AC Optimal Power Flows: Combining Lagrangian Dual and physics-Guided Graph Neutral Networks

Mei Yang ¹, Junyong Liu ², Gao Qiu ², Tingjian Liu ², and Zhiyuan Tang ²

¹sichuan university
²Affiliation not available

October 30, 2023

Abstract

Under high-dimensional and nonlinear stochastic power system environment, artificial intelligence (AI) is becoming a promising alternative towards the urgent demand of real-time AC optimal power flow (OPF). However, traditional AI only fit the settlement results of other method, and is unable to find in-depth law of power flow. This may result in an undesired poor generalizability. To address this issue, a unsupervised OPF model combining physics- guided graph neutral networks and Lagrangian dual is proposed. The model incorporates the physical constraints into training tensor graph, as well as a dual Lagrangian method to satisfy the physical and engineering constraints present in the OPF. At the same time, since topology of power system may vary, the branch characteristics are embeded into node graph to adapt to grid topology contingency. Numerical tests on benchmarks demonstrate that, in unseen operating condition, the proposed method enables a near or even better solution than conventional optimization algorithm, but consumes much more than 100 times calculation efficiency.
AC Optimal Power Flows: Combining Lagrangian Dual and physics-Guided Graph Neutral Networks

Mei Yang, Junyong Liu, Member, IEEE, Gao Qiu, Youbo Liu, Member, IEEE, Tingjian Liu, Zhiyuan Tang, Kai Liu, Senior Member, IEEE

Abstract—Under high-dimensional and nonlinear stochastic power system environment, artificial intelligence (AI) is becoming a promising alternative towards the urgent demand of real-time AC optimal power flow (OPF). However, traditional AI only fit the settlement results of other method, and is unable to find in-depth law of power flow. This may result in an undesired poor generalizability. To address this issue, a unsupervised OPF model combining physics-guided graph neutral networks and Lagrangian dual is proposed. The model incorporates the physical constraints into training tensor graph, as well as a dual Lagrangian method to satisfy the physical and engineering constraints present in the OPF. At the same time, since topology of power system may vary, the branch characteristics are embeded into node graph to adapt to grid topology contingency. Numerical tests on benchmarks demonstrate that, in unseen operating condition, the proposed method enables a near or even better solution than conventional optimization algorithm, but consumes much more than 100 times calculation efficiency.

Index Terms—Graph neural networks, Physical mechanism, Real-time optimal power flow, Unsupervised learning

I. INTRODUCTION

Due to manifold emerging events (e.g., highly penetrated renewable, varying topology, deregulated power market, etc.), power system is operating in a more volatile and stochastic manner, such that large-scale failures, even blackouts, are easier to occur than ever. This forces system operators to adjust the generators set points with increasing frequency in order to serve the power demand while ensuring stable network operations [1]. However, AC optimal power flow (AC-OPF) is usually a high-dimensional, nonlinear, and non-convex problem, due to sinusoidal nature of electrical characteristics. Thus it is challenging and time-consuming to solve AC-OPF [2], [3]. Usually, towards a fast and reliable dispatch, non-convexity is relaxed as convex models by using DC-OPF with small angle approximation or second-order cone trick. Unfortunately, DC-OPF approximation is no longer valid when power grids are heavily loaded [4], [5], and second-order cone may result in untighten decision feasible region and inaccurate solution, which may provide insecure dispatch [6]. In a word, traditional physic-based methods are in trouble with compromising decision precision and efficiency, as tedious iterations and serious non-convergent issues exist [7], [8], [9].

To address such problems, recent works entertain the idea of using a deep neural network (DNN) to predict the solutions of the AC-OPF. A common way is to conduct supervised learning (SL) or reinforcement learning (RL) to directly learn mapping between operations and AC-OPF decisions [10], [11], [12]. Once DNN model is well trained offline, OPF can be simply provided via feedforward parallel computing, computational burden is thus greatly reduced. In [13], the feasibility of SL-based AC-OPF solving scheme is proved. [12] utilized fully connected neural network to model AC-OPF decisions on extremely fast timescales, where inputs and outputs are load and AC-OPF solutions, in that order. However, SL families only enable capture of experienced AC-OPF patterns implicated in data, which is also called imitation learning (IL). Their decision quality depends on data distribution and SL models themselves. True physical patterns cannot be found by SL/IL such that robustness and generalizability can be limited. Exactly as stated in [14], pure SL/IL is prone to violate constraints. Therewith, two alternatives for this issue are came up with. One is still part of SL/IL, but several mandatory constraints are involved. For instance, [14] entailed OPF sensitivities into DNN training loss, which results in a sample-efficient learning process. [15] constrained load-dispatch mapping learning process via introducing constraint violation penalties into DNN loss, then Lagrangian duality was used to tune penalty weights during training. The numerical study showed that, ACOPF solving DNN becomes more precise with learning region enforced. Similar idea can be found in [16], where the authors used DNN to solve DC-OPF and a post projection process is set for online decision reliability. The other one is RL, which attempts to fit AC-OPF pattern via reward mechanism and interacting physical environment. [10] introduced RL to learn reward incentive that tends to solve AC-OPF. The trained agent enables real-time voltage control. Similarly, [11] successfully verified efficiency and effectiveness of RL in automatic power generation control. However, reward function must be carefully designed for a high performance RL. Even so, RL is still readily to fall in local optimum. Lagrangian method is then introduced to try to fix this issue. [17] incorporated Lagrangian into RL training, such that RL enables search OPF solutions following optimization path. The testing results show that, involving OPF physics indeed improves RL performance in dispatch task.
Overall, existing studies provide massive solution for ACOPF, but another major problem of generalizability to topology varying is under consideration and remains unresolved.

To further fix topology-varying challenge, spatial correlation inside data must be taken into account. Recent graph neural network (GNN) is exactly for this end. GNN can fully consider spatial features and is widely used in power system topology-related analysis, such as fault identification [18], [19], cascading failure search [20], etc. As salient performance of GNN is verified, GNN is gradually involved in power flow computation, SCED or ACOPF problem [21], [22], [23]. Specifically, supervised learning implementations are deployed in [21] and [22]. They simply imitate OPF solver response to decide unseen dispatch, thus may be poor in robustness and scalability. Few researchers therewith realize this, pure GNN imitation is still not enough to aware true operational traits of power grids. [23] proposed a GNN-based power flow analysis method, where GNN provides voltage features and takes power unbalance quantities directly as learning loss. But the method misses OPF modelling, and no topological adaptability is discussed.

In the context, this paper proposes a real-time ACOPF solving method based on Lagrangian duality and physics-guided GNN (L-P-GNN) and Lagrangian Dual Methods. Its main contributions include:

1. For the first time, a P-L-GNN model (ChebNet-4) is proposed towards a real-time AC-OPF solution. In this algorithm, power electrical physics are modelled and seamlessly incorporated with GNN. These differentiable physical equations are derived and forwarded to support training gradient backwards, such that GNN learning is dominantly guided by true power system operational law. Note that, the proposed GNN is in an unsupervised learning, so that computationally intensive SCED labeled dataset is circumvented.

2. The power system topology features are fully utilized in a physics-guided OPF learning event. This is benefited from the nature of GNN, which allows us to represent power system topology with a graph. Reconciled with physics, GNN generally emulates Kirchhoff’s law via neighborhood information aggregation, and learns operational patterns via message passing. The proposed GNN provides a mimic for power system energy flow, it thus well generalized to topology varying.

3. In pursuit of globally optimal OPF solutions for training, we recasts the OPF solution as the Lagrangian dual of the empirical risk minimization under constraints.

4. The test outcomes of the trained GNN provides that our proposed method holds salient generalizability and remarkable decision efficiency. Specifically, decision loss on dataset of unseen operations hits a well level of 4e-5, which is significantly better than traditional supervised learning family. And compared to mature MATPOWER toolbox, decision speed via our method is much faster than 200 times, and operational cost is almost the same. Consequently, among existing physic and data-driven methods, the proposed method is the best alternative that compromises dispatch efficiency and economy.

II. RESEARCH METHODOLOGY

A. ACOPF Model

ACOPF can be compactly formulated as follows

$$\min_u f(x, u) \quad (1a)$$

s.t. $$g(x, u) = 0 \quad (1b)$$

$$h(x, u) \leq 0 \quad (1c)$$

where \( g \) and \( h \) represent the set of equality and inequality constraints, respectively. \( x \) is the vector of state variables, including bus voltage magnitudes, bus voltage angle. \( u \) is the vector of control variables. \( f \) denotes objective function, which is usually total generation cost and renewable curtailment cost, i.e.,

$$f = \sum_{i \in S_G} f_i^C + \sum_{i \in S_R} f_i^R \quad (2a)$$

$$f_i^C = a_i (P_i^g)^2 + b_i P_i^g + c_i, \quad \forall i \in S_G \quad (2b)$$

$$f_i^R = P_i^{g,F} - P_i^g, \quad \forall i \in S_R \quad (2c)$$

where \( S_G \) and \( S_R \) respectively denote the set of traditional generators and renewable generators. \( P_i^F \) is practical generation, \( P_i^{g,F} \) is forecast of renewable generation.

\( x \) and \( u \) are constrained by several physics and security demands, e.g., Kirchhoff’s law, generation limits, voltage constraints and line flow constraints (\( \forall i \in S_G \cup S_R, \forall i \in S_B \)):

$$P_i = P_i^d - P_i^l = V_i \sum_{j \in i} V_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) \quad (3a)$$

$$Q_i = Q_i^d - Q_i^l = V_i \sum_{j \in i} V_j (G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}) \quad (3b)$$

$$P_i^{g,min} \leq P_i^g \leq P_i^{g,max} \quad (4a)$$

$$Q_i^{g,min} \leq Q_i^g \leq Q_i^{g,max} \quad (4b)$$

$$V_i^{min} \leq V_i \leq V_i^{max} \quad (5)$$

$$|S_{ij}| \leq s_{ij}^{max} \quad (6)$$

where \( S_B \) is the set of buses. (3) corresponds to equality \( g \). (4)-(6) are the inequalities \( h \). \( V_i \) represents voltage amplitude of ith bus, \( j \in i \) indicates that bus \( j \) is connected to bus \( i \) through branch \( ij \). \( G_{ij} \) and \( B_{ij} \) represent the conductance and the susceptance of branch \( ij \), respectively. \( \delta_{ij} \) is voltage phase angle difference of branch \( ij \). \( P_i \) and \( Q_i \) are active power and reactive power injection of ith bus respectively. \( V_i^{min} \) and \( V_i^{max} \) are the lower limit and upper limit of voltage amplitude, respectively. \( P_i^{g,max} \) and \( P_i^{min} \) are the maximum and minimum active generation of the ith generator, respectively. \( Q_i^{g,max} \) and \( Q_i^{min} \) are the maximum and minimum active generation of the ith generator, respectively.
B. Principle of Graph Neural Networks

A graph is the set of nodes and edges, denoted by $G(\mathcal{V}, \mathcal{E}, \mathcal{A})$, where $\mathcal{V}$ is the set of nodes, $\mathcal{E}$ is the set of edges, and $\mathcal{A}$ is the adjacency matrix representing topology. Laplace matrix of a graph is defined as $L = D - A$, where $D$ is degree matrix, which is a diagonal matrix and represents the number of neighbors node.

GNN is a neural network built upon graph theory. It was first proposed by Scarselli et al. [24], and developed a series of models, such as ChebNet [25], GCN [26], and GAT [27], etc. They have been demonstrated to be high-powered in topology-related field, e.g., graph classification [28], vertex classification [29] and relational reasoning [30], etc. Typical framework of GNN is shown in Fig.1, where $\mathbf{x}_i \in \mathbb{R}^{F_i}$ is the input feature of node $i$, and $F_{in}$ represents dimension of inputs. $h_l^i \in \mathbb{R}^{F_l}$ is obtained by the fully connected transformation of $h_{l-1}^i$, and $F_l$ represents the number of neurons of the $l$th layer. $h_{l-1}^i \rightarrow h_l^i$ implements node information aggregation. In GNN, aggregating and aggregation are the cores to feedforward and transform features. Specifically, messaging completes longitudinal transmission. After that, hidden features $h_l^i$ are obtained through transverse aggregation of neighbors $h_{l-1}^j$. Such mechanism allows GNN to extract and represent in-depth spatial features of input topologies. This indeed matches the characteristics of power systems.

C. ChebNet

GNN can be divided into two categories, i.e., vertex domain and spectral domain. The main idea of spectral domain method is to study properties of graph by using eigenvalues and eigenvectors of Laplacian matrix. Notably, Laplacian matrix enables reflection of energy flows [31]. The gain of node $i$ caused by the disturbance of node $j$ is shown in Fig.2, where $z$ stands for branch impedance. It can be found that the energy flows of Laplacian matrix satisfies all operational constraints (the inequality constraint is transformed into equality constraint) with low cost and renewables curtailment, so the optimization problem above is equivalent

\[
\min_{\theta} \quad f^C(\mathbf{x}, \phi^\theta(\mathbf{x})) \tag{8a}
\]

s.t.

\[
f^R(\mathbf{x}, \phi^\theta(\mathbf{x})) = 0 \tag{8b}
\]
\[
f^Q(\mathbf{x}, \phi^\theta(\mathbf{x})) = 0 \tag{8d}
\]
\[
f^V(\mathbf{x}, \phi^\theta(\mathbf{x})) = 0 \tag{8e}
\]
\[
f^Q(\mathbf{x}, \phi^\theta(\mathbf{x})) = 0 \tag{8f}
\]

where (8a) correspond to cost constraints (2b). (8b) correspond to renewables curtailment constraints (2c). (8c) and (8d)
Karush-Kuhn-Tucker conditions

\[\phi\]

where \(N\) is the number of buses; and \(P\) variable rally, the output of the model increases the Lagrangian dual described above can be solved by Lagrangian duality. Naturally, demands gradients of all physical and parametrized models.

To instantiate the structure above, we give the generation process of active power balance and generation constraints \(f^P(x, \phi^θ(x))\). According to (3a) and (4a)

\[f^P(x, \phi^θ(x)) = \begin{cases} \sum_{i=1}^{N}(P_i + P_i^d - P_i^{g_{\text{max}}}), & \text{if } P_i + P_i^d > P_i^{g_{\text{max}}} \\ \sum_{i=1}^{N}(P_i^{g_{\text{min}}} - P_i - P_i^d), & \text{if } P_i + P_i^d < P_i^{g_{\text{min}}} \\ 0, & \text{else} \end{cases}\]

where \(N\) is the number of buses; and \(P_i\) is calculated from \(\phi^θ(x)\).

In order to train the model effectively, AC-OPF solution described above can be solved by Lagrangian duality. Naturally, the output of the model increases the Lagrangian dual variable \(\lambda = [\lambda^R, \lambda^P, \lambda^Q, \lambda^V, \lambda^S]\). Construct the Lagrangian function according to (8)

\[
\mathcal{L}^θ(x) = f^C(x, \phi^θ(x)) + \lambda^R f^R(x, \phi^θ(x)) + \lambda^P f^P(x, \phi^θ(x)) + \lambda^Q f^Q(x, \phi^θ(x)) + \lambda^V f^V(x, \phi^θ(x)) + \lambda^S f^S(x, \phi^θ(x))
\]

Furthermore, the loss function is obtained according to the Karush-Kuhn-Tucker conditions

\[
\mathcal{J}^θ(x) = \left(\frac{\partial \mathcal{L}^θ(x)}{\partial x}\right)^2 + \left(\frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^R}\right)^2 + \left(\frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^P}\right)^2 + \left(\frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^Q}\right)^2 + \left(\frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^S}\right)^2
\]

\[
+ \left(\frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^V}\right)^2 + \left(\frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^S}\right)^2 + \left(\frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^S}\right)^2
\]

B. Algorithm Implementation Process and Model Training

The implementation steps of the proposed framework are shown in Fig. 4, which includes offline optimization and online application for SCOPF solutions. Learning such a PGGNN demands gradients of all physical and parametrized models.

This can be done by analyzing (10) and (11), of which the process can be generally derived as

\[
\theta \leftarrow \theta + \frac{\partial J^θ(x)}{\partial \theta} = \theta + \frac{\partial J^θ(x)}{\partial \phi^θ(x)} \frac{\partial \phi^θ(x)}{\partial \theta}
\]

(12a)

\[
\frac{\partial J^θ(x)}{\partial \phi^θ(x)} = 2\frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^R} \frac{\partial \lambda^R}{\partial \phi^θ(x)} + 2\frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^P} \frac{\partial \lambda^P}{\partial \phi^θ(x)} + \frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^Q} \frac{\partial \lambda^Q}{\partial \phi^θ(x)} + \frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^V} \frac{\partial \lambda^V}{\partial \phi^θ(x)} + \frac{\partial \mathcal{L}^θ(x)}{\partial \lambda^S} \frac{\partial \lambda^S}{\partial \phi^θ(x)}
\]

(12b)
Algorithm 1: AC-OPF based on L-P-GNN

Input: \( x = [x_1, x_2, \ldots, x_i, \ldots, x_N] \), A, Iterations : \( t \)
Output: \( [v_i, \delta_i, \lambda_{R_i}, \lambda_{P_i}, \lambda_{Q_i}, \lambda_{V_i}, \lambda_{S_i}] \)

1. Loss \( \leftarrow J^\theta(x) \)
2. for \( k < t \) do
   3. \( J^\theta(k)(x) \leftarrow [v(k), \delta(k), \lambda_{R(k)}, \lambda_{P(k)}, \lambda_{Q(k)}, \lambda_{V(k)}, \lambda_{S(k)}]; \)
   4. \( \theta(k+1) \leftarrow \theta(k) + \frac{\partial J^\theta(k)}{\partial \theta}; \)
   5. if \( J^\theta(k+1) - J^\theta(k) \leq \epsilon \) then
      6. save model;
   end
3. end

the training efficiency. If topology scenarios with minimal probability occur in practice, we can directly adopt L-P-GNN’s training architecture for transfer learning. Noted that there is no need for unseen label preparation, which significantly reduce transfer learning computational-burden.

In order to enhance the learning efficiency of new sample migration, we first adopted clustering to obtain \( k \) typical scenes of historical running data (the value of \( k \) was determined by clustering indicators sum of squared err (SSE), Calinski-Harabasz index (CHI) and Davies-Bouldin index (DBI)), and then packed the new topology and these groups of typical scene data to PGNN for training. By numerical experiments, we proved numerically that the method could effectively adapt the Unseen topology, and the transfer learning efficiency is high.

IV. CASE STUDY

A. Test System and Parameter Settings

To verify the effectiveness of the proposed method, the simulation test is carried out on 2 modified IEEE systems. Generator (the range is 0 to 100% of the rating) and load (the range is 60% to 100% of the rating) data are collected historically, and with period of 5-minute timestamp over a year. 100,000 groups were randomly selected as training data and the remaining is as test set.

Case1 : Modified IEEE 30-bus system as shown in Fig.5. We added renewables on bus 6, 12, 10, 15 and 27; The maximum output is set as 60MW, 70MW, 55MW, 65MW, and 65MW separately. Thermal unit on bus 2 is replaced with hydropower with rated capacity of 126MW. The renewable penetration accounts for 39.62% of total installed capacity. The topological changes are the transmission lines switching at line 37 (from bus 29 to bus 30), line 1 (from bus 1 to bus 2), line 8 (from bus 5 to bus 7), line 7 (from bus 4 to bus 6), and line 2 (from bus 1 to bus 3).

Case2 : Modified IEEE 118-bus system as shown in Fig.6. the generators on bus 10, 25, 26, 46, 49 and 54 are replaced with renewables, and the maximum output is set as 560MW, 328MW, 185MW, 118MW, 304MW, and 148MW separately. The topological changes are the transmission lines switching at line 44 (from bus 15 to bus 33), line 39 (from bus 17 to bus 31), line 45 (from bus 19 to bus 34), line 38 (from bus 26 to bus 30), line 51 (from bus 37 to bus 38), and line 98 (from bus 49 to bus 66).

For comparison, the interior point method (IPM) is used as a benchmark solver for the SCOPF problem. 6 unsupervised methods (ChebNet-2, ChebNet-3, ChebNet-4, ChebNet-5, GCN, GAT) are tested to determine the optimal order of aggregate neighbor nodes, and a supervised model (ChebNet-4-label) is used to analyze the guiding effect of physical mechanism on the model. ChebNet-2, ChebNet-3, ChebNet-4, ChebNet-5 indicate the aggregation of 2-hop, 3-hop, 4-hop, and 5-hop neighbor features respectively. As for the PGNN, its hidden layer consists of three graph convolution layers, and output layer is set to fully connected layer. Leaky ReLU function is used as the activation function. Other parameters are shown in Tab. I.

Our testing hardware environment is a computer with CPU Processor Intel(R) Core(TM) I7, main frequency 3.30GHz, memory 16G, and GPU GTX2080Ti. The codes are built upon Pytorch.

B. Comparative Study

To illustrate the training accuracy of proposed method, the loss curves tested on IEEE 30-bus system are shown in Fig.7.
TABLE I
SOME PARAMETERS OF THE MODE

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Case1</th>
<th>Case2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of neurons per layer</td>
<td>[64 128 256 512]</td>
<td>[64 128 256 512]</td>
</tr>
<tr>
<td>The hop of neighbor nodes</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Number of training samples</td>
<td>50 000</td>
<td>50 000</td>
</tr>
<tr>
<td>Number of test samples</td>
<td>6820</td>
<td>6820</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>Batchsize</td>
<td>256</td>
<td>256</td>
</tr>
</tbody>
</table>

Fig. 7. Loss curves training set and test set.

To facilitate comparison and observation, the loss of ChebNet-4-label increases to 100 times of the actual value.

From Fig.7 we can see that ChebNet-4 outperforms others. From comparison between testing loss and training loss, the testing losses of GCN, GAT and ChebNet models are similar with their training loss, while testing loss of supervised method (e.g., ChebNet-4-label) is significantly larger than training loss of it, and the loss of GCN and GAT is greater than that of other PGNN models. In GAT model, every vertex performs attention operation for any vertex on the graph, which breaks away from the topology structure of graph. However, the power flow is transmitted from one node to the connected node, which is closely related to the topological structure, so GAT didn’t do as well as ChebNet model. The order of aggregated neighbors represents the order of the polynomial used to fit each frequency component. The essence of GCN is ChebNet-1, which only aggregates the 1-hop feature, so the accuracy of constraint conditions reflected by training loss cannot well meet the requirements. Loss of ChebNet-2, ChebNet-3 and ChebNet-4 decreases with the increase of neighbor order. Loss of ChebNet-5 is greater than that of ChebNet-4, as order increase makes GNN over smooth. What stands out in these comparative studies is that ChebNet-4 features better generalizability.

To further illustrate the ability of the proposed method to satisfy physical constraints, box diagram of constraint violation of generator active and reactive power, branch power flow and bus voltage amplitude are shown in Fig.8. It is found that accuracy of ChebNet-4 is superior to other ChebNet models. Reactive power unbalance is less than 4e-5 and active power unbalance is less than 2e-5. Violations of ChebNet-2, ChebNet-3 and ChebNet-4 decreases with the increase of order. Phase angle loss reflected well by the supervised method (ChebNet-4-label), reaches 0.04rad.

To illustrate the computational efficiency, training time () and test time of each model are shown in Tab.II. In training stage, the training time of unsupervised model is greatly increased compared with supervised method for the model physics-guided needs a lot of trigonometry operation. Both GCN and GAT models only aggregate the features of 1-hop neighbors, and the test time is relatively small. With the increase of aggregation neighbor order, test time increases slightly, but it is far less than the calculation time of IPM. For case1, IPM and ChebNet-4 tested at 1.372s and 0.0061s respectively, with an acceleration ratio of 224.92. The test time of IPM and ChebNet-4 is 2.5310s and 0.0101s respectively, and the acceleration ratio is 250.59 for case 2.

Generation cost and renewable curtailment are shown in Fig.9. You can see that the cost and renewable curtailment calculated by methods ChebNet and IPM are pretty close. This supports a conclusion that, the proposed PGNN indeed enables capture of electric physics.

To analysis the stability, 8620 sets of test data were randomly selected, with a total of 258600 bus voltages. The box diagrams of voltage amplitude calculated by ChebNet-4-label and ChebNet-4 are shown in Fig.10.

As shown in Fig.10, voltage amplitude calculated by Chebnet-4 is located in a secure range of [0.94, 1.06]. Look backward to traditional supervised learning (i.e., ChebNet-4-label), numerous operating conditions are run beyond security...
boundary. This shows that the physics-guided ChebNet-4 significantly outperform conventional experience dominated learning way in terms of stability and reliability. This striking improvement of our method is undoubtedly profited from better physical law digging.

To detailly show performance of each tested methods, Tab.III is given, where operational cost, renewable curtailment (RC), and power loss in 4 testing scenarios are included.

The calculation result of ChebNet-4-label shows that the renewable curtailment is negative (scenarios 1 and 4), which is contrary to the reality and cannot meet the actual demand. When the load is large and the renewable cannot fully meet the demand, the computational cost of ChebNet-4 is slightly lower than that of IPM, while when the load is small and the renewable is sufficient, the cost of the proposed method is similar with that of IPM. The branch power flow corresponding to the above four scenarios is shown in Fig.11. From the graph below we can see that the power flow distribution calculated by the proposed method is basically consistent with that calculated by IPM and satisfy the branch power flow limits.

To verify if the proposed method could provide feasible operation that satisfy Kirchhoff’s law, 1,000 unseen scenarios were randomly selected. Here, only error box diagram by ChebNet-4 is given in Fig.12, since others resulted in the similar outcomes.

It can be found that the bus voltage amplitude calculated by ChebNet-4 model is basically consistent with that calculated by power flow verification calculation, with the error less than 0.015 and the phase Angle error less than 3 degrees. The power flow calculation is verified, indicating that the model based on physics has high reliability.

### C. Comparative Study with Other Topology-Adaptive Learning Methods

To illustrate the topology adaptability of the proposed method, traditional deep learning method (T-DL), traditional deep unsupervised learning method (T-DUL), supervised method using branch admittance as topological feature (TO-DL), and unsupervised method (TO-DUL) using branch admittance as topological feature, are compared with the proposed method. We designed index $P_P$, $P_Q$, the probability that the bus power unbalance exceeds average value 0.05MW, 0.06MVar, to measure the satisfaction of the overall power flow constraint for large number of scenarios. The index calculation results of training set and test set are shown in Tab. IV.

Tab. IV shows that three unsupervised methods physics-guided can complete optimal power flow calculation. For case IEEE 30 with small amount of data, the results of the three...
models can well satisfy the power constraints. The probability of node power unbalance greater than 0.05MW, 0.6MVar increases as the node number increase. There is no topological information in T-DUL, which results in large unbalance of bus power flow in training stage.

The branch admittance data is used to represent the topological structure in TO-DUL, but its variation is small and easy to be drowned by a large amount of other variation information, so the deep neural network cannot effectively reflect the power flow characteristics of IEEE 118. Probability of power unbalance exceeding 0.5MV, 0.6MVar is 90 and 50 times lower than that of deep learning without considering topology and deep learning through topology encoding. What’s more, the probability of proposed method in training set and test set is basically the same, while probability of the supervised methods in test set are much higher than that in training set.

Because topology information is not effectively utilized, the probability of node power flow unbalance in the test set is 3-4 times higher than that in the training set. By comparison, it can be found that the GNN model introduced with topology structure can effectively extract the topology information of the power grid, and the model has high training accuracy. The good performance on the test set indicates that PGNN better in generalization.

### V. Conclusion

Real-time optimal power flow (OPF) is the key to realize safe and efficient operation of renewable-penetrated power grid. However, traditional physics-based method is unable to meet the timeliness demand of fast varying power grid operations. Emerging data-driven method is fast enough, but pure data-driven methods suffer from limited generalizability, especially under the circumstance that power grid topology frequently changes. To overcome these challenges, this paper proposed a physics-guided graph neural network (PGNN)-based method towards real-time OPF. Benefited from physics-guided framework, the proposed PGNN can automatically learn ACOPF solving patterns without any additional preparation of labeled data. Moreover, nature merit of GNN allows us to seamlessly involve power grid topology features into offline learning, such that a topology-adaptability can be readily achieved, which is a fairly hard issue for other data-driven methods. Numerical studies justified generalizability, solving accuracy and remarkable efficiency of the proposed method.

### References


