	**Converter 4**	**Converter 5**	**Converter 6**	**Converter 7**	**Converter 8**	**Converter 9**	**Converter 10**
V_{dcrefm} (kV),	Network 1	12, −, −	12, −, 0.65	12, −, -0.6	12, -8, −	/	/	/	/	/	/
P_{dcrefm} (MW),	Network 2	12, −, −	12, 8, −	12, −, -0.6	12, -8, −	12, −, 0.1	12, −, 0	/	/	/	/
I_{dcrefm} (kA)	Network 3	20, −, −	20, 15, −	20, −, -0.25	20, −, -0.25	20, −, -0.25	20, −, -0.25	20, -5, −	20, -5, −	/	/
	Network 4	20, −, −	20, 10, −	20, 10, −	20, −, -0.25	20, −, -0.25	20, −, -0.25	20, −, -0.25	20, -5, −	20, -5, −	/
	Network 5	70, −, −	70, −, 0.7	70, 50, −	70, −, -0.65	70, -45, −	70, −, -0.65	70, −, -0.15	70, -10, −	70, −, 0.15	70, −, 0
Control Method	Network 1	V	I/V (0.69)	I	P	/	/	/	/	/	/
	Network 2	V	P/V (8.33)	I	P	I/V (0.35)	I/V (0.35)	/	/	/	/
	Network 3	V	P/V (10)	I/V (0.2)	I/V (0.2)	I	I	P	P	/	/
	Network 4	V	P/V (7.5)	P/V (7.5)	I/V (0.2)	I/V (0.2)	I	I	P	P	/
	Network 5	V	I/V (0.12)	P/V (8.57)	I/V (0.10)	P/V (7.14)	I/V (0.10)	P	I	I/V (0.03)	I/V (0.02)
Line Parameters	**Line 1**	**Line 2**	**Line 3**	**Line 4**	**Line 5**	**Line 6**	**Line 7**	**Line 8**	**Line 9**	**Line 10**	**Line 11**
Distance, R (Ω/km), L (mH/km)	1	4, 0.1, 0.02	2, 0.1, 0.02	5, 0.1, 0.02	/	/	/	/	/	/	/	/
	2	4, 0.1, 0.02	5, 0.1, 0.02	5, 0.1, 0.02	4, 0.1, 0.02	3, 0.1, 0.02	/	/	/	/	/	/
	3	6, 0.1, 0.02	8, 0.1, 0.02	8, 0.1, 0.02	6, 0.1, 0.02	4, 0.1, 0.02	4, 0.1, 0.02	4, 0.1, 0.02	/	/	/	/
	4	6, 0.1, 0.02	8, 0.1, 0.02	8, 0.1, 0.02	8, 0.1, 0.02	8, 0.1, 0.02	6, 0.1, 0.02	4, 0.1, 0.02	4, 0.1, 0.02	4, 0.1, 0.02	/	/
	5	12, 0.1, 0.02	16, 0.1, 0.02	16, 0.1, 0.02	16, 0.1, 0.02	16, 0.1, 0.02	12, 0.1, 0.02	8, 0.1, 0.02	8, 0.1, 0.02	8, 0.1, 0.02	8, 0.1, 0.02	6, 0.1, 0.02

Note 1: The units of I/V and P/V droop constants are kA/kV and MW/kV, respectively.
Note 2: The listed line distance refers to the distance for each pole.

is determined by the proportion of data deemed qualified $N_{qualified}$ to the entire test set N_{total} (49). Test data is considered qualified when the error between the sum of actual MVDC voltage values predicted by the NNs and the provided samples is less than a predefined error $\mathcal{E}$, expressed as (50).

$$A_{test} = \frac{N_{qualified}}{N_{total}} \times 100\% \tag{49}$$

$$error_{sum}^{actual} = \left| \sum_{i=1}^{n} V_{dcm,i}^{act} - \sum_{i=1}^{n} V_{dcm,i}^{tar} \right| < \mathcal{E}, \tag{50}$$

which can be further converted into:

$$error_{sum}^{p.u.} = \left| \sum_{i=1}^{n} V_{dcm,i}^{act(p.u.)} - \sum_{i=1}^{n} V_{dcm,i}^{tar(p.u.)} \right| < \frac{\mathcal{E}}{V_{dcm}^{rated}} \tag{51}$$

if per unit calculations are adopted, and 0.01 p.u. error in each terminal is used that $\left| V_{dcm,i}^{act(p.u.)} - V_{dcm,i}^{tar(p.u.)} \right| < \mathcal{E}/(n \cdot V_{dcm}^{rated})$.

C. Fixed Network Topology (Case 1):

1) Trained Data Distribution Range: The topology of all MT-MVDC networks remains identical, where only the values of converter references, and line resistances are uniformly sampled. The five networks are trained separately and the prediction accuracy is verified in MLP, FCN and proposed PI-FCN. The comparison results are listed in Table II. The results indicate that the proposed PI-FCN outperforms other NNs, with its accuracy most notably improving as the number of iterations increases, especially in the 10T-MVDC system. The curves illustrating loss reduction in training/validation set for different NNs under 10T-MVDC systems over 1000 epochs are presented in Fig. 6. It reveals that the proposed PI-FCN converges more rapidly and achieves lower loss than both the conventional FCN and MLP.

Table III lists the training time comparison in one epoch. The weight-sharing characteristics of FCNs lead to a decrease in training time compared to the MLP, while the proposed PI-FCN requires more time due to the introduction of additional operation layers. During the PF prediction stage, all NNs have significant speed advantages compared to conventional NR method, as shown in Table IV. Despite the integration of physics operation in the forward PF prediction, this extra step only involves the simple solution of two linear equations (28) and (29), necessitating minimal additional time.

2) Untrained Data Distribution Range: To further demonstrate the superiority of physics-informed channel reformulation, a contrast experiment is conducted in both 4T-MVDC (Fig. 7) and 10T-MVDC (Fig. 8) networks. Three NNs are used in this experiment, including MLP, FCN and PI-FCN (PI-FCN refers to PI-FCN$^{(3)}$ in the subsequent description), to assess their performance under untrained sampling range based on MVDC voltage prediction errors $error_{sum}^{p.u.}$ (the sum of all dc voltages expressed in p.u. value).

The dc power/current references of load-connected or grid-connected converters are constant in each test, while other parameters are changed following the sampling method described in Section V-A. Figs. 7(a), 8(a) and Figs. 7(b), 8(b) show the prediction accuracy comparison (mean value of 4000 test samples), when the dc power/current references of load-connected converters ranges from 0.3 to 1.7 and dc power/current references of grid-connected converters change from 0.7 to 1.3 with 0.1 step, respectively. In both 4T- and 10T-MVDC distribution systems, all three NNs have acceptable performance within the trained areas, but the proposed PI-FCN has higher prediction accuracy in untrained areas.

TABLE II
PREDICTION ACCURACY COMPARISON OF DIFFERENT NNS UNDER CASE 1.

Network	4T-MVDC	6T-MVDC	8T-MVDC	9T-MVDC	10T-MVDC
MLP	97.5%	95.8%	91.4%	88.1%	88.2%
FCN	97.5%	94.7%	92.2%	90.0%	87.9%
PI-FCN$^{(1)}$	99.3%	99.5%	98.0%	98.2%	97.8%
PI-FCN$^{(2)}$	99.8%	99.8%	98.7%	98.6%	98.1%
PI-FCN$^{(3)}$	**99.9%**	**99.7%**	**99.1%**	**99.4%**	**98.7%**

$^{(1)}$: 1 iteration, $^{(2)}$: 5 iterations, $^{(3)}$: 20 iterations.

D. Varying Network Topology with Random Line Disconnection (Case 2):

1) Trained Line Disconnection: Random line disconnection is included in addition to the uniform sampling of converter

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

TABLE III
TRAINING TIME COMPARISON (ONE EPOCH).

Network	MLP	FCN	PI-FCN$^{(1)}$	PI-FCN$^{(2)}$	PI-FCN$^{(3)}$
4T-MVDC	4.01s 84.9it/s	2.98s 113.5it/s	3.32s 102.0it/s	3.40s 98.5it/s	4.23s 80.7it/s
10T-MVDC	7.20s 53.3it/s	5.93s 65.0it/s	6.18s 61.9it/s	6.49s 59.0it/s	7.91s 48.5it/s

TABLE IV
CALCULATION TIME COMPARISON (4000 SAMPLES) IN CPU.

Prediction Method		NR	MLP	FCN	PI-FCN$^{(1)}$	PI-FCN$^{(2)}$	PI-FCN$^{(3)}$
Time(s)	4T-MVDC	3.775	0.213	0.221	0.229	0.301	0.457
	10T-MVDC	23.879	0.450	0.502	0.526	0.692	0.983

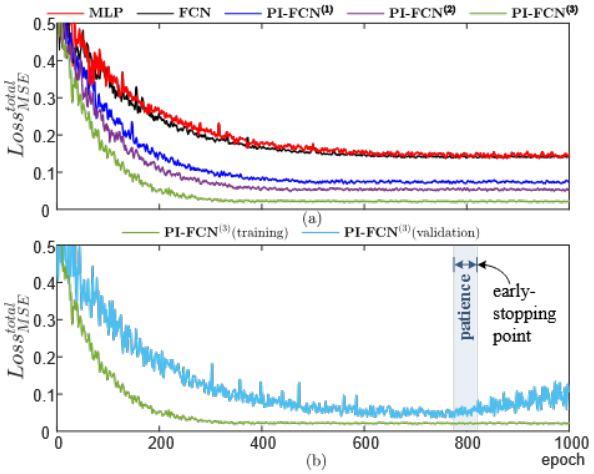


Fig. 6. Loss reduction in the training/validation set under 10T-MVDC: (a) training set in all NNs, and (b) training and validation sets in the PI-FCN with 20 iterations.

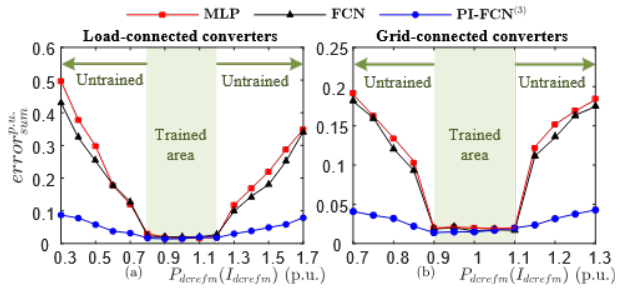


Fig. 7. Prediction error comparison for 4T-MVDC distribution system under different NNs: (a) dc power/current reference value changes of load-connected converters, and (b) dc power/current reference value variations of grid-connected converters.

references and line resistances. Single and double line disconnections are considered together to mimic $N-1$ and $N-2$ contingencies. Table V lists the prediction accuracy considering random single line disconnections in the five MVDC systems. The prediction accuracy of MLP and conventional FCN declines compared to fixed topology due to the increased training complexity. The proposed PI-FCN maintains good performance, despite a slight decrease in prediction accuracy.

2) Untrained Line Disconnection: A test for generalization capability is conducted, wherein the system is trained on

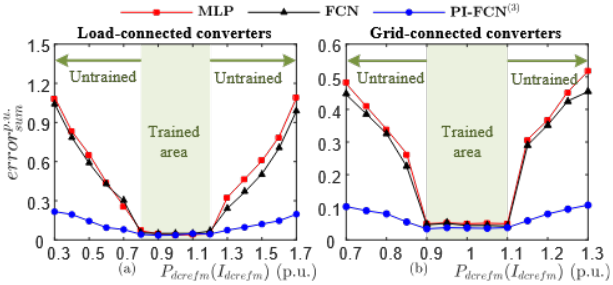


Fig. 8. Prediction error comparison for 10T-MVDC distribution system under different NNs: (a) dc power/current reference value changes of load-connected converters, and (b) dc power/current reference value variations of grid-connected converters.

single line disconnections and then tested on double line disconnections ($N-2$ contingency). The accuracy comparison for untrained double line disconnections is presented in Table VI, excluding 4T-MVDC and 6T-MVDC networks due to the presence of independent terminals for any double line disconnections. The results demonstrate the proposed PI-FCN has better generalization capability in the case of unseen double line disconnection scenarios.

Fig. 9 shows a detailed comparison of prediction errors (mean error of 4000 samples), examining the effects of single-line disconnection (line 1 or line 7) and concurrent disconnections of line 1 & line 3 or line 7 & line 11, within the 10T-MVDC system. It should be noted that to ensure the uninterrupted power supply to load 1 following the simultaneous tripping of line 1 and line 3, the tie switch connecting the two load terminals (T7 and T8) is closed. The proposed PI-FCN (green bars) outperforms others in both single and unseen double line tripping scenarios, owing to its integration of physics-based information. The increase in errors from single to double line disconnection of the proposed PI-FCN are also the smallest, being $\Delta error_{sum}^{p.u.} = 0.028$ for lines 1 and 3, and $\Delta error_{sum}^{p.u.} = 0.030$ for lines 7 and 11 disconnections, respectively. In contrast, the MLP exhibits the poorest performance in scenarios involving unseen double line disconnections ($error_{sum}^{p.u.}$ is increased to around 0.3) because it flattens the two-dimensional conductance matrix into one-dimensional inputs, a process which results in the loss of critical spatial structural information.

TABLE V
PREDICTION ACCURACY COMPARISON OF DIFFERENT NNS UNDER CASE 2 - SINGLE LINE DISCONNECTION.

Neural Network	4T-MVDC	6T-MVDC	8T-MVDC	9T-MVDC	10T-MVDC
MLP	92.4%	86.7%	82.8%	78.9%	78.3%
FCN	94.2%	90.7%	88.5%	84.9%	83.6%
PI-FCN	**99.5%**	**99.2%**	**98.6%**	**98.1%**	**98.2%**

TABLE VI
PREDICTION ACCURACY COMPARISON OF DIFFERENT NNS UNDER CASE 2 - UNTRAINED DOUBLE LINE DISCONNECTION.

Neural Network	8T-MVDC	9T-MVDC	10T-MVDC
MLP	76.5%	72.4%	68.2%
FCN	79.3%	76.5%	72.3%
PI-FCN	**93.4%**	**92.2%**	**91.4%**

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266

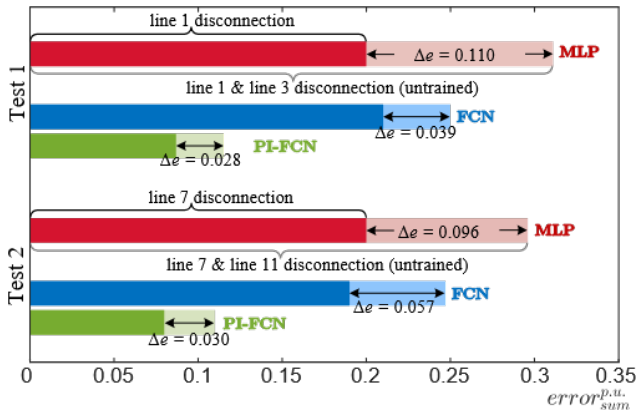


Fig. 9. Prediction error comparison for random line disconnection occurred in 10T-MVDC distribution system under different NNs.

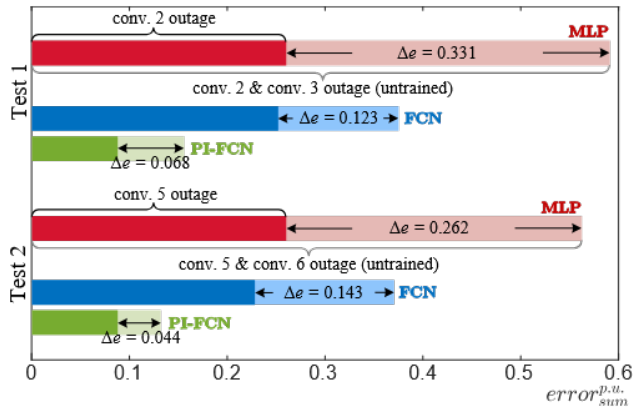


Fig. 10. Prediction error comparison for random converter outage occurred in 10T-MVDC distribution system under different NNs.

E. Varying Network Topology with Random Converter Outage (Case 3):

1) Trained Converter Outage: This case is the most complex because line disconnection and converter isolation are included in the training process together. During sampling, in addition to isolating single or double grid-connected converters ($N-1$ and $N-2$ contingencies), lines connected to isolated converters are also disconnected. Moreover, the outage of a dc voltage controlled-converter (converter 1) is not included during the sampling process for providing a voltage magnitude reference for the NR algorithm initialization. Similar to the NN performance comparison of consistent sampling in Case 2, Table VII presents the prediction accuracy for random single converter outages across five MVDC systems. Despite a decrease in the prediction accuracy of the proposed PI-FCN owing to greater training complexity, it still achieves the highest precision relative to the MLP and the conventional FCN.

2) Untrained Converter Outage: Table VIII lists the accuracy comparison for general untrained double converter outages. The proposed PI-FCN has the highest prediction accuracy in scenarios involving unseen double converter failures, especially in the 4T-MVDC system where only simultaneous outages of converters 1 and 3 or converters 2 and 4 are allowed.

Fig. 10 presents a comparison of prediction errors to assess the generalization ability of the proposed PI-FCN, which is trained on single-converter outages and tested on separate instances of individual converter outages (either converter 2 or converter 5) as well as on simultaneous outages of converters 2 and 3, or converters 5 and 6, in the 10T-MVDC system. The errors of all NNs significantly increase when tested on double converter outages, compared to those on unknown line disconnections, due to the substantial divergence in system structures between the training and test sets. However, the proposed PI-FCN still has the smallest errors although the prediction errors from single to double converter outage increase by 75.6% and 51.8% for converters 2 & 3, and for converters 5 & 6 outages, respectively.

TABLE VII
PREDICTION ACCURACY COMPARISON OF DIFFERENT NNS UNDER CASE 3 - TRAINED SINGLE CONVERTER OUTAGE.

Neural Network	4T-MVDC	6T-MVDC	8T-MVDC	9T-MVDC	10T-MVDC
MLP	90.4%	82.0%	76.8%	74.9%	73.0%
FCN	91.1%	84.1%	82.5%	78.5%	77.7%
PI-FCN	**98.6%**	**96.3%**	**95.8%**	**94.2%**	**93.5%**

TABLE VIII
PREDICTION ACCURACY COMPARISON OF DIFFERENT NNS UNDER CASE 3 - UNTRAINED DOUBLE CONVERTER OUTAGE.

Neural Network	4T-MVDC	6T-MVDC	8T-MVDC	9T-MVDC	10T-MVDC
MLP	72.7%	72.9%	69.8%	67.5%	66.5%
FCN	81.0%	80.3%	78.5%	74.4%	73.6%
PI-FCN	**95.8%**	**92.8%**	**91.6%**	**91.0%**	**90.2%**

VI. CONCLUSION

This paper proposes a novel PI-FCN with integrated physics guidance to predict the PF in MT-MVDC distribution networks, addressing the accuracy reduction issue in conventional NNs under unseen data distributions and topology structures. The PI-FCN simplifies its structure by merging various MVDC bus-type inputs into three channels using a channel combination layer. This reduces the NN computation burden and strengthens the feature extraction capability of PI-FCN by information concentration across different channels. In the presented physics-informed channel reformulation method, the physical relation between voltage, current, and line conductance in an MVDC network is embedded by reconstructing the input voltage and current matrices. It allows a more thorough extraction of information, elevates prediction accuracy and strengthens generalization capability. Three different cases, including fixed MVDC network topology, varying topologies with line disconnection and converter outage, demonstrate the improved prediction accuracy and the potential generalization ability of proposed PI-FCN.

The PI-FCN offers an effective FCN channel reconstruction method based on physical information for MT-MVDC PF prediction in different operation conditions. Given that MVDC is still in the early stages of commercial application, it is significant to identify the optimal operating conditions and strategize for the future expansion of existing projects. By

using the proposed PI-FCN, the efficiency and accuracy of decision-making processes can be enhanced to better facilitate the PF analysis in response to generation/load changes, renewable energy integration and grid upgrades in practical MVDC systems. The presented channel reconstruction method also offers crucial insights for future physics-informed NN-based PF estimation method in power system analysis.

REFERENCES

[1] CIGRE WG C6/B4.37, "Medium voltage DC distribution systems," CIGRE, Tech. Brochure 875, Jul. 2022.

[2] P. Simiyu et al., "Review of the DC voltage coordinated control strategies for multi-terminal VSC-MVDC distribution network," *The J. of Eng.*, vol. 2019, no. 16, pp. 1462–1468, Dec. 2018.

[3] P. Sun, G. Li, G. Town, and G. Konstantinou, "Identifying opportunities for medium voltage dc systems in australia," in *2022 IEEE PES 14th Asia-Pacific Power Energy Eng. Conf. (APPEEC)*, Nov. 2022, pp. 1–6.

[4] Y. Ji, Z. Yuan, J. Zhao, C. Lu, Y. Wang, Y. Zhao, Y. Li, and Y. Han, "Hierarchical control strategy for MVDC distribution network under large disturbance," *IET Generation, Transmission & Distribution*, vol. 12, no. 11, pp. 2557–2565, May 2018.

[5] A. Garces, "On the convergence of newton's method in power flow studies for DC microgrids," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5770–5777, Sep. 2018.

[6] E. Aprilia, K. Meng, M. Al Hosani, H. H. Zeineldin, and Z. Y. Dong, "Unified power flow algorithm for standalone ac/dc hybrid microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 639–649, Jan. 2019.

[7] S. Petridis, O. Blanas, D. Rakopoulos, F. Stergiopoulos, N. Nikolopoulos, and S. Voutetakis, "An efficient backward/forward sweep algorithm for power flow analysis through a novel tree-like structure for unbalanced distribution networks," *Energies*, vol. 14, no. 4, Feb. 2021.

[8] P. Sun, R. Wu, G. Li, M. Khalid, G. Town, and G. Konstantinou, "Decoupled sequential power flow study in MT-MVDC distribution systems based on novel NR/estimation-correction algorithm," in *2023 11th Int. Conf. Power Electron. ECCE Asia (ICPE 2023 - ECCE Asia)*, May 2023, pp. 1464–1469.

[9] M. Karimi et al., "Application of Newton-based load flow methods for determining steady-state condition of well and ill-conditioned power systems: A review," *Int. J. of Elect. Power & Energy Syst.*, vol. 113, pp. 298–309, Dec. 2019.

[10] Y. Chen, C. Wu, and J. Qi, "Data-driven power flow method based on exact linear regression equations," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 3, pp. 800–804, May 2022.

[11] V. Paucar and M. J. Rider, "Artificial neural networks for solving the power flow problem in electric power systems," *Elect. Power Syst. Res.*, vol. 62, no. 2, pp. 139–144, Jun. 2002.

[12] R. Nellikkath and S. Chatzivasileiadis, "Physics-informed neural networks for AC optimal power flow," *Elect. Power Syst. Res.*, vol. 212, p. 108412, Nov. 2022.

[13] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, 2018.

[14] P. Simiyu, A. Xin, K. Wang, G. Adwek, and S. Salman, "Multiterminal medium voltage DC distribution network hierarchical control," *Electronics*, vol. 9, no. 3, Mar. 2020.

[15] Y. Tan, Y. Chen, Y. Li, and Y. Cao, "Linearizing power flow model: A hybrid physical model-driven and data-driven approach," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 2475–2478, May 2020.

[16] 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.

[17] J. Zhang, Z. Wang, X. Zheng, L. Guan, and C. Y. Chung, "Locally weighted ridge regression for power system online sensitivity identification considering data collinearity," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1624–1634, Mar. 2018.

[18] P. Li, W. Wu, X. Wang, and B. Xu, "A data-driven linear optimal power flow model for distribution networks," *IEEE Trans. Power Syst.*, vol. 38, no. 1, pp. 956–959, Jan. 2023.

[19] J. Yuan and Y. Weng, "Support matrix regression for learning power flow in distribution grid with unobservability," *IEEE Trans. Power Syst.*, vol. 37, no. 2, pp. 1151–1161, Mar. 2022.

[20] A. Karami and M. Mohammadi, "Radial basis function neural network for power system load-flow," *Int. J. of Elect. Power & Energy Syst.*, vol. 30, no. 1, pp. 60–66, Jan. 2008.

[21] 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.

[22] W. Dong, Z. Xie, G. Kestor, and D. Li, "Smart-pgsim: Using neural network to accelerate ac-opf power grid simulation," in *Int. Conf. High Perform. Comput., Netw., Storage Anal.*, Nov. 2020, pp. 1–15.

[23] X. Pan, T. Zhao, M. Chen, and S. Zhang, "DeepOPF: A deep neural network approach for security-constrained DC optimal power flow," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 1725–1735, May 2021.

[24] M. K. Singh, V. Kekatos, and G. B. Giannakis, "Learning to solve the AC-OPF using sensitivity-informed deep neural networks," *IEEE Trans. Power Syst.*, vol. 37, no. 4, pp. 2833–2846, Jul. 2022.

[25] X. Pan, M. Chen, T. Zhao, and S. H. Low, "DeepOPF: A feasibility-optimized deep neural network approach for AC optimal power flow problems," *IEEE Syst. J.*, vol. 17, no. 1, pp. 673–683, Mar. 2023.

[26] R. Khalitov, T. Yu, L. Cheng, and Z. Yang, "Chordmixer: A scalable neural attention model for sequences with different lengths," 2023.

[27] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, Jun. 2022.

[28] Y. Zhou, W.-J. Lee, R. Diao, and D. Shi, "Deep reinforcement learning based real-time AC optimal power flow considering uncertainties," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 5, pp. 1098–1109, Sep. 2022.

[29] Y. Du, F. Li, J. Li, and T. Zheng, "Achieving 100x acceleration for N-1 contingency screening with uncertain scenarios using deep convolutional neural network," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3303–3305, Jul. 2019.

[30] Y. Jia and X. Bai, "A CNN approach for optimal power flow problem for distribution network," in *2021 Power System and Green Energy Conference (PSGEC)*. IEEE, Aug. 2021.

[31] Y. Jia, X. Bai, L. Zheng, Z. Weng, and Y. Li, "ConvOPF-DOP: A data-driven method for solving AC-OPF based on CNN considering different operation patterns," *IEEE Trans. Power Syst.*, vol. 38, no. 1, pp. 853–860, Jan. 2023.

[32] D. Wang, K. Zheng, Q. Chen, X. Zhang, and G. Luo, "A data-driven probabilistic power flow method based on convolutional neural networks," *International Transactions on Electrical Energy Systems*, vol. 30, no. 7, p. e12367, Mar. 2020.

[33] F. Aksan, Y. Li, V. Suresh, and P. Janik, "CNN-LSTM vs. LSTM-CNN to predict power flow direction: A case study of the high-voltage subnet of northeast germany," *Sensors*, vol. 23, no. 2, p. 901, Jan. 2023.

[34] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, May 2021.

[35] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *2015 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 3431–3440.

[36] P. Sun, R. Wu, Z. Shen, G. Li, M. Khalid, G. Town, and G. Konstantinou, "Sequential power flow algorithm and post-event steady-state power distribution analysis in hybrid AC/MT-MVDC systems," *Int. J. of Elect. Power & Energy Syst.*, vol. 157, p. 109828, Jun. 2024.

[37] S. Gao, Y. Chen, S. Huang, and Y. Xia, "Efficient power flow algorithm for AC/MTDC considering complementary constraints of VSC's reactive power and AC node voltage," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 2481–2490, May 2021.

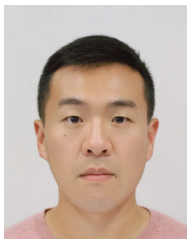
[38] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," 2015.

[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778.

[40] 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, Mar. 2023.

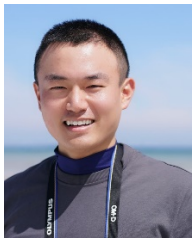
[41] S. Imambi, K. B. Prakash, and G. Kanagachidambaresan, "Pytorch," *Programming with TensorFlow: Solution for Edge Computing Applications*, pp. 87–104, 2021.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3382266



Pingyang Sun (Graduate Student Member, IEEE) received the B.Eng. degree in Electrical Engineering from Qilu University of Technology (Shandong Academy of Sciences), Jinan, China, in 2018, the M.Phil degree in electrical engineering at the School of Electrical engineering and Telecommunications, the University of New South Wales (UNSW), Sydney, Australia, in 2021. He is currently pursuing his Ph.D. degree at UNSW. He was an exchange student at the School of Electronics and Information, Kyung Hee University (KHU), Suwon, Korea, in 2017, and a guest Ph.D student at the Technical University of Denmark (DTU), Ballerup, Denmark, in 2024. His research interests include system modeling and benchmarking from MVDC to HVDC systems.

Rongcheng Wu (Student Member, IEEE) received the Master of Philosophy (M.Phil) from UNSW and is currently a PhD candidate at UTS, specializing in computer vision and AI security, with a particular focus on 3D vision and defense against backdoors in AI models. Prior to embarking on his PhD journey at UTS, he served as a research assistant in the UTS Data Science group, where he gained valuable experience and insights into data-driven technologies and methodologies.



Hongyi Wang (Student Member, IEEE) received the M.Sc. degrees in electrical engineering from the Shenyang University of Technology, Shenyang, China, in 2019. He is currently working toward the Ph.D. degrees in energy technology from the Department of AAU Energy, Aalborg University, Aalborg Denmark. His research interests include nonlinear dynamic system modeling for power electronic converter-based systems, Aggregation modeling for large-scale wind farms, dictionarty learning, digital twins for power electronic converter-based systems.

Gen Li (Senior Member, IEEE) received the B.Eng. degree in Electrical Engineering from Northeast Electric Power University, Jilin, China, in 2011, the M.Sc. degree in Power Engineering from Nanyang Technological University, Singapore, in 2013 and the Ph.D. degree in Electrical Engineering from Cardiff University, Cardiff, U.K., in 2018.

He is now an Associate Professor at the Technical University of Denmark (DTU), Denmark. From 2013 to 2016, he has been a Marie Curie Early Stage Research Fellow funded by the European Commission's MEDOW project. He has been a Visiting Researcher at China Electric Power Research Institute and State Grid Smart Grid Research Institute, China, at Elia, Belgium and at Toshiba International (Europe), U.K. He was a Research Associate at the School of Engineering, Cardiff University from 2018 to 2022. His research interests include control and protection of HV and MV DC technologies, offshore wind, offshore energy islands, reliability modelling and evaluation of power electronics systems.

Dr. Li is a Chartered Engineer in the U.K., a Young Editorial Board Member of Applied Energy, an Associate Editor of the CSEE Journal of Power and Energy Systems and IET Energy Systems Integration, an Editorial Board Member of CIGRE ELECTRA and Global Energy Interconnection and an IET Professional Registration Advisor. His Ph.D. thesis received the First CIGRE Thesis Award in 2018. He is now a Steering Committee Member of CIGRE Denmark NGN, a Member of CIGRE Working Group B4.96 and a Member of IEEE PELS Publicity Committee.

Muhammad Khalid (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the School of Electrical Engineering Telecommunications (EET), Australia, in 2011. He was initially a Postdoctoral Research Fellow, for three years, and then he continued as a Senior Research Associate with the Australian Energy Research Institute, School of EET, UNSW, for two years. He is currently an Associate Professor with the Electrical Engineering Department, King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia. He is also a Research Affiliate with the Interdisciplinary Research Center for Sustainable Energy Systems, KFUPM. He has authored/coauthored several journals and conference papers in the field of control and optimization for renewable power systems. His research interests include the optimization and control of battery energy storage systems for large-scale grid-connected renewable power plants (particularly wind and solar), cyber physical systems, networked control system, distributed power generation and dispatch, hybrid energy storage, hydrogen systems, EVs, AI, machine learning, and smart grids. He was a recipient of the Highly Competitive Postdoctoral Writing Fellowship from UNSW, in 2010. He was a recipient of many academic awards and research fellowships. In addition, he has been a reviewer of numerous international journals and conferences.

Georgios Konstantinou (Senior Member, IEEE) received the B.Eng. degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2007 and the Ph.D. degree in electrical engineering from UNSW Sydney (The University of New South Wales), Australia, in 2012. From 2013 to 2016, he was a Senior Research Associate with the University of New South Wales, Sydney, NSW, Australia, where he was part of the Australian Energy Research Institute. Since 2017, he has been with the School of Electrical Engineering and Telecommunications, UNSW Sydney, where he is currently an Associate Professor with the Energy Systems Research Group and the Real-time Simulations Laboratory (RTS@UNSW). He is an Associate Editor for IEEE Transactions on Power Electronics.