A Physics-Guided Graph Convolution Neural Network for Optimal Power Flow

Maosheng Gao ⓘ, Juan Yu ⓘ, *Senior Member, IEEE*, Zhifang Yang ⓘ, *Member, IEEE*, and Junbo Zhao ⓘ, *Senior Member, IEEE*

Abstract—The data-driven method with strong approximation capabilities and high computational efficiency provides a promising tool for optimal power flow (OPF) calculation with stochastic renewable energy. However, the topology change dramatically increases the learning difficulties and the demand for learning samples. In this work, we propose a physics-guided graph convolution neural network (GCNN) for OPF calculation with consideration of varying topologies, including the physics-guided graph convolution kernel, feature construction, and loss function formulation. Specifically, a physics-embedded graph convolution kernel is derived by aggregating the features from local neighborhoods utilizing the nodal OPF model formulation. An iterative feature construction method is also developed that encodes both the physical feature and practical constraints into the node vector. Finally, a correlative learning loss function to optimize the unbalanced power injection is developed. Extensive numerical results carried out on various IEEE test systems show that the prediction accuracy of OPF using the proposed method under varying topology changes can be improved by an average of 13.30% and up to 32.63% compared with state-of-the-art methods.

Index Terms—Graph convolution neural network (GCNN), neighborhood aggregation, optimal power flow (OPF), varying topology.

I. INTRODUCTION

OPTIMAL power flow (OPF) plays an essential role in power system operation. It is widely used for power system planning, marketing clearing, economic dispatch, etc. The OPF problem is a nonconvex optimization problem with NP-hard computational complexity [1], [2]. For application scenarios like the probability analysis, the OPF calculation needs to be executed an enormous number of times on randomly generated samples to get the probability statistics. The huge computational burden makes it difficult for practical applications and this calls for new approach developments. Existing studies propose to accelerate the OPF calculation of a single sample, such as simplifying the OPF model [3], [4], using heuristic algorithms [5], [6], or advanced numerical algorithms [7], [8]. However, the accumulated computational burden in the probabilistic analysis is still high.

With the development of artificial intelligence technologies, there is an increasing interest in using them to efficiently solve the OPF problem with significantly enhanced computational efficiency [9]. For example, the stacked extreme learning machine [10], deep neural network (DNN) [11], [12], convolution neural network (CNN) [13], and graph convolution neural network (GCNN) [14] are exploited as a data-driven proxy to the OPF calculation. Those methods settle the heavy computational burden in two different approaches, the direct learning approach, and the indirect learning approach.

The direct learning approach is to predict the solutions by learning the mapping from the load and renewable energy to the OPF outcomes [15]. The direct learning approach can yield 20-100 times faster than the conventional optimization-based methods [11], [12]. In addition to the neural network (NN), other machine learning methods, such as the Gaussian process [16], are also applied to predict the OPF solutions. To ensure the satisfaction of practical constraints like the limits of power generations and the transmission limit of power lines, [11] and [17] construct a new loss function with the regularization part derived from the Karush-Kuhn-Tucker condition or power flow model.

The indirect learning approach is to speed up the conventional solving algorithm by predicting the active constraints or setting the warm start point. [15] uses the NN to predict the solution of OPF and sets the warm-start point to find the nearby feasible solution rather than directly output the solution. [18] predicts active and inactive inequality constraints before solving AC OPF. Taking advantage of the nature that few constraints could be active, solving the size-reduced OPF problem is much faster than the original one.

However, for those data-driven OPF methods, a key concern is whether they can provide a reliable solution given varying operating conditions, especially changing topologies. This brings discrete changes to the projections between the input and output of the data-driven OPF calculations. To cope with the complexity considering varying topology, model-informed feature engineering, and meta-learning are helpful approaches. Instead of encoding topology via a binary input vector, [19] takes the difference of voltage magnitude between the original

TABLE I
THE RELATED WORKS CONSIDERING THE VARYING TOPOLOGY

Methods	Description	Potential Improvements
Topology feature engineering	Regarding the diagonal element of susceptance matrix as topology feature [20].	Constructed features do not consider the physics relationship.
	Regarding the magnitude and angle of impedance as topology feature [13].	
	Regarding the voltage difference between topology as topology feature [19].	
Meta-learning	Using the meta-learning algorithm to find good initialization points to accelerate new topology learning [21].	The volume of training data is huge.
Topology embedded GCNN	Embedding the topology with the adjacent matrix [22].	The graph convolution kernel is not designed following the physical model of OPF.
	Exploiting the impedance-based Gaussian kernel as graph convolution kernel [14].	
	Transforming the impedance-based Gaussian kernel based on spectral convolution theory [23].	

topology and a new topology as the input topology feature. [20] regards the diagonal element of the admittance matrix as a topology feature and [13] takes advantage of the magnitude and angle of the diagonal elements of the admittance matrix. [21] treats the learning of mapping under a given topology as an independent task and applies the meta-learning algorithm to find good initialization points to accelerate new topology learning.

Although the topology feature engineering and the meta-leaning approach make it possible to predict the OPF solution or active constraints under varying topologies, there are still potential improvements. Specifically, the aforementioned feature engineering has feature loss and it is impossible to input the complete topology information due to the size of the feature dimension. Besides, directly concatenating of load vector and engineered topology feature vector does not reveal the complicated physics relationship between load and the topology. As a result, the combinatory explosion of topology exists and only a few topologies can be considered in practical applications. Instead of topology feature engineering, the GCNN uses the built-in topology to extract the topology feature and is promising to achieve the data-driven approximation of OPF considering topology change. The topology embedding and feature extraction are not limited to the input topology information but extracting the valuable information during the training process. Therefore, the GCNN is a promising tool to address the data-driven OPF problem considering varying topologies.

Some studies have built the GCNN to solve the OPF problem. [22] embeds the topology into the GCNN and exploits the adjacent matrix-based graph convolution kernel to extract the topology feature. On account of the topological representation capability of the physical variable like impedance matrix and admittance matrix, the correlated feature between loads and topologies can be effectively learned. [14] also constructs a topology-embedded GCNN to predict the OPF solutions, which replaces the graph convolution kernel with an impedance-based Gaussian kernel. [23] transforms the impedance-based Gaussian kernel in [14] based on spectral convolution theory to a new convolution kernel. However, the linear graph convolution in [14], [22], and [23] ignores the coupling and the non-linear characteristic of the OPF problem. As a result, their prediction accuracy cannot be guaranteed. To conclude, the characteristics

and potential improvements of the related works are summarized in Table I.

To this end, a physics-guided GCNN for OPF considering topology feature learning is proposed and the main contributions are summarized as follows:

1) A novel physics-embedded aggregating method is proposed to construct the graph convolution kernel, which improves the prediction accuracy. This is achieved by decomposing the AC power flow equations based on Gaussian-Seidel iteration, followed by the physics-embedded neighborhood aggregation formulation. This allows the feature propagation in the NNs to follow the physical law, yielding reduced learning difficulty. Thanks to the iteration of the physics-embedded aggregation, the nodal feature vector can be obtained by encoding the topology and inter-relationship between loads. In addition, to sufficiently consider the physical constraints, the clipping technique is used during the feature construction.

2) A new constrained loss function is proposed to consider the physical correlations among the outputs. It is an extended-term that describes the coupling relationship among physical variables, e.g., node voltage and power generation. This allows us to restrict the outliers that do not obey the physical relationship and therefore improve the overall accuracy effectively.

The effectiveness of the proposed method is demonstrated in several IEEE and Polish test systems. Compared to other data-driven OPF methods, the proposed method has stronger topology adaptability and can more accurately predict the OPF solution. The average probabilistic prediction accuracy can be improved by an average of 13.30% and up to 32.63% compared to the state-of-art data-driven methods, like DNN and CNN.

II. PROBLEM FORMULATION

A. Problem Description

The OPF problem is to decide the optimal control of power equipment in the power system, like the generation, voltage magnitude, and operational state (on/off) of power generators. It would involve mixed-integer decisions or multiple time-step

optimizations. While adding new technologies such as the management of renewable resources, the OPF problem would be more complicated and bring unseen injection distribution and topology. But they share the same modeling basis with the power flow constraints and branch flow limits, which is named the general OPF problem. In this paper, we mainly consider the general OPF problem and study a basic framework for the complicated one. The general OPF problem is to find the low-cost and feasible set-points of generators while satisfying the equality and inequality constraints. The mathematical formulation is shown as:

$$\min C = \sum_{i \in S_G} \left(a_i PG_i^2 + b_i PG_i + c_i \right), \tag{1}$$

$$\left.\begin{array}{l} PG_i - PD_i = \Gamma\left(V, \theta, z\left(G\right), z\left(B\right)\right) \\ QG_i - QD_i = \Psi\left(V, \theta, z\left(G\right), z\left(B\right)\right) \end{array}\right\} (i, j \in S_B), \tag{2}$$

$$\left.\begin{array}{l} \underline{PG_i} \le PG_i \le \overline{PG_i} \\ \underline{QG_i} \le QG_i \le \overline{QG_i} \end{array}\right\} (i \in S_G), \tag{3}$$

$$\underline{V_i} \le V_i \le \overline{V_i} \,(i \in S_B), \tag{4}$$

$$\left.\begin{array}{l} PL_{ij} = V_i V_j (z(G)_{ij} \cos\theta_{ij} + z(B)_{ij} \sin\theta_{ij}) \\ \qquad\qquad - V_i^2 z(G)_{ij} \\ -\underline{PL_{ij}} \le PL_{ij} \le \overline{PL_{ij}} \end{array}\right\} (i, j \in S_B), \tag{5}$$

where PG_i and VG_i are the control variables for the ith generator representing the active power generation and voltage magnitude, respectively; a_i, b_i and c_i are the generation cost coefficients for ith generator; PD_i and QD_i are the active and reactive power demand at bus i; V_i and θ_i are the voltage magnitude and phase angle at bus i; θ_{ij} is the phase angle difference between ith and jth bus; G_{ij} and B_{ij} denote the conductance and susceptance between the ith and jth bus, respectively; $z(.)$ represents the topology change function; $\Gamma(.)$, $\Psi(.)$ are the power flow equations; PL_{ij} represents the active branch power between ith node and jth node; S_G, S_B and S_K are the index set of generators, buses and branches.

Differing from the general OPF model, the varying topology based on the default one is considered in the formulation. The topology changes function $z(.)$ indicates the topology of the power system can be occasionally different because of the power line contingency or regular maintenance. It operates on the original conductance or susceptance matrix and the result is still the conductance or susceptance matrix referring to the varying topology. Hence, in (5), $z(G)_{ij}$ and $z(B)_{ij}$ represent the element in the ith row and jth column of $z(G)$ and $z(B)$.

B. Framework of the Proposed Method

The data-driven methods in solving the OPF problem can be viewed as identifying a mapping from the load fluctuation, uncertain renewable energy, and topology to the set-points of generators. They approximate this mapping from the historic or simulated data. Hence, when solving the OPF problem considering varying topology, the conventional data-driven methods input the feature vector containing the topology representation to a NN or embed the topology with a linear relationship into the
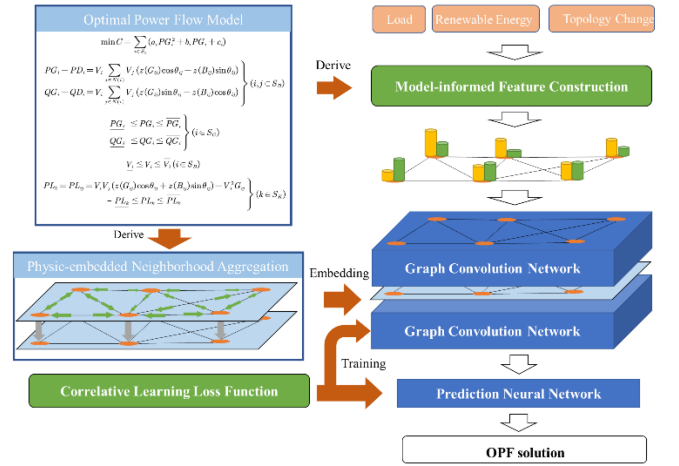

Fig. 1. The framework of the proposed method considering varying topology.

NN structure. After the projection from the input via the NN, the final OPF solution is obtained. However, since the nonlinearity and coupling physical relationship under different topologies are not input or embedded into the NN, there could be large biases in the OPF solutions.

In the proposed method, the feature construction, NN structure, and learning loss are improved according to the OPF formulation to reduce learning complexity, see Fig. 1. Firstly, the model-informed feature construction method is proposed. It encodes the loads, renewable energy, and topology into a unique feature space, where the physical relationship and the correlated topology are included. Secondly, the physics-embedded NN is designed to embed the topology and include the coupling physical relationships. While the unique features are input to the NN, the physics-embedded graph convolution can further aggregate neighborhoods' features and convolute them to a new high-dimensional space. This embedding of non-linear and coupling physical relationships can significantly reduce the learning complexity. Lastly, the correlative learning loss function is used after the input feature and NN structure are determined. It calculates the unbalance power injection at each node and back-propagates them to enhance the physical relationship with the outputs.

III. PROPOSED PHYSICS-GUIDED GCNN

The section briefly introduces the GCNN. Then, the physics-embedded graph convolution and the physics-guided GCNN structure with physics-embedded graph convolution are proposed. Besides, the model-informed features are constructed.

A. A Brief Introduction of GCNN

GCNN is one kind of NN operating over graph-structured data $G = (\mathcal{V}, \mathcal{E})$ where G represents the graph data; $\mathcal{V} \subset R^n$ and $\mathcal{E} \subset R^b$ are the vertex set and edge set. To describe the relationship among different nodes, matrix $A \in R^{n \times n}$ denotes the weighted adjacent matrix. Let $a_{ij} \ne 0$ while node i and node j are connected, and otherwise, $a_{ij} = 0$. The input feature of

each node is $\boldsymbol{X} \in R^{n \times m}$. Then the graph convolution in a layer can be shown as:

$$Y = f(AXW + B), \qquad (6)$$

where $\boldsymbol{Y} \in R^{n \times k}$ is the output matrix of a graph convolution layer; $\boldsymbol{W} \in R^{m \times k}$ is the trainable parameter matrix; $\boldsymbol{B} \in R^k$ is the trainable bias matrix; $f(\cdot)$ is the activation function.

In (6), $\boldsymbol{A}$ can be constant or trainable, but to decrease the size of trainable parameters, $\boldsymbol{A}$ usually utilizes the constant form derived from the adjacent matrix. For the power system problems, there are physical variables in the branch, and $\boldsymbol{A}$ can be defined based on the physical impedance. $\boldsymbol{A}$ presents the weights that extract the information from adjacent nodes and $\boldsymbol{AX}$ denotes the extraction process named neighborhood aggregation. When the relationship of adjacent nodes is non-linear and cannot be accurately extracted by matrix multiplication, the neighborhood aggregation can be represented by a function $\phi(\textbf{.})$. So (6) can be rewritten as:

$$Y = f(\phi(X, A)W + B). \qquad (7)$$

In the OPF problem, the nature of non-linearity and non-convexity indicates that the neighborhood aggregation with a single weighted matrix cannot effectively extract features from adjacent nodes. Besides, the physical variables in branches like conductance and susceptance are coupled with node voltage and power. It makes feature extraction difficult but also provides an opportunity to extract features by utilizing the coupled relationship and integrating the physical law into the neighborhood aggregation process.

B. Physics-Embedded Graph Convolution

For the OPF model in Section II, (1) is the objective function and (2)–(5) are the constraints. Especially, (2) is the AC power flow equations that lead to complex non-linearity. The power flow equations also include topology information and reveal the internal physical relationship between the adjacent nodes. Taking (2) as the feature transformation function, if the power flow equations can be embedded into the neighborhood aggregation, the NN will be capable of efficiently extracting the topology information. By setting $e_i = V_i \cos \theta_i$ and $f_i = V_i \sin \theta_i$ as the nodal feature in the graph and writing the power flow equations in the Cartesian coordinate system, (2) can be rewritten as:

$$PG_i - PD_i = e_i \sum_{j \in N(i)} \left(z(G)_{ij} e_j - z(B)_{ij} f_j \right)$$
$$+ f_i \sum_{j \in N(i)} \left(z(G)_{ij} f_j + z(B)_{ij} e_j \right), \qquad (8)$$

$$QG_i - QD_i = f_i \sum_{j \in N(i)} \left(z(G)_{ij} e_j - z(B)_{ij} f_j \right)$$
$$- e_i \sum_{j \in N(i)} \left(z(G)_{ij} f_j + z(B)_{ij} e_j \right). \qquad (9)$$

Based on the Gaussian-Seidel iteration that separates the diagonal elements of the admittance matrix, there are two kinds of features in the feature transformation function. One is the self-transformed feature which is only affected by the feature at the central node i, such as $M_i = (e_i^2 + f_i^2) z(G)_{ii}$ which is the sum of the components with G_{ii} and B_{ii} in (8). The others are the extracted features from adjacent nodes, such as $M_j = e_i(z(G)_{ij} e_j - z(B)_{ij} f_j)$ which indicates the extracted feature from the node j. Hence, by moving the self-transformed feature to the left side and keeping the extracted features on the right side, the feature transformation function can be rewritten as:

$$\delta_i = e_i \alpha_i + f_i \beta_i, \qquad (10)$$

$$\lambda_i = f_i \alpha_i - e_i \beta_i, \qquad (11)$$

where:

$$\alpha_i = \sum_{j \in N(i), j \neq i} \left(z(G)_{ij} e_j - z(B)_{ij} f_j \right), \qquad (12)$$

$$\beta_i = \sum_{j \in N(i), j \neq i} \left(z(G)_{ij} f_j + z(B)_{ij} e_j \right), \qquad (13)$$

$$\delta_i = PG_i - PD_i - \left(e_i^2 + f_i^2 \right) z(G)_{ii}, \qquad (14)$$

$$\lambda_i = QG_i - QD_i + \left(e_i^2 + f_i^2 \right) z(B)_{ii}. \qquad (15)$$

In (12) and (13), α_i, β_i are the aggregated feature from all the neighbors. δ_i, λ_i are the total self-transformed features at node i. Hence, to complete the aggregation and obtain a new feature state, the coupled relationship between (10) and (11) can be simplified by solving the linear equations while assuming e_i^{l+1}, f_i^{l+1} are the new feature state at node i. Eventually, the function to aggregate features from adjacent nodes under a given topology is obtained as:

$$e_i^{l+1} = \frac{\delta_i \alpha_i - \lambda_i \beta_i}{\alpha_i^2 + \beta_i^2}, \qquad (16)$$

$$f_i^{l+1} = \frac{\delta_i \beta_i + \lambda_i \alpha_i}{\alpha_i^2 + \beta_i^2}. \qquad (17)$$

In the OPF problem, the generation of active power and reactive power are unknown. In the meanwhile, the graph convolution is to extract the important feature. Hence, the generations of active power and reactive power at all nodes are set to be zeros. By substituting (16) and (17) into (7), the physic-embedded graph convolution is obtained.

$$Y = \begin{bmatrix} e^{l+1} \\ f^{l+1} \end{bmatrix} = f \left(\begin{bmatrix} \frac{\delta \odot \alpha - \lambda \odot \beta}{\alpha \odot \alpha + \beta \odot \beta} & 0 \\ 0 & \frac{\delta \odot \beta + \lambda \odot \alpha}{\alpha \odot \alpha + \beta \odot \beta} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \right), \qquad (18)$$

where $\odot$ is the Hadamard (entry-wise) product; $\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{B}_1$ and $\boldsymbol{B}_2$ are the trainable parameters in the graph convolution layer; e^{l+1}, f^{l+1} are the features matrix in the $l+1$ layer; and:

$$\boldsymbol{\alpha} = z(\boldsymbol{G}_{\text{ndiag}}) \boldsymbol{e}^l - z(\boldsymbol{B}_{\text{ndiag}}) \boldsymbol{f}^l, \qquad (19)$$

$$\boldsymbol{\beta} = z(\boldsymbol{G}_{\text{ndiag}}) \boldsymbol{f}^l + z(\boldsymbol{B}_{\text{ndiag}}) \boldsymbol{e}^l, \qquad (20)$$

$$\boldsymbol{\delta} = -\boldsymbol{PD} - \left(\boldsymbol{e}^l \odot \boldsymbol{e}^l + \boldsymbol{f}^l \odot \boldsymbol{f}^l \right) z(\boldsymbol{G}_{\text{diag}}), \qquad (21)$$

$$\boldsymbol{\lambda} = -\boldsymbol{QD} - \left(\boldsymbol{e}^l \odot \boldsymbol{e}^l + \boldsymbol{f}^l \odot \boldsymbol{f}^l \right) z(\boldsymbol{B}_{\text{diag}}), \qquad (22)$$

where G_{ndiag}, B_{ndiag} are the admittance and susceptance matrices without diagonal elements; e^l, f^l are the features matrices in the l layer; G_{diag}, B_{diag} are the diagonal matrices of the admittance and susceptance matrices.

In (18), the non-linear power flow and topology features are embedded in the forward-propagation of a graph convolution layer. Therefore, the physic-embedded graph convolution that can sufficiently extract topology features and physical features, is developed.

C. Model-Informed Feature Construction

In the data-driven OPF method, despite the physics-embedded graph convolution can significantly reduce the learning difficulty, the data processing including feature construction also affects the prediction performance. Hence, this paper presents a model-informed feature construction strategy.

Based on the physics-embedded neighborhood aggregation method in the graph convolution, the proposed feature construction strategy encodes both the topology and the physical features including the physical relationship as well as the practical constraints like the limitation of the generators. In the solving process of the OPF problem, the voltage magnitude, phase angle, and active and reactive power generation are unknown. But in the aggregation function, all the variables of the initial state should be known. Hence, based on the property of the power system, the voltage magnitude and phase angle for all nodes are set to be 1 and 0. Then, the PG and QG can be calculated by (8) and (9).

However, when directly substituting the PG and QG calculated by (8) and (9) into the neighborhood aggregation function, the main feature of load and renewable energy can be omitted because of the identical relationship of PG and QG in both the aggregation function and power flow equations. Therefore, before substituting calculated PG and QG into the aggregation function, the practical limitations of the power generation in all nodes are considered via (23) and (24).

$$PG_i = \max\left(\min\left(PG_i, \overline{PG_i}\right), \underline{PG_i}\right). \tag{23}$$

$$QG_i = \max\left(\min\left(QG_i, \overline{QG_i}\right), \underline{QG_i}\right). \tag{24}$$

After aggregation using the limited PG and QG, the new nodal features are obtained. But, a single neighborhood aggregation merely assembles the adjacent features. To assemble the feature from the entire graph, an iterative feature construction strategy is proposed in Fig. 2. With the increase in iteration number, the features extracted can reflect the influence of other nodes far away from the central node in the graph.

While implementing the iteration of neighborhood aggregation, the feature extraction process may be divergent. For instance, when applying the proposed feature construction strategy in the IEEE 57-bus system and visualizing the features at buses 2, 3, and 4, the results are shown in Fig. 3. It can be seen that when the number of iterations is greater than 9, the features obtained by aggregation diverge, resulting in the inapplicability of the features constructed by this method. Hence, it is necessary to normalize the nodal features and avoid diverging.
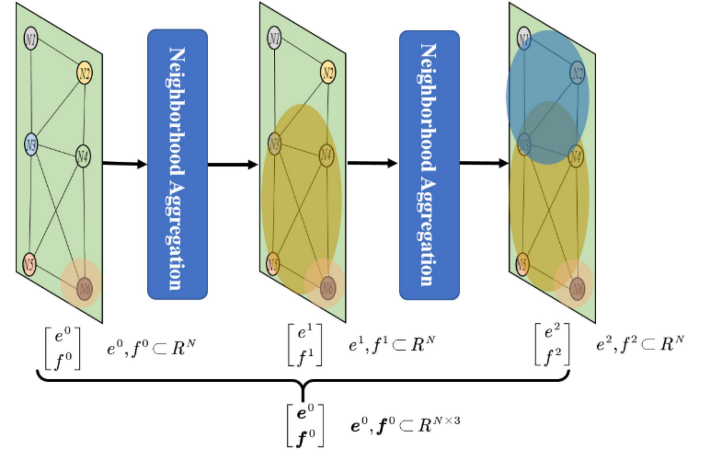

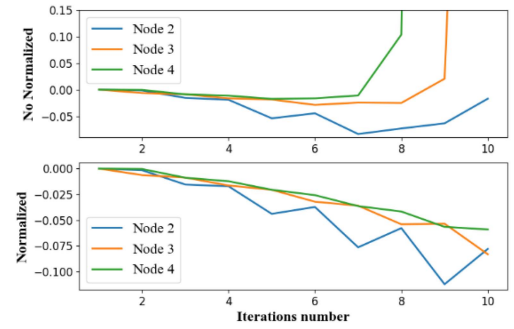Fig. 2.　The iterative feature construction strategy.

Fig. 3.　The comparison of feature construction with and without normalization.

According to the fact that $e^2 + f^2 \approx 1$, the normalization can be formulated as:

$$e = \frac{e}{\sqrt{e^2 + f^2}}, f = \frac{f}{\sqrt{e^2 + f^2}} \tag{25}$$

Eventually, the model-informed feature construction is developed and the implementation is shown in Fig. 4.

D. Physics-guided Graph Convolution Neural Network

Setting the constructed node feature e^0, $f^0 \subset R^{N \times k}$, power demand PD, $QD \subset R^{N \times 1}$, and physic topology information G_{ndiag}, B_{ndiag}, G_{diag}, B_{diag}, the physics-guided GCNN is developed to predict the OPF solution. It consists of two blocks, one is the feature extraction block and the other is the prediction block, as depicted in Fig. 5.

The feature extraction block composed of several graph convolution layers plays an important role in extracting the topological and physical features. Due to the assumption in the derivation of the physics-guided neighborhood aggregation, the constructed feature in Section-C cannot accurately indicate the distribution of the OPF solution. Besides, the OPF model does not include the equality constraints of power flow equations. The inequality constraints of power flow limits in power lines based on the practical and engineering conditions also have a
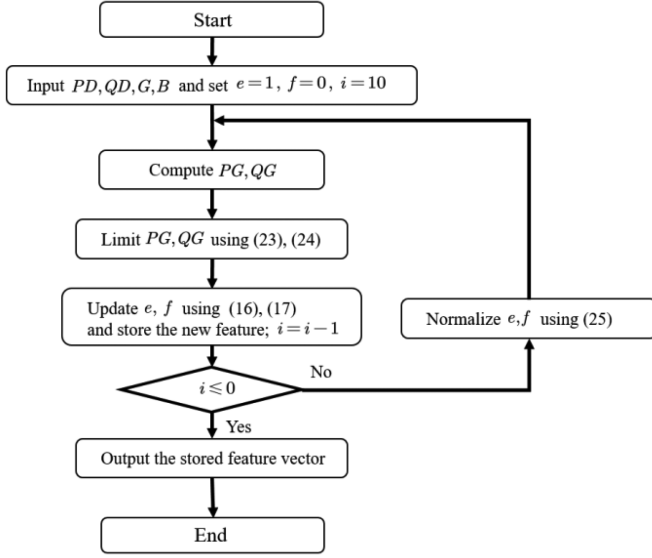
Fig. 4. The flow chart of the feature construction strategy, where i is the iteration number.

big impact. Therefore, the enhanced learning ability with multiple graph convolution layers is necessary to ensure sufficient features extraction. The mathematical formulation of the graph convolution is presented in (18), and the activation function $f(\cdot)$ is the *tanh* so that e_i and f_i are in $[-1, 1]$. Because of the strong dependence on the physical model, the pooling layers that change the physical topology are not added in the feature extraction block.

The prediction block makes up with several fully connected layers. Through the model-informed feature construction and the trainable feature extraction block, the topology feature and physical relationship can be transformed into a general and high dimension space. Then, the final step is to approximate the OPF solution using the extracted features. Specifically, the feature matrix of all nodes is flattened to 1 dimension and there are several fully connected layers to predict the final OPF solution after the flattened layer.

Eventually, combining the two blocks, the complete architecture of physics-guided GCNN for the OPF problem is designed, as shown in Fig. 5. In the architecture, the topology of the IEEE 6-bus system is taken as an example. There are 2 trainable graph convolution layers and 2 fully connected layers.

E. Extension of the Proposed Method

Here, we discuss the potential of embedding the proposed method into other learning techniques, such as the reinforcement learning (RL) method. The RL approach is a straightforward solution for optimal power flow. It explores the action space based on environmental rewards and does not require any labels to supervise the learning of neural networks. However, even though the RL approach is model-free, the feature extraction ability of neural networks that are used as actors and critics has a considerable impact on their performance. Therefore, the proposed GCNN could be exploited as the actor and critic neural

network in the actor-critic RL approaches to strengthen the feature extraction capability. The detailed structure is shown in Fig. 6.

Besides, RL is also an ideal technique to address the multiple time period dispatch problem [25], [26]. Its goal is to maximize the long-term reward and it can consider the coupling relationship among different time steps. In this circumstance, a robust neural network is required to extract the complex features, where the proposed physics-guided GCNN with robust feature extraction capability can be a desired choice.

IV. THE CORRELATIVE LEARNING

Essentially, the learning algorithm for NN is a process to approximate the mapping of the OPF problem. In the NN, many trainable variables $\theta = \{w, b\}$ need to be optimized based on the learning algorithm. Hence, the training of NN is an optimization problem and the common loss function is the mean squared error for the approximating problem, shown as

$$L_{supervised} = E\left[(y_{out} - y_{label})^2\right] \quad (26)$$

where y_{out} is the predicted value of the NN; y_{label} is the corresponding label.

However, (26) ignores the relationship along errors of all outputs. For instance, the predicted voltage and power generation should satisfy the power flow equation. Therefore, the extended terms to describe the physical relationship are added. Based on the OPF model, the extended-terms in the loss function are defined as:

$$L_{\triangle,PG} = E\left[(PG_{out} - f_{PG}(V_{out}))^2\right], \quad (27)$$

$$L_{\triangle,QG} = E\left[(QG_{out} - f_{QG}(V_{out}))^2\right], \quad (28)$$

$$L_{\triangle,PL} = E\left[(PL_{out} - f_{PL}(V_{out}))^2\right], \quad (29)$$

$$L_{\triangle,QL} = E\left[(QL_{out} - f_{QL}(V_{out}))^2\right], \quad (30)$$

where $f_{PG}, f_{QG}, f_{PL}, f_{QL}$ are the variants of power flow equations to calculate PG, QG, PL, QL based on the predicted node voltage magnitude and angle.

The extended terms are to optimize the power difference between the predictions and the one calculated by the predicted voltage via the physical equations. Compared with (26), no labels are needed. It can strengthen the physical relationship learning rather than being limited to minimizing errors. Using $L_{\triangle,PG}$ at the node i and updating the w_{mn} in the hidden layer which indicates the weight from m neuron to n neuron, which leads to $\triangle w_{mn}$:

$$\triangle w_{mn} = -\frac{\partial(PG_{out,i} - f_{PG}(V_{out,i}))^2}{\partial w_{mn}}, \quad (31)$$

where the unbalanced power between the directly predicted $PG_{out,i}$ and $f_{PG}(V_{out,i})$ is used to calculate the gradient and determine $\triangle w_{mn}$. To demonstrate the improvements of this proposed loss function, adding the labels $PG_{label,i}$ in (31) and
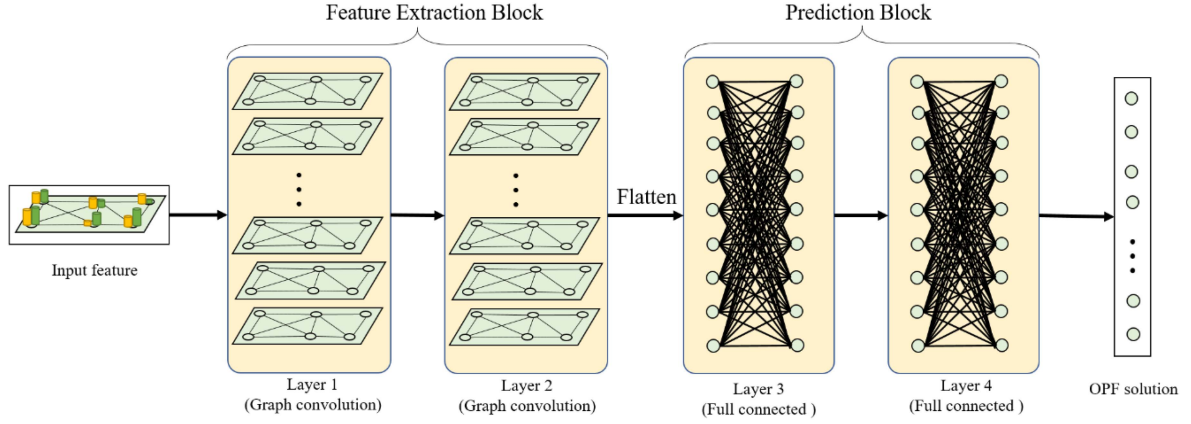
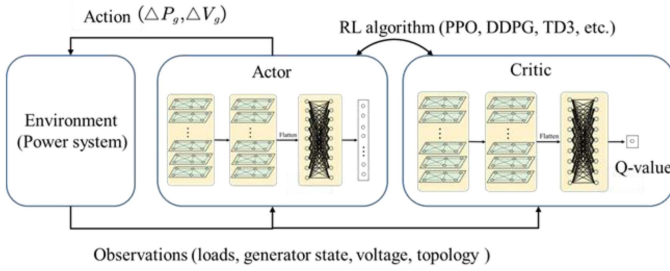Fig. 5.   The architecture of the physics-guided GCNN.

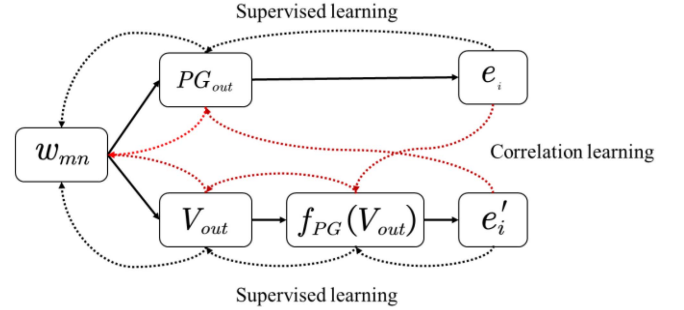Fig. 6.   The application of the proposed GCNN in the actor-critic RL approach.

Fig. 7.   The components of back-propagation in the proposed loss function.

after simplification, we get:

$$\triangle w_{mn} = -2e_i \frac{\partial PG_{out,i}}{\partial w_{mn}} - 2e'_i \frac{\partial f_{PG}(V_{out,i})}{\partial w_{mn}}$$
$$+ 2e' \frac{\partial PG_{out,i}}{\partial w_{mn}} + 2e_i \frac{\partial f_{PG}(V_{out,i})}{\partial w_{mn}} \qquad (32)$$

where:

$$e_i = PG_{out,i} - PG_{label,i} \qquad (33)$$
$$e'_i = f_{PG}(V_{out,i}) - PG_{label,i} \qquad (34)$$

Then, the first and second components in (32) are the parameter change computed by the back-propagation algorithm. The two components guide the NN to learn the error directly related to the outputs. The rest two components update w_{mn} according to the physical correlation, which is shown in Fig. 7. For instance, e'_i calculated by $f_{PG}(V_{out,i})$ can also train w_{mn} through $PG_{out,i}$ in the back-propagation process. Therefore, the proposed method does not only optimize the errors but also mine the physics correlation.

In the power system, the main unbalance term is the nodal active power. Therefore, to improve the training efficiency, the correlative learning loss of PG is added with a weight κ to the final loss function, that is:

$$L = L_{supervised} + \kappa L_{\triangle,PG} \qquad (35)$$

V. CASE STUDY

To demonstrate the effectiveness of the proposed OPF method, various case studies have been carried out on the IEEE 39-bus, 57-bus, 118-bus, and 300-bus systems.

A. Simulation Settings

The settings of different cases are how to simulate the uncertain renewable energy, load fluctuations, and varying topologies. Since it is hard to ensure the full exploration of the possible state, the sample generation is still a bottleneck for the current data-driven methods. In the simulation, certain conditions of the renewable energy, load and topology changes are given for training and testing sample generation. The wind farms and PVs are integrated to simulate the uncertainty of renewable energies. The wind velocity obeys the Weibull distribution with parameters $\lambda = 5.089$ and $k = 2.016$, and the solar irradiance obeys Beta distribution with parameters $\alpha = 2.06$ and $\beta = 2.5$. Randomly connecting some wind farms and PVs to different nodes, the renewable energy penetration rates of each testing system are shown in Table II. The load fluctuation obeys the normal distribution where the mean equals the case's default value and the standard deviation is 0.1. To include various topologies in the training and testing data, the N-1 contingency of some branches is included. Then, the training and testing data are generated using PYPOWER library on a PC with Intel(R) Core

TABLE II
RENEWABLE ENERGY PENETRATION RATE OF DIFFERENT IEEE TESTING SYSTEM

IEEE testing system	Renewable energy penetration rate
IEEE 39-bus	50.7%
IEEE 57-bus	35.77%
IEEE 118-bus	28.25%
IEEE 300-bus	22.19%

(TM) i7-10700K CPU@ 3.80GHz 3.80 GHz, 16 GB RAM, and NIVIDIA GeForce RTX 2080Ti. The comparison methods and corresponding descriptions are shown in Table III. The GCNNs (M1-M5) have three graph convolution layers with 8 channels and three fully connected layers with 1000 neurons each layer. The CNNs (M6-M7) consist of four convolution layers with 3 × 3 kernels and [64, 128, 256, 256] channels and three fully connected layers with 1000 neurons in each layer. The stride of each convolution is (2, 1). The DNNs (M8-M9) have three layers and 1000 neurons per layer. All of the neural networks are implemented in Tensorflow 2.8.0 python platform. The Adam optimizer is utilized to train all the NNs.

To improve the training efficiency and avoid the impact of outliers, the z-score normalization method is applied in all the testing cases. It is formulated as:

$$y_{out} = \begin{cases} (y - y_{mean})/y_{std}, & if \, y_{std} \neq 0 \\ y - y_{mean}, & if \, y_{std} = 0 \end{cases} \quad (36)$$

where y_{mean}, y_{std} are the mean value and standard deviation, respectively.

In the case study, we mainly discuss the prediction of the setpoints of the generators. The performance is measured by the probabilistic accuracy, as shown in (37). It reflects the probability of the prediction error less than a threshold. The thresholds for PG and VG are 1MW and 0.001 p.u., respectively.

$$p = P\left(\left|T - \hat{T}\right| < thr\right) \quad (37)$$

where T and $\hat{T}$ are the predicted value and actual value, respectively, thr is the threshold.

B. Algorithm Validation Under Fixed Topology

In this subsection, the effectiveness of the proposed method under a fixed topology is demonstrated. Five GCNNs constructed with identical structures but different graph convolution kernels are compared. Besides, the prediction block in GCNN is composed of three fully connected layers. They are tested in the IEEE 57-bus system under the default topology. After generating 10000 samples for training and 2000 samples for testing, we train the four neural networks with 2000 epochs. The learning rate is 0.001 in the first 1000 epochs and changes to 0.0001 in the remaining epochs. The loss curve is shown in Fig. 8 and the probability accuracies are displayed in Table IV. As for M5, it utilizes the trained model of M4 and fine-tunes the model with $\kappa = 1$ in 200 epochs. Besides, the mean absolute errors of generator voltage are plotted in Fig. 9.
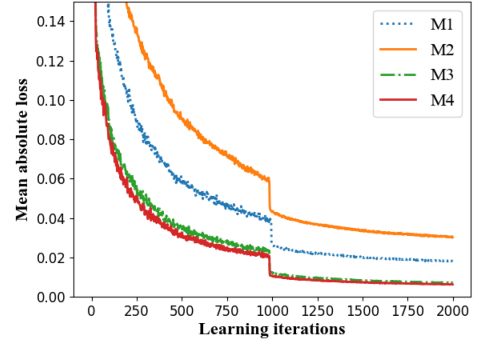

Fig. 8. The loss curve of the four neural networks. The initial learning rate is 0.001 and it becomes 0.0001 after 1000 epochs.
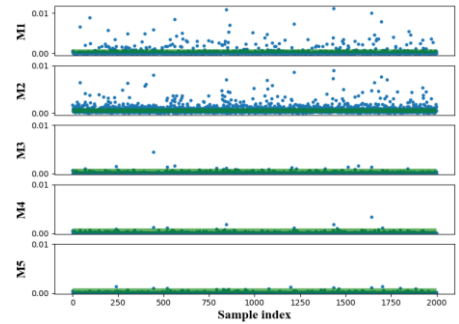

Fig. 9. The mean absolute error of generator voltage.

Compared with M1, M3 can converge faster and has smaller errors at the end of training. It indicates that the physics-embedded graph convolution has better convergence. Besides, the probabilistic accuracy in the testing samples increases by 3.37% for the prediction of generator voltage. Therefore, the physics-embedded graph convolution can improve the convergence as well as the prediction accuracy as compared to M1. By contrast, utilizing the impedance matrix to construct the graph convolution kernel in M2 decreases the convergence and the prediction accuracy as compared to M3. Hence, by simply embedding topology with the impedance matrix, the physical relationship cannot be effectively extracted. The reason is that the physical variable, such as the impedance matrix, reveals a coupled physical relationship while ignoring the self-relationship that misleads the NN's feature extraction. As for M1, although the graph convolution kernel of M1 is a single matrix derived from the adjacent matrix, it is the theoretical kernel for a general graph based on the spectral graph theory.

The comparison between M3 and M4 indicates that the model-informed feature construction provides a better input feature for the physics-guided GCNN and promotes the convergence of training. In Fig. 9, it can be observed that M5 has the least outliers in the testing samples even though the prediction accuracies are very close to M4.

According to the test results, our proposed method can achieve desired performance in fixed topology scenarios. Moreover, the correlative learning loss function to fine-tune the physics-guided

TABLE III
THE DESCRIPTION OF DIFFERENT METHODS

Method	Neural network	Description		
M0		Traditional Interior-Point algorithm (benchmark of data-driven methods)		
		Physics-embedded graph convolution kernel	Model-guided feature construction	Correlative learning loss function
M1	Original GCNN [24]	×	×	×
M2	GCNN [14]	×	×	×
M3		○	×	×
M4	Physics-guided GCNN	○	○	×
M5		○	○	○
M6	CNN [20]	The diagonal elements of admittance matric is regarded as topology input feature.		
M7	CNN [13]	The $|Y_{\text{diag}}|$ and $\angle Y_{\text{diag}}$ is regarded as topology input feature.		
M8	DNN [18]	The voltage difference of each bus are used as the topology feature.		
M9	DNN	The diagonal elements of B are used as the topology feature.		

[1] $|Y_{\text{diag}}|$ and $\angle Y_{\text{diag}}$ are the magnitude and angle of the diagonal elements in admittance matrix.

TABLE IV
THE PROBABILISTIC ACCURACY OF THE FOUR NEURAL NETWORKS UNDER
THE DEFAULT TOPOLOGY

Methods	Testing time(s)	P_{VG}	P_{PG}
M1	0.5974	93.71%	99.56%
M2	0.5395	77.23%	95.29%
M3	0.5495	97.08%	99.44%
M4	0.4936	99.34%	99.85%
M5	0.4617	99.41%	99.63%

TABLE V
THE PERFORMANCE OF THE FOUR NEURAL NETWORKS UNDER VARYING
TOPOLOGY

System		M5	M6	M7	M8	M9
39-bus	P_{VG}	**97.76%**	86.06%	86.15%	93.09%	88.28%
	P_{PG}	**98.05%**	84.39%	86.31%	92.72%	89.82%
57-bus	P_{VG}	**99.99%**	84.43%	71.02%	82.24%	77.57%
	P_{PG}	**99.76%**	93.36%	82.51%	96.37%	94.35%
118-bus	P_{VG}	**93.49%**	83.29%	70.05%	82.47%	83.65%
	P_{PG}	**98.47%**	92.89%	85.45%	93.18%	93.05%
300-bus	P_{VG}	**92.72%**	69.83%	77.46%	60.09%	71.72%
	P_{PG}	**94.05%**	70.85%	78.26%	70.26%	72.63%

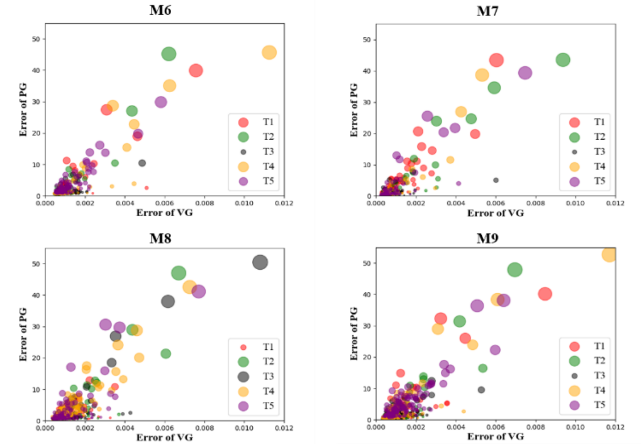GCNN is helpful to avoid outliers by strengthening the physical correlation.

Fig. 10.   The error distribution of different topology. The bubbles in the figure above represent samples and their size indicate the magnitude of the error.

C. Validation Under Varying Topologies

In this subsection, the effectiveness of the proposed method under varying topologies is demonstrated. There are two kinds of NNs considering the varying topologies in the literature, i.e., DNN and CNN. So the methods M5-M9 are compared in IEEE 39-bus, 57-bus, 118-bus, and 300-bus systems with multiple topologies. Note that M5 is the proposed method. In those cases, the statuses of branches are changed to generate varying topologies. 10000 and 2000 samples are respectively generated for training and testing under five types of line contingencies. The prediction accuracies of generator real power and voltage are shown in Table V.

Compared to other methods, M5 has the best prediction accuracy for all systems. In the IEEE 57-bus system, the prediction accuracy of M5 is up to 99% while the other methods are almost lower than 90%. By contrast, in the IEEE 118-bus and 300-bus systems, the prediction accuracies of generator voltage are only 93.49% and 92.72%, but still significantly outperform other methods. In general, the average improvement is 13.30% and the maximum improvement is up to 32.63%. In the meanwhile, it can be observed that the predication accuracies decrease as the size of system increases. The 10000 training samples cannot ensure accurately prediction even though the physics are embedded into the GCNN.

To analyze the prediction errors in each topology, the testing results of IEEE 300-bus systems are displayed in Fig. 10. Samples are plotted as bubbles in the figure, where the x-axis

represents the mean absolute error of generator voltage and the y-axis is the mean absolute error of active generation. Their sizes also indicate the magnitudes of errors. It can be observed that the bubbles with a larger size in M5 are almost red, which indicates that the corresponding samples have identical topology. The main errors in M5 come from T1, indicating the trained physics-guided GCNN can accurately predict the set-points of generators under T2-T5 but not T1. However, in other methods, the big-sized bubbles cover almost the five topologies. Those methods are capable of predicting one topology accurately but have weak adaptability to other topologies.

Therefore, the proposed OPF method can predict the solution more accurately and has better topology adaptability as compared to the state-of-the-arts methods. It could be a promising tool to implement probabilistic analysis in data-ahead operation or real-time risk monitoring, while considering the multiple topologies.

D. Demonstration for the Extension of the Proposed Method

To demonstrate the effectiveness of the proposed GCNN with RL, it is tested in the Proximal Policy Optimization (PPO) RL method to deal with the optimal power flow problem. The environment is built based on the IEEE 30-bus system and the load profile is sampled from a normal distribution whose average is the default value and the standard derivation is 0.5. Besides, there is a branch randomly chosen to be out of service. The observations include active and reactive loads, current generation and voltage magnitude of power plants, the conductance matrix and susceptance matrix. The reward function is shown as (38) which is referring to [13].

$$r = \begin{cases} -5000, & \text{PF solver is diverged.} \\ R_{pg,v} + R_{v,v} + R_{br,v}, & \text{constraints violations exist.} \\ -0.01 C_{\text{generators}}, & \text{system is normal.} \end{cases} \tag{38}$$

where $R_{pg,v}, R_{v,v}, R_{br,v}$ are the total amount of violations as negative rewards; $C_{generators}$ is the total cost of all generators. To verify the feature extraction ability of the proposed GCNN with the PPO training algorithm, we train the GCNN, CNN, and DNN with a 0.0001 learning rate. DNN consists of two fully connected layers with 256 neurons. There are two convolution layers with a 4×4 kernel, a flatten layer, and two fully connected layers with 256 neurons in CNN. The proposed GCNN has two graph convolution layers, a flatten layer, and two fully connected layers with 256 neurons. After 180000 steps of training, the reward curves of different neural networks are shown in Fig. 11. What's more, to verify effectiveness of the proposed GCNN in OPF problem considering the time-dependent constraints, the ramping rate of generators is added to the simulated environment. We set the ramping rate to be 10% of the maximum active generation and retrain the three kinds of neural networks. Results are shown in Fig. 12.

According to the results, it can be observed that when the PPO RL is combined with the proposed GCNN, it can converge faster than those with the CNN and DNN using the identical training parameters. The final reward of PPO RL with the proposed GCNN is the largest one, which demonstrate that the proposed
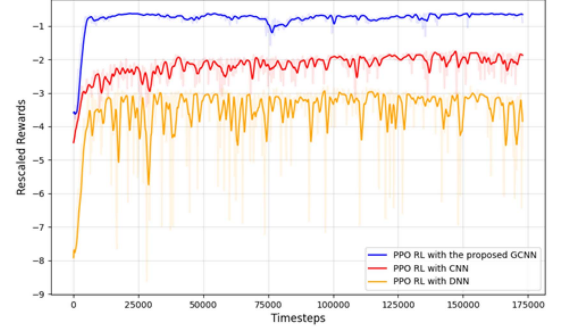

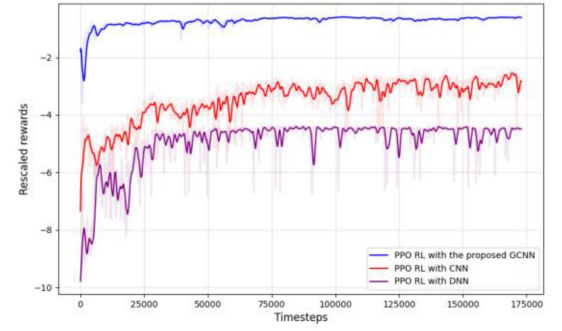Fig. 11. The reward curves of different neural networks when training with the PPO RL algorithm.

Fig. 12. The reward curves of different neural networks when considering the ramp rate.

GCNN still has the robust feature extraction capability when combining with the PPO RL method. When considering ramping rate of generators, the performance of the proposed GCNN is still the best one. To conclude, the proposed GCNN can be applied in the RL approaches and it is a helpful tool to address complicated OPF problems, such as the multiple-period co-optimization.

VI. CONCLUSION

The paper proposes a data-driven OPF approach considering topology feature learning. Compared to other data-driven OPF methods, taking advantage of the model-informed feature construction and training the neural network with an extended-term loss function to learn the physical correlation, it improves the feature extraction ability by the physics-embedded graph convolution kernel. Meanwhile, because the complex physical relationships like the coupling and physical topology are directly embedded to the graph convolution network, the learning difficulty of the OPF mapping from loads to generation output is reduced with the help of model-driven GCNN architecture. An evidence is that the prediction performance of the proposed model-guided GCNN is improved by an average of 13.30% and up to 32.63% compared with state-of-the-art methods.

REFERENCES

[1] K. Lehmann, A. Grastien, and P. Van Hentenryck, "AC-feasibility on tree networks is NP-hard," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 798–801, Jan. 2016.

[2] J. Lavaei and S. H. Low, "Zero duality gap in optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 92–107, Feb. 2012.

[3] Z. Yang, H. Zhong, A. Bose, T. Zheng, Q. Xia, and C. Kang, "A linearized OPF model with reactive power and voltage magnitude: A pathway to improve the MW-only DC OPF," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1734–1745, Mar. 2018.

[4] D. Deka and S. Misra, "Learning for DC-OPF: Classifying active sets using neural nets," in *Proc. IEEE Milan PowerTech*, 2019, pp. 1–6.

[5] V. J. Gutierrez-Martinez, C. A. Cañizares, C. R. Fuerte-Esquivel, A. Pizano-Martinez, and X. Gu, "Neural-network security-boundary constrained optimal power flow," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 63–72, Feb. 2011.

[6] M. Niu, C. Wan, and Z. Xu, "A review on applications of heuristic optimization algorithms for optimal power flow in modern power systems," *J. Modern Power Syst. Clean Energy*, vol. 2, no. 4, pp. 289–297, Dec. 2014.

[7] M. Yan, M. Shahidehpour, A. Paaso, L. Zhang, A. Alabdulwahab, and A. Abusorrah, "A convex three-stage SCOPF approach to power system flexibility with unified power flow controllers," *IEEE Trans. Power Syst.*, vol. 36, no. 3, pp. 1947–1960, May 2021.

[8] S. Mhanna and P. Mancarella, "An exact sequential linear programming algorithm for the optimal power flow problem," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 666–679, Jan. 2022.

[9] F. Hasan, A. Kargarian, and A. Mohammadi, "A survey on applications of machine learning for optimal power flow," in *Proc. IEEE Texas Power Energy Conf.*, 2020, pp. 1–6.

[10] X. Lei, Z. Yang, J. Yu, J. Zhao, Q. Gao, and H. Yu, "Data-driven optimal power flow: A physics-informed machine learning approach," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 346–354, Jan. 2021.

[11] 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.

[12] A. S. Zamzam and K. Baker, "Learning optimal solutions for extremely fast AC optimal power flow," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. for Smart Grids (SmartGridComm)*, 2020, pp. 1–6.

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

[14] D. Owerko, F. Gama, and A. Ribeiro, "Optimal power flow using graph neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2020, pp. 5930–5934.

[15] A. Velloso and P. Van Hentenryck, "Combining deep learning and optimization for preventive security-constrained dc optimal power flow," *IEEE Trans. Power Syst.*, vol. 36, no. 4, pp. 3618–3628, Jul. 2021.

[16] P. Pareek and H. D. Nguyen, "Gaussian process learning-based probabilistic optimal power flow," *IEEE Trans. Power Syst.*, vol. 36, no. 1, pp. 541–544, Jan. 2021.

[17] Z. Yan and Y. Xu, "Real-time optimal power flow: A Lagrangian based deep reinforcement learning approach," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 3270–3273, Jul. 2020.

[18] F. Hasan, A. Kargarian, and J. Mohammadi, "Hybrid learning aided inactive constraints filtering algorithm to enhance ac OPF solution time," *IEEE Trans. Ind. Appl.*, vol. 57, no. 2, pp. 1325–1334, Mar./Apr. 2021.

[19] M. Xiang, J. Yu, Z. Yang, Y. Yang, H. Yu, and H. He, "Probabilistic power flow with topology changes based on deep neural network," *Int. J. Elect. Power Energy Syst.*, vol. 117, 2020, Art. no. 105650.

[20] 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.

[21] Y. Chen, S. Lakshminarayana, C. Maple, and H. V. Poor, "A meta-learning approach to the optimal power flow problem under topology reconfigurations," *IEEE Open Access J. Power Energy*, vol. 9, pp. 109–120, 2022, doi: 10.1109/OAJPE.2022.3140314.

[22] C. Kim, K. Kim, P. Balaprakash, and M. Anitescu, "Graph convolutional neural networks for optimal load shedding under line contingency," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, 2019, pp. 1–5.

[23] D. Wang, K. Zheng, Q. Chen, G. Luo, and X. Zhang, "Probabilistic power flow solution with graph convolutional network," in *Proc. IEEE PES Innov. Smart Grid Technol. Europe*, 2020, pp. 650–654.

[24] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Representations*, 2017, pp. 1–14.

[25] C. Hu, Z. Cai, Y. Zhang, R. Yan, Y. Cai, and B. Cen, "A soft actor-critic deep reinforcement learning method for multi-timescale coordinated operation of microgrids," *Protection Control Modern Power Syst.*, vol. 7, no. 3, pp. 423–432, 2022.

[26] H. Jin, Y. Teng, T. Zhang, Z. Wang, and Z. Chen, "A deep neural network coordination model for electric heating and cooling loads based on IoT data," *CSEE J. Power Energy Syst.*, vol. 6, no. 1, pp. 22–30, Mar. 2020.

Maosheng Gao received the B.S. degree in 2020 from the School of Electrical Engineering, Chongqing University, Chongqing, China, where he is currently working toward the Ph.D. degree in electrical engineering. His research interests include deep learning and its applications in power and energy systems.
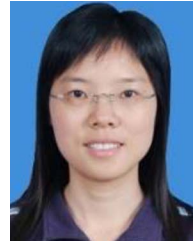
Juan Yu (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Chongqing University, Chongqing, China, in 2007. She is currently a Full Professor with Chongqing University. Her research interests include Big Data analytics and power system analysis.

Zhifang Yang (Member, IEEE) received the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2018. He is currently a Full Professor with Chongqing University, Chongqing, China. His research interests include power system analysis and electricity market.

Junbo Zhao (Senior Member, IEEE) received the Ph.D. degree from the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA, in 2018. He is currently an Assistant Professor with the University of Connecticut, Storrs, CT, USA. He was an Assistant Professor and Research Assistant Professor with Mississippi State University, Starkville, MS, USA, and Virginia Tech, Blacksburg, VA, USA, from 2019 to 2021 and 2018 to 2019, respectively. From May to August 2017, he did the summer internship with Pacific Northwest National Laboratory. He is currently the Chair of IEEE Task Force on Power System Dynamic State and Parameter Estimation, the Chair of Cyber-Physical Interdependency for Power System Operation and Control working group, and the Secretary of the IEEE Task Force on Synchrophasor Applications in Power System Operation and Control.