Recurrent Neural Network for Nonconvex Economic Emission Dispatch

Jiayu Wang, Xing He, Junjian Huang, and Guo Chen

***Abstract*—In this paper, an economic emission dispatch (EED) model is developed to reduce fuel cost and environmental pollution emissions. Considering the development of new energy sources in recent years, the EED problem involves thermal units with the valve point effect and WTs. Meanwhile, it complies with demand constraint and generator capacity constraints. A recurrent neural network (RNN) is proposed to search for local optimal solution of the introduced nonconvex EED problem. The optimality and convergence of the proposed dynamic model are given. The RNN algorithm is verified on a power generation system for the optimization of scheduling and minimization of total cost. Moreover, a particle swarm optimization (PSO) algorithm is compared with RNN under the same problematic frame. Numerical simulation results demonstrate that the optimal scheduling given by RNN is more precise and has lower total cost than PSO. In addition, the dynamic variation of power load demand is considered and the power distribution of eight generators during 12 time periods is depicted.**

***Index Terms*—Recurrent neural network (RNN), nonconvex economic emission dispatch, optimization problem.**

I. Introduction

In order to control the total cost of power generation and reduce the environmental effect of gaseous pollutants, the economic emission dispatch (EED) problem has become a hot topic in power industry. The EED problem, which is subject to the demand constraint and generator capacity constraints at the same time, aims to optimize the scheduling of all power generators. However, EED problems are not convex objectives due to the valve point effect (VPE). Many local minimums that represent nonconvex feature are introduced to the objective problem because of the VPE, which increases the difficulty of finding the optimal solution.

To cope with EED problems, many studies have been carried out in electric system using diverse optimization algorithms. Various nature-inspired heuristic techniques still play an important role in solving these problems. For example, an improved particle swarm optimization (PSO) [1] was demonstrated to solve an EED problem. In the previous year, a PSO with time varying operators was constructed to find global optimal solution, which was designed for optimization problems with hard combination constraints such as nonconvex discontinuity problem [2]. A hybrid approach was employed in [3]. PSO and evolutionary programming (EP) were integrated to solve nonconvex economic load dispatch (ELD) problem. It had been observed from simulation results that all algorithms could find a reasonable global solution over a period of time. A gravitational acceleration based PSO (GAEPSO) [4] could search for optimal solution of economic dispatch (ED) problem, which was applied to a power system containing wind turbines (WTs) and thermal generators. An ED problem was equated with a two-objective problem [5]. Firefly algorithm (FFA) and real genetic algorithm (RGA) were adopted to solve the objective optimization problem. The effectiveness of the introduced method was backed by simulation results. To solve the ELD problem with continuous and nonconvex function, an efficient optimization approach [6] was proposed on the strength of real-coded genetic algorithm (RCGA). Numerical results manifested that the RCGA held higher operation efficiency than binary-coded genetic algorithm (BCGA). Considering the VPE, artificial bee colony (ABC) as a novel neural network algorithm was adopted to deal with ED problem, which proved the validity of the proposed approach in multifarious test system [7].

Most EED problems are settled by nature-inspired heuristic techniques so far. They can converge to optimal value at a fast rate. However, nature-inspired heuristic techniques are insufficient in handling constraints, accurate local search. Meanwhile, they lack the support of mathematical theoretic background. The techniques converge to different solutions instead of a precise optimal solution in each run. Compared with nature-inspired heuristic techniques, the neurodynamic optimization methods own a more obvious advantage in

Manuscript received: December 14, 2018; accepted: November 8, 2019. Date of CrossCheck: November 8, 2019. Date of online publication: April 9, 2020.

This work was supported by the Fundamental Research Funds for the Central Universities (No. XDJK2019B010), the Natural Science Foundation of China (No. 61773320), the Natural Science Foundation Project of Chongqing Science and Technology Commission (CSTC) (No. cstc2018jcyjAX0583, No. cstc2018 jcyjAX0810), the Research Foundation of Key Laboratory of Machine Perception and Children's Intelligence Development funded by Chongqing University of Education (CQUE) (No. 16xjpt07), and the Foundation of Chongqing University of Education (No. KY201702A).

J. Wang and X. He (corresponding author) are with Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China (e-mail: 741112320@qq.com; hexingdoc@swu.edu.cn).

J. Huang is with Key Laboratory of Machine Perception and Children's Intelligence Development, Chongqing University of Education, Chongqing 400067, China (e-mail: hmomu@sina.com).

G. Chen is with the School of Electrical Engineering and Computer Science, University of Newcastle, Newcastle, NSW 2308, Australia (e-mail: guo.chen@unsw.edu.au).

DOI: 10.35833/MPCE.2018.000889

searching for exact local solutions under strict constraints, because the veracity and stability of objective solutions could be justified mathematically. Many neurodynamic optimization approaches had been applied to deal with optimization problems due to these advantages. A modified back propagation neural network was designed in [8], which was applied to solve the combined economic and emission dispatch (CEED) problem. Dynamic neural networks (DNNs) were demonstrated to solve a CEED problem with a fixed load demand, and the convergence could be proved [9]. An adaptive Hopfield neural network (HNN) could solve the CEED problem [10], by means of changing the slope of activation function to seek solutions of proximate Pareto. A novel method was developed to get the best compromise solution of the ED model under real-time state [11]. It was observed from numerical results that the proposed abductive reasoning network (ARN) was superior than some nature-inspired heuristic techniques. The method of conic programming [12] was utilized to deal with a nonconvex ED problem. Numerical results based on an IEEE 33-bus system showed the validity of the represented technique.

The optimization problems can be found in many technical and scientific applications [13]-[15]. Based on the ED problem, a oblivious routing economic dispatch (ORED) algorithm [16] was designed to optimize the total power generation. This algorithm was suitable for both radial and non-radial networks. The superiority of the proposed method in managing congestion and minimizing power was elucidated. To cope with the uncertainties in renewable generators and load demand, an algorithm for solving semidefinite programs (SDPs) sequences was adopted in [17] based on iterative re-weighted norm approximation.

A three-level decomposition algorithm [18] is presented to settle down a distribution expansion planning problem, which was transformed into a convex problem in advance. In [19], an iterative AC optimal power flow (AC-OPF) was adopted to deal with mixed-integer linear programming (MILP) problem. Non-convex models can be easily linearized or convexed by using the MILP method. The validity of the algorithm was proved on 33-bus distribution networks. A recurrent neural network (RNN) was proposed for solving nonconvex optimization problems [20], which showed stability for the objective minimum solution. The solutions were unstable at the maximum and saddle points. When specific conditions were unsatisfied for the minimum solution, it could be solved by constructing an equivalent problem based on original problem by means of p-power method.

In this paper, considering the stability and superiority of neurodynamic optimization, an RNN algorithm is adopted to solve a nonconvex EED problem. A PSO algorithm is compared with the proposed algorithm to solve the nonconvex optimization problem. Our main contributions are as follows:

1) An EED model is introduced which incorporates thermal units and WTs. The model is subject to demand constraints and generator capacity constraints.

2) An RNN algorithm is proposed and compared with PSO. With the same power load demand, simulation results show that the optimal scheduling given by RNN is more precise and has lower total cost than PSO. The RNN algorithm can be applied to solve the nonconvex optimization problem.

3) The power distribution of eight generators during 12 time periods is derived by using the RNN algorithm. Simulation results indicate that the proposed algorithm is applied to multi-period economic dispatch.

The remainder of this paper is as follows. Section II addresses the problem formulation. In Section III, the RNN algorithm is described. In Section IV, optimization solutions are obtained and a PSO algorithm combined with the penalty function method is introduced. Finally, the paper is concluded in Section V.

II. Problem Representation

Reducing the emissions of gaseous pollutants is crucial to power industry. In recent years, renewable energy sources start playing an important role in power generation system. As shown in Fig. 1, we introduce an architecture of the proposed EED model which consists of conventional thermal units, WTs and energy storage system (ESS). Popularly, the objective optimization functions and constraints are indicated in mathematical terms as follows.

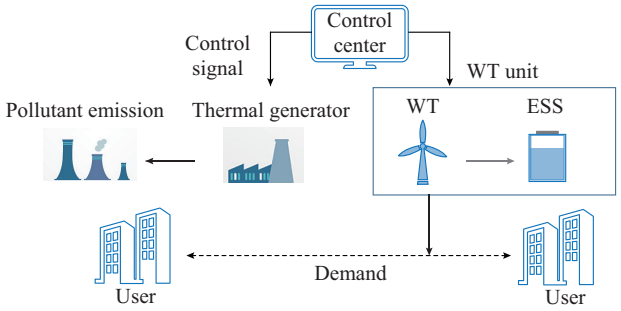

Fig. 1. Architecture of proposed EED model.

A. Objective Problem

1) Minimization of Fuel Cost

For each thermal unit, fuel cost is a quadratic function generally if the VPE is ignored [21]. But VPE brings ripples in the heat rate curve [22], which results in non-convexity of the fuel cost function as shown in Fig. 2.

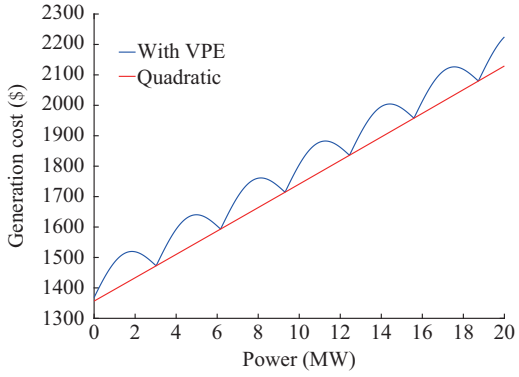

Fig. 2. Effect caused by VPE.

Currently, the fuel cost function can be described as:

$$F\left(P_{G,i}\right)=a_i+b_iP_{G,i}+c_iP_{G,i}^2+\left|e_i\sin\left(f_i\left(P_{G,i}^{\min}-P_{G,i}\right)\right)\right| \quad (1)$$

where $P_{G,i}$ is the real output power of the i^{th} thermal unit; a_i, b_i and c_i are the cost coefficients for the i^{th} unit; e_i and f_i are the coefficients of VPE; and $P_{G,i}^{\min}$ is the minimum generation limit of the i^{th} thermal unit. The cost function contains sine term, and its generalized gradient is remarked in Appendix A.

2) Minimization of Pollutant Emissions Cost

At present, fossil fuels are still the main source of power generation. However, fossil fuel power generation creates a lot of gaseous pollutants. The pollutant emission cost is described as:

$$E\left(P_{G,i}\right)=\sum_{i=1}^{N_G}10^{-2}\left(\alpha_i+\beta_iP_{G,i}+\gamma_iP_{G,i}^2\right) \quad (2)$$

where N_G is the total number of thermal units in the grid; and α_i, β_i, γ_i are the pollutant emission cost coefficients of the i^{th} unit.

3) Minimization of WT Generation Cost

ESSs considered in this study are currently utilized for integration support of wind power so as to improve the controllability of intermittent wind power. The power output of WTs can be ensured, which is equivalent to the predetermined wind power. The underestimation and overestimation costs of wind power are also considered [23]. Thus, the cost of WT is a fixed model in [24], [25], which is denoted as:

$$W\left(Y_{W,j}\right)=\delta_jY_{W,j}+C_{pwj}E\left(Y_{W,j}^{ue}\right)+C_{rwj}E\left(Y_{W,j}^{oe}\right) \quad (3)$$

where $Y_{W,j}$ is the power output of the j^{th} WT; δ_j is the liner cost coefficient of the j^{th} WT; C_{pwj} is the cost coefficient of the underestimated availability for the j^{th} WT; C_{rwj} is the cost coefficient of the overestimated availability; $C_{pwj}E\left(Y_{W,j}^{ue}\right)$ is the penalty cost that all available wind power cannot be used absolutely; and $C_{rwj}E\left(Y_{W,j}^{oe}\right)$ is the overestimated cost, which means the cost of purchasing power from ESSs due to insufficient wind power. $E\left(Y_{W,j}^{ue}\right)$ is described as:

$$\begin{aligned}E\left(Y_{W,j}^{ue}\right)=&\left(W_r-Y_{W,j}\right)\left(\exp\left(-\frac{v_r^k}{c^k}\right)-\exp\left(-\frac{v_{out}^k}{c^k}\right)\right)+\\&\left(\frac{W_rv_{in}}{v_r-v_{in}}+Y_{W,j}\right)\left(\exp\left(-\frac{v_r^k}{c^k}\right)-\exp\left(-\frac{v_1^k}{c^k}\right)\right)+\\&\frac{W_rc}{v_r-v_{in}}\left[\Gamma\left(1+\frac{1}{k},\left(\frac{v_1}{c}\right)^k\right)-\Gamma\left(1+\frac{1}{k},\left(\frac{v_r}{c}\right)^k\right)\right]\end{aligned} \quad (4)$$

where k is the shape factor of the Weibull distribution for wind; c is the scale factor of the Weibull distribution for wind; W_r is the rated wind power; v_r is the wind speed; v_{in} and v_{out} are the cut-in and cut-out wind speeds, respectively; v_1 is an intermediary parameter defined as $v_1=v_{in}+\left(v_r-v_{in}\right)Y_{W,j}/W_r$; $\exp(\cdot)$ is the exponential function; and $\Gamma(\cdot)$ is the standard incomplete gamma function. The generalized gradient of the standard gamma function is given in Appendix A. Similarly, $E\left(Y_{W,j}^{oe}\right)$ is defined as:

$$\begin{aligned}E\left(Y_{W,j}^{oe}\right)=&Y_{W,j}\left(1-\exp\left(-\frac{v_{in}^k}{c^k}\right)+\exp\left(-\frac{v_{out}^k}{c^k}\right)\right)+\\&\left(\frac{W_rv_{in}}{v_r-v_{in}}+Y_{W,j}\right)\left(\exp\left(-\frac{v_{in}^k}{c^k}\right)-\exp\left(-\frac{v_1^k}{c^k}\right)\right)+\\&\frac{W_rc}{v_r-v_{in}}\left[\Gamma\left(1+\frac{1}{k},\left(\frac{v_1}{c}\right)^k\right)-\Gamma\left(1+\frac{1}{k},\left(\frac{v_{in}}{c}\right)^k\right)\right]\end{aligned} \quad (5)$$

4) Objective Function

Three objective functions are combined into a single function to search for the compromise generator schedule. Inspired by the method of assigning prices to emissions [26], the single-objective EED problem is expressed as follows:

$$FO=F\left(P_{G,i}\right)+\sigma_{Tr}E\left(P_{G,i}\right)+W\left(Y_{W,j}\right) \quad (6)$$

where σ_{Tr} is the price penalty factor that is combined with emission function for obtaining the total fuel cost function. The method to obtain σ_{Tr} is given in Appendix B.

B. Equality and Inequality Constraints

In this paper, equality and inequality constraints are considered, which contain demand constraint and generator capacity constraints. These constraints are formulated as follows.

1) Power Demand Constraint

Transmission losses should be considered in practical production. According to [23], [27], transmission losses are denoted by quadratic equations. The transmission losses P_{WTL} induced by the j^{th} WT can be given as $P_{WTL}=B_{W,j}Y_{W,j}^2$, and the transmission losses P_{GTL} caused by the i^{th} thermal unit can be expressed as $P_{GTL}=B_{G,i}P_{G,i}^2$, where $B_{W,j}$ and $B_{G,i}$ are the coefficients of power losses.

The total power load demand P_L is less than or equal to the total power generation minus the total real transmission losses. The EED problem complies with the demand constraint under the power load demand P_L as follows:

$$\sum_{i=1}^{N_G}\left(P_{G,i}-B_{G,i}P_{G,i}^2\right)+\sum_{j=1}^{N_W}\left(Y_{W,j}-B_{W,j}Y_{W,j}^2\right)\geq P_L \quad (7)$$

where N_W is the total number of WTs in the system.

2) Generator Capacity Constraint

To ensure stable operation, power output of each generator will be limited between its minimum limit and maximum limit. The power limits of thermal units and WTs are as follows:

$$P_{G,i}^{\min}\leq P_{G,i}\leq P_{G,i}^{\max} \quad (8)$$

$$Y_{W,j}^{\min}\leq Y_{W,j}\leq Y_{W,j}^{\max} \quad (9)$$

C. Problem Formulation

The EED problem with the purpose of minimizing (6) is subject to some constraints denoted above. At last, the optimization problem is mathematically expressed as:

$$\begin{cases} \min FO = F\left(P_{G,i}\right) + \sigma_{Tr} E\left(P_{G,i}\right) + W\left(Y_{W,j}\right) \\ \text{s.t.} \ \sum_{i=1}^{N_G}\left(P_{G,i} - B_{G,i}P_{G,i}^2\right) + \sum_{j=1}^{N_W}\left(Y_{W,j} - B_{W,j}Y_{G,j}^2\right) \geq P_L \\ P_{G,i}^{\min} \leq P_{G,i} \leq P_{G,i}^{\max} \\ Y_{W,j}^{\min} \leq Y_{W,j} \leq Y_{W,j}^{\max} \end{cases} \quad (10)$$

III. Algorithm Model Description

In this section, an RNN is introduced [20], [28]. Its network architecture is represented in Fig. 3.

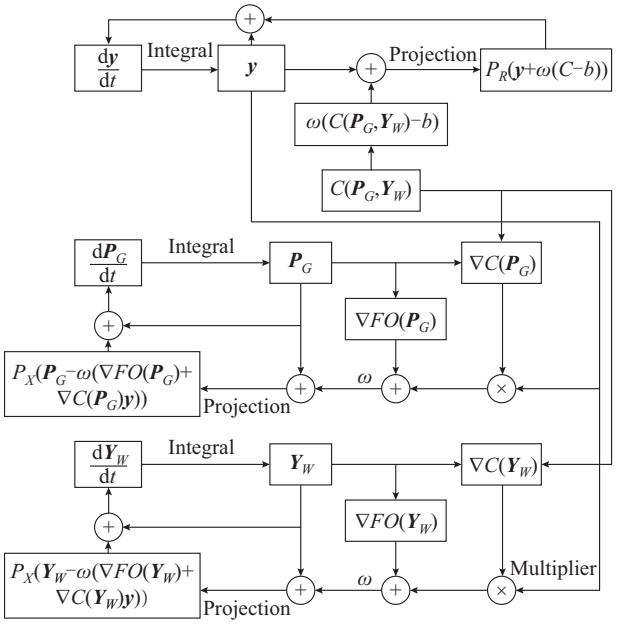

Fig. 3. Network architecture of proposed RNN.

The projection operator and the gradient are used to build RNN model. The following neural network to solve objection problem is obtained.

The dynamic equation is given as below:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\boldsymbol{P}_G \\ \boldsymbol{Y}_W \\ \boldsymbol{y}\end{bmatrix} = \begin{bmatrix} P_X\left(\boldsymbol{P}_G - \omega\left(\nabla FO\left(\boldsymbol{P}_G\right) + \nabla C\left(\boldsymbol{P}_G\right)\boldsymbol{y}\right)\right) - \boldsymbol{P}_G \\ P_X\left(\boldsymbol{Y}_W - \omega\left(\nabla FO\left(\boldsymbol{Y}_W\right) + \nabla C\left(\boldsymbol{Y}_W\right)\boldsymbol{y}\right)\right) - \boldsymbol{Y}_W \\ P_R\left(\boldsymbol{y} + \omega\left(C - b\right)\right) - \boldsymbol{y} \end{bmatrix} \quad (11)$$

where $\omega > 0$ is a control factor; $y \in \mathbf{R}$ is a Lagrange multiplier; $\nabla(\cdot)$ is the gradient; and $b = -P_L$. Define $P_{\Omega}\left(X_o\right)$ as the projection operator expressed as $P_{\Omega}\left(X_o\right) = \arg\min\left\|X_o - y\right\|$, where $\boldsymbol{\Omega}$ is the projection object and X_o is an abstract projection variable. At the same time, the inequality constraint (12) is equivalent to the demand constraint (7).

$$C = -\sum_{i=1}^{N_G}\left(P_{G,i} - B_{G,i}P_{G,i}^2\right) - \sum_{j=1}^{N_W}\left(Y_{W,j} - B_{W,j}Y_{W,j}^2\right) \leq -P_L \quad (12)$$

Specially, $\boldsymbol{P}_X\left(\boldsymbol{P}_G\right) = \left[P_X\left(P_{G,1}\right), P_X\left(P_{G,2}\right), ..., P_X\left(P_{G,N_G}\right)\right]^{\mathrm{T}}$.

For a box set $X = \left\{\boldsymbol{P}_G \in \mathbf{R}^m \,\middle|\, \boldsymbol{P}_G^{\min} \leq \boldsymbol{P}_G \leq \boldsymbol{P}_G^{\max}\right\}$, where $\boldsymbol{P}_G^{\min} = \left[P_{G,1}^{\min}, P_{G,2}^{\min}, ..., P_{G,N_G}^{\min}\right]$ and $\boldsymbol{P}_G^{\max} = \left[P_{G,1}^{\max}, P_{G,2}^{\max}, ..., P_{G,N_G}^{\max}\right]$, $P_X\left(P_{G,i}\right)$ and $P_X\left(Y_{W,j}\right)$ are defined as (13) and (14), respectively.

$$P_X\left(P_{G,i}\right) = \begin{cases} P_{G,i}^{\min} & P_{G,i} \leq P_{G,i}^{\min} \\ P_{G,i} & P_{G,i}^{\min} < P_{G,i} < P_{G,i}^{\max} \\ P_{G,i}^{\max} & P_{G,i}^{\max} \leq P_{G,i} \end{cases} \quad (13)$$

$$P_X\left(Y_{W,j}\right) = \begin{cases} Y_{W,j}^{\min} & Y_{W,j} \leq Y_{W,j}^{\min} \\ Y_{W,j} & Y_{W,j}^{\min} < Y_{W,j} < Y_{W,j}^{\max} \\ Y_{W,j}^{\max} & Y_{W,j}^{\max} \leq Y_{W,j} \end{cases} \quad (14)$$

Simultaneously, $P_R\left(y\right)$ is defined as:

$$P_R\left(y\right) = \begin{cases} 0 & y \leq 0 \\ y & y > 0 \end{cases} \quad (15)$$

Two necessary theoretical problems should be explained for achieving the optimal solution of (10). Firstly, equilibrium points of proposed RNN correspond to Karush-Kuhn-Tucker (KKT) points under some necessary and sufficient conditions. Secondly, the convergence of RNN algorithm is proved. The detailed theoretical analysis is shown in Appendix C. For further proof, the following definition is given.

Definition 1: a solution $\boldsymbol{J} = \left[\boldsymbol{P}_G^{\mathrm{T}}, \boldsymbol{Y}_W^{\mathrm{T}}, \boldsymbol{y}^{\mathrm{T}}\right]^{\mathrm{T}}$ of the RNN (11) represents a feasible solution. When the gradients of $C_j, \nabla C_j, \forall j \in \left\{\boldsymbol{J} \,\middle|\, C + P_L = 0\right\}$ are linearly independent, a feasible solution J is said to be a regular point.

IV. Simulation Results

A. Results of RNN Approach

In this part, the proposed RNN algorithm is verified on an eight-generator system which consists of six thermal units ($P_{G,i}$, $i = 1, 2, ..., 6$) and two WTs ($Y_{W,j}$, $j = 1, 2$). The simulation data are given in Table I and Table II, which can be found in [24], [26].

TABLE I
Data for Six Thermal Units

Unit	$P_{G,i}^{\min}$ (MW)	$P_{G,i}^{\max}$ (MW)	a_i (\$/h)	b_i (\$/(MW²·h))	c_i (\$/h)	e_i (\$/h)	f_i (\$/h)	α_i (\$/h)	β_i (\$/(MW²·h))	γ_i (\$/h)	B_{Gi} (\$/h)
1	70	470	1356.659	38.2704	0.01799	84.0	0.08400	42.8955	–0.5112	0.00460	0.00031
2	20	80	950.606	39.5804	0.10908	52.5	0.10610	350.0056	–3.9524	0.04652	0.00011
3	20	130	800.705	39.5104	0.12111	63.0	0.10010	330.0056	–3.9023	0.04625	0.00041
4	47	120	900.705	36.5104	0.12511	67.2	0.09660	330.0056	–3.9023	0.04625	0.00022
5	50	160	756.799	38.5390	0.15274	63.0	0.08820	13.8593	0.3277	0.00420	0.00045
6	10	55	1000.430	40.5407	0.12951	69.3	0.01073	360.0012	–3.9864	0.04702	0.00021

TABLE II
DATA FOR TWO WTS

Parameter	Value
v_{in} (m/s)	5
v_{out} (m/s)	45
v_r (m/s)	15
c (m/s)	8
k (m/s)	2
δ_j ($/h)	6
C_{pwj} ($/h)	3.1
C_{rwj} ($/h)	3.1
W_r (MW)	160
B_{Wj} ($/h)	0.00033

During the simulation, the step size of the algorithm is set as $\theta=0.001$, $\omega=15$, and the completion deadline $\tau=16$. Figures 4-6 show the convergence of the eight-generator system in the EED problem. In simulation, the power load demand P_L is set as 580 MW. For each generator, ten arbitrary initial values are selected. The optimal result of Lagrange multiplier y is also given. The Lagrange multiplier is the optimal value of the total system. It is obvious that initial values converge to their optimal solutions after the calculation. The value of convergence does not change when the initial value changes.

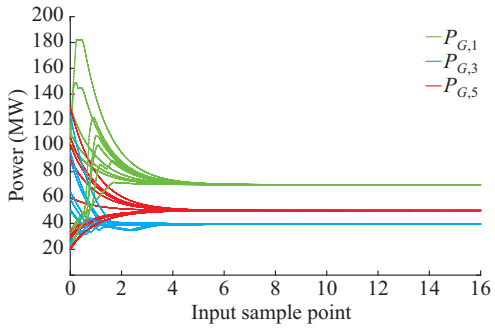

Fig. 4. Optimal results of $P_{G,1}$, $P_{G,3}$, and $P_{G,5}$.

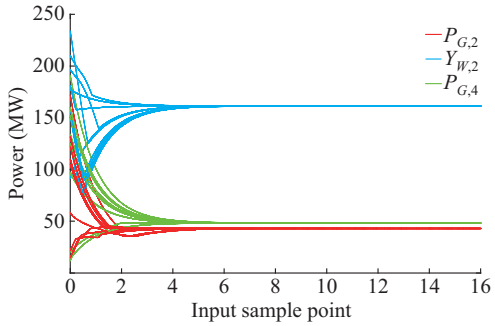

Fig. 5. Optimal results of $P_{G,2}$, $P_{G,4}$, and $Y_{W,2}$.

The above results are obtained based on a fixed power load demand, which is inconsistent with practical applications. A variable load system within 12 time periods of a typical day is utilized to test the proposed RNN method. The load variations of 12 time periods is illustrated in Fig. 7, one of which indicates two hours. Figures 8 and 9 present the load variation and power distribution of four generators during 12 time periods, respectively. Power load demands in 12 time periods are elucidated in Table III. The detailed power distribution values are given in Table IV. It can be seen that the optimal power of two WTs is almost approximate.

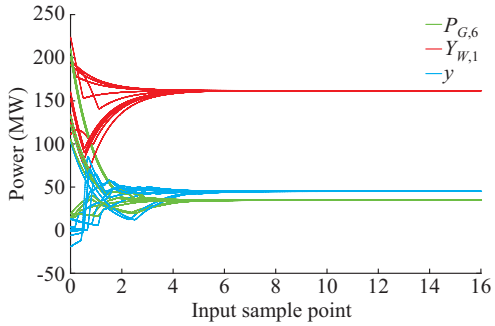

Fig. 6. Optimal results of $P_{G,6}$, $Y_{W,1}$, and y.

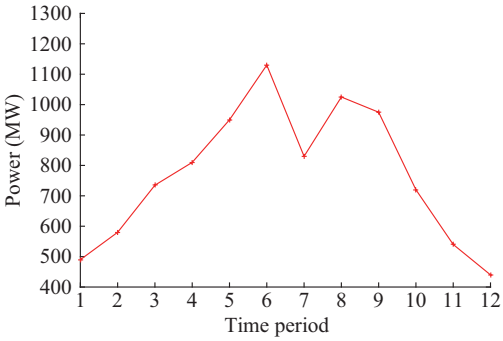

Fig. 7. Load variations in 12 time periods.

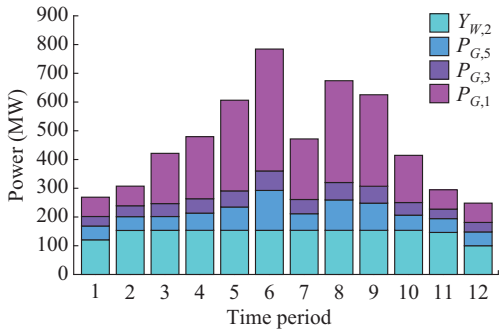

Fig. 8. Power distribution of four generators in 12 time periods.

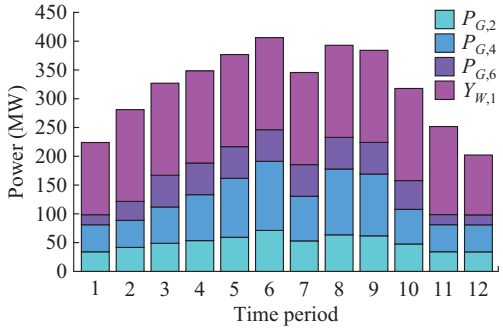

Fig. 9. Power distribution of four generators in 12 time periods.

TABLE III
LOAD VARIATIONS OF 12 TIME PERIODS ON A TYPICAL DAY

Time period	Load (MW)
1	490
2	580
3	735
4	810
5	950
6	1130
7	830
8	1025
9	975
10	720
11	540
12	440

B. Comparison with PSO Approach

In a classical PSO approach, each particle in a swarm stands for a potential solution. Their initial positions are randomly generated and then updated in a specific way. Each particle has the ability to learn experience from its own and its neighbors. Compared with other stochastic methods, PSO can obtain high-quality solutions in shorter computation time and has more stable convergence characteristic [31], [32].

However, PSO lacks the ability to precisely find local optimal solution and handle complex constraints. The method of penalty function can transform a constrained problem into an equivalently unconstrained problem. Many related researches have been done [33] - [37]. In this paper, a PSO algorithm combined with the penalty function is introduced to seek local optimal solution.

TABLE IV
POWER DISTRIBUTION OF EIGHT GENERATORS IN 12 TIME PERIODS

Power load (MW)	$P_{G,1}$ (MW)	$P_{G,2}$ (MW)	$P_{G,3}$ (MW)	$P_{G,4}$ (MW)	$P_{G,5}$ (MW)	$P_{G,6}$ (MW)	$Y_{W,1}$ (MW)	$Y_{W,2}$ (MW)
490	70.0000	34.0404	34.4908	47.0000	50.0000	17.4646	125.4650	125.6138
580	71.0638	41.6987	39.5341	47.0000	50.0000	32.8518	159.5350	159.5350
735	182.1369	48.7500	46.9003	63.1913	50.0000	55.0000	160.0000	160.0000
810	224.9520	53.5094	52.1057	79.8193	62.4147	55.0000	160.0000	160.0000
950	328.6094	59.2444	58.4348	102.5037	84.5272	55.0000	160.0000	160.0000
1130	441.7915	71.2183	70.6367	120.0000	144.5669	55.0000	160.0000	160.0000
830	219.5369	53.0121	51.5566	77.5490	60.1725	55.0000	160.0000	160.0000
1025	369.0382	63.5956	63.0957	114.2751	110.1540	55.0000	160.0000	160.0000
975	331.6637	61.8129	61.2114	107.2863	98.5072	55.0000	160.0000	160.0000
720	171.0769	47.6572	45.7272	60.1714	55.0000	50.0000	160.0000	160.0000
540	70.0000	34.1303	34.5878	47.0000	50.0000	17.6097	152.9196	152.3581
440	70.0000	33.9469	34.3897	47.0000	50.0000	17.2278	104.0327	104.2030

1) Penalty Function Method

Through the extensive work [38], the objective function is transformed into a penalty function as (16).

$$G(P,\rho,\varsigma,M)=Q(FO(P)-M)+\rho\left[\sum_{j\in I_p}P(C(P)+P_L)\right]^{\varsigma} \quad (16)$$

where $M\in\mathbf{R}$ is the objective penalty parameter; ρ and ς are the constraint penalty parameters and $\rho>0$, $\varsigma>0$; $j\in I_p=\{1,2,...,p\}$; Q and P are both continuous functions satisfying Q, P: $\mathbf{R}^1\to\mathbf{R}^1\cup\{+\infty\}$ and (17). $R(t)$: $\mathbf{R}^1\to\mathbf{R}^1\cup\{+\infty\}$ is a continuous function satisfying (18).

$$\begin{cases} Q(t)=0 & t\le 0\\ P(t)=0 & t\le 0\\ Q(t)>0 & t>0\\ P(t)>0 & t>0\\ Q(t_2)>Q(t_1) & t_2>t_1>0\\ P(t_2)>P(t_1) & t_2>t_1>0 \end{cases} \quad (17)$$

$$\begin{cases} R(t)=0 & t=0\\ R(t)>0 & t\ne 0\\ R(t_2)>R(t_1) & |t_2|>|t_1| \end{cases} \quad (18)$$

The penalty function is constructed as follows:

$$G(P,\rho,\varsigma,M)=\max(FO(P)-M)+\rho\left[\max(C(P)+P_L)\right]^{\varsigma} \quad (19)$$

The optimal solution of penalty function is approximately equal to the optimal solution of objective function, which has been proofed in [38] with more information.

2) PSO

$\boldsymbol{U}_P^i=\left[U_{P,1}^i,U_{P,2}^i,...,U_{P,m}^i\right]^{\mathrm{T}}\in\mathbf{R}^m$ denotes the position of the i^{th} particle, which is generated in a random way. And $\boldsymbol{V}^i=\left[V_{i,1},V_{i,2},...,V_{i,n}\right]^{\mathrm{T}}\in\mathbf{R}^n$ denotes the velocity of the i^{th} particle, where n denotes the dimension of each particle. $\boldsymbol{U}_{pb}^i=[U_{pb,1}^i,U_{pb,2}^i,...,U_{pb,n}^i]^{\mathrm{T}}$ is defined as the best position of the i^{th} particle obtained on the basis of its own experience, and $\boldsymbol{U}_{gb}^i=[U_{gb,1}^i,U_{gb,2}^i,...,U_{gb,n}^i]^{\mathrm{T}}$ is the best particle position based on overall swarm experience. Each particle updates the forward velocity and position vector as follows [39]:

$$V_{i,j}(t+1)=\varpi V_{i,j}(t)+c_1 r_1\left(U^i_{pb,j}-U^i_{P,j}(t)\right)+c_2 r_2\left(U^i_{gb,j}-U^i_{P,j}(t)\right) \tag{20}$$

$$U^i_{P,j}(t+1)=U^i_{P,j}(t)+U_{i,j}(t+1) \tag{21}$$

where ϖ is the inertia weight; c_1 and c_2 are two weighting parameters; and r_1 and r_2 are two random numbers in the range of [0, 1].

If the number of iterations reaches a predefined maximum or a best position, the movement of the swarm will stop. $\kappa(\kappa=0,1,...,\kappa_{\max})$ is an iteration step, where $\kappa_{\max}$ is the maximum iteration step. The conditions for terminating iteration are the same as [40].

PSO has shortcomings in searching exact local optimal solution. In the next part, a comparison of the proposed RNN and PSO is given under the same problem frame.

3) Comparison of Simulation Results

The introduced PSO algorithm is verified on this eight-generator system. The size of the PSO swarm is 100 and the the inertia weight ϖ is 0.9. The weighting parameters c_1, c_2 are both set as 2 [41]. Three different iterations $\kappa=50$, 100, 200 are selected in the simulation. Besides, the object penalty parameters are set as $M=-10$, $\rho=200$ and $\varsigma=2$ according to [38].

Figure 10 presents the optimal total cost of eight-generator system by PSO. The optimal scheduling of eight generators by PSO algorithm are compared with that by RNN algorithm in Table V. It can be seen from the results that PSO algorithm gives different optimal solutions in each run, but RNN algorithm would give stable optimal values of the objective function. It is theoretically ensured that the local optimal solution is found through the RNN algorithm, it cannot be mathematically proved that the local optimal solution can be found through PSO. With the same power load demand $P_L=580$ MW, the optimal scheduling given by RNN is more accurate and has lower total cost than that by PSO.

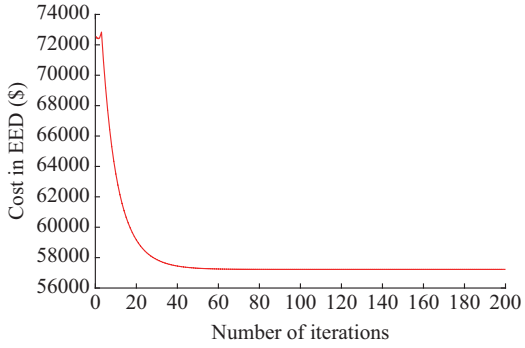

Fig. 10. Optimal total cost by PSO.

As a nature-inspired heuristic technique, PSO lacks mathematically theoretic support, so it may converge to different solutions instead of accurate optimal solution in each run. The proposed RNN algorithm owns a more obvious advantage in searching for accurate local solution with strict constraints, because the stability and veracity of the solutions can be justified mathematically.

TABLE V
COMPARISON OF SIMULATION RESULTS

Constituent	Power distribution of eight-generator system (MW)			
	RNN	PSO ($\kappa=50$)	PSO ($\kappa=100$)	PSO ($\kappa=200$)
$P_{G,1}$	71.0638	73.6879	72.5561	72.3551
$P_{G,2}$	41.6987	42.5431	40.5673	40.8463
$P_{G,3}$	39.5341	39.9431	38.6541	38.8052
$P_{G,4}$	47.0000	51.5678	51.2231	50.3521
$P_{G,5}$	50.0000	53.4721	52.0587	52.1925
$P_{G,6}$	32.8518	35.4321	33.4721	33.7331
$Y_{W,1}$	159.5350	128.6878	154.3741	155.0350
$Y_{W,2}$	159.5350	155.5879	159.2145	158.6878
Total cost ($)	56659	57551	57164	57135

V. CONCLUSION

The EED problem aims to reduce the total fuel cost and pollutant emissions of the power system. An EED model containing thermal units and WTs is introduced in this paper. Meanwhile, it needs to meet the demand constraint and generator capacity constraints. To solve the nonconvex EED problem, an RNN algorithm is proposed to search for local optimal solution. The optimality and convergence of the proposed algorithm are verified by mathematical theory. The RNN algorithm is verified on an eight-generator system to optimize the scheduling and minimize the total cost. Moreover, a PSO algorithm is compared with RNN in solving nonconvex optimization problem. Numerical simulation results in Table V demonstrate that the proposed RNN owns a more obvious advantage in searching for exact local solution than PSO. The optimal scheduling of the system with the same power load demand given by RNN is more precise and has lower total cost. Finally, a variable load system within 12 time periods of a day is utilized to test the proposed RNN method, and the power distribution of eight generators during 12 time periods is depicted.

Only quadratic static constraints and inequality constraints are considered in this paper, and we are going to take into account the network constraints as a future research direction in the proposed problem frame. Also, it is expected to see that more stable and effective algorithms can be used to solve non-convex optimization problem.

APPENDIX A

The generator cost function (1) has sine term. Through the extensive work in [27], the generalized gradient $\partial F_{\sin}\left(P_{G,i}\right)$ is limited to $\left[-e_i f_i,+e_i f_i\right]$ at the non-derivative points. $F_{\sin}\left(P_{G,i}\right)$ is described as:

$$F_{\sin}\left(P_{G,i}\right)=\left|e_i \sin\left(f_i\left(P^{\min}_{G,i}-P_{G,i}\right)\right)\right| \tag{A1}$$

At derivative points, $\partial F_{\sin}\left(P_{G,i}\right)$ is defined as:

$$\partial F_{\sin}\left(P_{G,i}\right)=e_i f_i \cos\left(\left(P_{G,i}-P^{\min}_{G,i}\right)f_i-\left|\frac{\left(P_{G,i}-P^{\min}_{G,i}\right)f_i}{\pi}\right|\pi\right) \tag{A2}$$

where $\lceil\cdot\rceil$ is the round operation realized by *fix*(·) function in MATLAB. In this paper, $\partial\Gamma(\cdot)$, which is the partial differential of the standard incomplete gamma function $\Gamma(\cdot)$, is defined as:

$$\partial\Gamma\left(1+\frac{1}{k},\left(\frac{v_1}{c}\right)^k\right)=-\frac{kv_1^k\left(v_r-v_{in}\right)}{W_r c^{k+1}}\exp\left(-\left(\frac{v_1}{c}\right)^k\right) \tag{A3}$$

From (4) and (A3), we can obtain:

$$\partial E\left(Y_{W,j}^{ue}\right)=\exp\left(-\frac{v_{out}^k}{c^k}\right)-\exp\left(-\frac{v_1^k}{c^k}\right) \tag{A4}$$

Similarly, from (5) and (A3), we can obtain:

$$\partial E\left(Y_{W,j}^{oe}\right)=1+\exp\left(-\frac{v_{out}^k}{c^k}\right)-\exp\left(-\frac{v_1^k}{c^k}\right) \tag{A5}$$

APPENDIX B

The price penalty factor σ_{Tr} for each unit can be calculated as follows [26]:

$$\sigma_{Tr}(i)=\frac{F\left(P_{G,i}^{\max}\right)}{E\left(P_{G,i}^{\max}\right)} \tag{B1}$$

where $F\left(P_{G,i}^{\max}\right)$ is the fuel cost when the output power of the i^{th} generator is the maximum. $F\left(P_{G,i}^{\max}\right)$ can be described as:

$$F\left(P_{G,i}^{\max}\right)=\left(a_i+b_iP_{G,i}^{\max}+c_iP_{G,i}^{\max 2}+\left|e_i\sin\left(f_i\left(P_{G,i}^{\min}-P_{G,i}^{\max}\right)\right)\right|\right) \tag{B2}$$

Similarly, $E\left(P_{G,i}^{\max}\right)$ is the emission when the output power of the i^{th} generator is the maximum. $E\left(P_{G,i}^{\max}\right)$ can be defined as:

$$E\left(P_{G,i}^{\max}\right)=\sum_{i=1}^{N_G}10^{-2}\left(\alpha_i+\beta_iP_{G,i}^{\max}+\gamma_i\left(P_{G,i}^{\max}\right)^2\right) \tag{B3}$$

APPENDIX C

In this part, the existence and convergence of local optimal solution of the proposed model are given.

A. Optimal Analysis

Theorem 1: making a hypothesis that the objective problem is a nonconvex function within a given feasible region. The equilibrium point $\bar{\boldsymbol{J}}=\left[\bar{\boldsymbol{P}}_G^{\mathrm{T}},\bar{\boldsymbol{Y}}_W^{\mathrm{T}},\bar{\boldsymbol{y}}^{\mathrm{T}}\right]^{\mathrm{T}}$ of (11) satisfies the KKT condition of (10).

Proof: the KKT condition for the objective problem (10) can be described as:

$$\begin{cases}P_{G,i}^{\min}\le P_{G,i}\le P_{G,i}^{\max}\\ Y_{W,j}^{\min}\le Y_{W,j}\le Y_{W,j}^{\max}\quad \boldsymbol{y}\ge\boldsymbol{0}\\ C+P_L\le 0\quad \boldsymbol{y}^{\mathrm{T}}\left(C+P_L\right)=\boldsymbol{0}\end{cases} \tag{C1}$$

$$\begin{cases}\left(\nabla FO\left(\boldsymbol{P}_G\right)+\nabla C\left(\boldsymbol{P}_G\right)\boldsymbol{y}\right)_i\le 0 & P_{G,i}=P_{G,i}^{\min}\\ \left(\nabla FO\left(\boldsymbol{P}_G\right)+\nabla C\left(\boldsymbol{P}_G\right)\boldsymbol{y}\right)_i=0 & P_{G,i}^{\min}\le P_{G,i}\le P_{G,i}^{\max}\\ \left(\nabla FO\left(\boldsymbol{P}_G\right)+\nabla C\left(\boldsymbol{P}_G\right)\boldsymbol{y}\right)_i\ge 0 & P_{G,i}^{\max}=P_{G,i}\end{cases} \tag{C2}$$

$$\begin{cases}\left(\nabla FO\left(\boldsymbol{Y}_W\right)+\nabla C\left(\boldsymbol{Y}_W\right)\boldsymbol{y}\right)_j\le 0 & Y_{W,j}=Y_{W,j}^{\min}\\ \left(\nabla FO\left(\boldsymbol{Y}_W\right)+\nabla C\left(\boldsymbol{Y}_W\right)\boldsymbol{y}\right)_j=0 & Y_{W,j}^{\min}<Y_{W,j}<Y_{W,j}^{\max}\\ \left(\nabla FO\left(\boldsymbol{Y}_W\right)+\nabla C\left(\boldsymbol{Y}_W\right)\boldsymbol{y}\right)_j\ge 0 & Y_{W,j}^{\max}=Y_{W,j}=Y_{W,j}^{\min}\end{cases} \tag{C3}$$

According to the KKT theorem [29], the mentioned KKT conditions can be equivalently translated into the following projection formulation:

$$\begin{cases}P_X\left(\boldsymbol{P}_G-\omega\left(\nabla FO\left(\boldsymbol{P}_G\right)+\nabla C\left(\boldsymbol{P}_G\right)\boldsymbol{y}\right)\right)=\boldsymbol{P}_G\\ P_X\left(\boldsymbol{Y}_W-\omega\left(\nabla FO\left(\boldsymbol{Y}_W\right)+\nabla C\left(\boldsymbol{Y}_W\right)\boldsymbol{y}\right)\right)=\boldsymbol{Y}_W\\ P_R\left(\boldsymbol{y}+\omega\left(C+P_L\right)\right)=\boldsymbol{y}\end{cases} \tag{C4}$$

where $\omega>0$. $\boldsymbol{J}^*=\left[\left(\boldsymbol{P}_G^*\right)^{\mathrm{T}},\left(\boldsymbol{Y}_W^*\right)^{\mathrm{T}},\left(\boldsymbol{y}^*\right)^{\mathrm{T}}\right]^{\mathrm{T}}$ satisfies the above formulation if $\left[\left(\boldsymbol{P}_G^*\right)^{\mathrm{T}},\left(\boldsymbol{Y}_W^*\right)^{\mathrm{T}}\right]^{\mathrm{T}}$ is a local optimal solution of (11), which can be expressed as:

$$\begin{cases}P_X\left(\boldsymbol{P}_G^*-\omega\left(\nabla FO\left(\boldsymbol{P}_G^*\right)+\nabla C\left(\boldsymbol{P}_G^*\right)\boldsymbol{y}\right)\right)=\boldsymbol{P}_G^*\\ P_X\left(\boldsymbol{Y}_W^*-\omega\left(\nabla FO\left(\boldsymbol{Y}_W^*\right)+\nabla C\left(\boldsymbol{Y}_W^*\right)\boldsymbol{y}\right)\right)=\boldsymbol{Y}_W^*\\ P_R\left(\boldsymbol{y}^*+\omega\left(C+P_L\right)\right)=\boldsymbol{y}^*\end{cases} \tag{C5}$$

$\bar{\boldsymbol{J}}=\left[\bar{\boldsymbol{P}}_G^{\mathrm{T}},\bar{\boldsymbol{Y}}_W^{\mathrm{T}},\bar{\boldsymbol{y}}^{\mathrm{T}}\right]^{\mathrm{T}}$ is the equilibrium point of RNN. The following relationships must exist:

$$\begin{cases}P_X\left(\bar{\boldsymbol{P}}_G-\omega\left(\nabla FO\left(\bar{\boldsymbol{P}}_G\right)+\nabla C\left(\bar{\boldsymbol{P}}_G\right)\boldsymbol{y}\right)\right)-\bar{\boldsymbol{P}}_G=0\\ P_X\left(\bar{\boldsymbol{Y}}_W-\omega\left(\nabla FO\left(\bar{\boldsymbol{Y}}_W\right)+\nabla C\left(\bar{\boldsymbol{Y}}_W\right)\boldsymbol{y}\right)\right)-\bar{\boldsymbol{Y}}_W=0\\ P_R\left(\bar{\boldsymbol{y}}+\omega\left(C+P_L\right)\right)-\bar{\boldsymbol{y}}=0\end{cases} \tag{C6}$$

Therefore, the equilibrium points of the RNN are in a one-to-one correspondence with the KKT points of (10) exactly. The optimal solution $\bar{\boldsymbol{J}}=\left[\bar{\boldsymbol{P}}_G^{\mathrm{T}},\bar{\boldsymbol{Y}}_W^{\mathrm{T}},\bar{\boldsymbol{y}}^{\mathrm{T}}\right]^{\mathrm{T}}$ is also satisfied with the KKT conditions simultaneously.

B. Convergence Analysis

The Lagrangian function associated with (10) can be defined as:

$$L\left(\boldsymbol{P},\boldsymbol{y}\right)=FO\left(\boldsymbol{P}\right)+\boldsymbol{y}\left(C\left(\boldsymbol{P}\right)+P_L\right) \tag{C7}$$

where $\boldsymbol{P}=\left[\boldsymbol{P}_G^{\mathrm{T}},\boldsymbol{Y}_W^{\mathrm{T}}\right]^{\mathrm{T}}$ satisfying the constraints is a feasible solution of (10). It follows that the Hessian of the Lagrangian function is:

$$\nabla_P^2L\left(\boldsymbol{P},\boldsymbol{y}\right)=\nabla^2FO\left(\boldsymbol{P}\right)+\sum_{i=1}^{m}y_i\nabla^2C_i\left(\boldsymbol{P}\right) \tag{C8}$$

Theorem 2: the proposed neural network can be stable at a KKT point $\boldsymbol{J}^*$ for any initial point $J(t_0)$ if $\nabla_P^2L\left(\boldsymbol{J}^*\right)$ is positive semidefinite on feasible solution, where $\boldsymbol{J}\left(t_0\right)=\left[\boldsymbol{P}_G^{\mathrm{T}}\left(t_0\right),\boldsymbol{Y}_W^{\mathrm{T}}\left(t_0\right),\boldsymbol{y}^{\mathrm{T}}\left(t_0\right)\right]^{\mathrm{T}}$. Simultaneously, the output trajectory of the proposed neural network converges to a global minimum.

Proof: $\boldsymbol{P}^*$ is an equilibrium point, and at the same time, $\boldsymbol{J}^*=$

$\left[\left(\boldsymbol{P}^*\right)^{\mathrm{T}},\boldsymbol{y}^{\mathrm{T}}\right]^{\mathrm{T}}$ is a KKT point. $\boldsymbol{P}^*\in\boldsymbol{P}$, so there exists a set $N\in\boldsymbol{P}$ that satisfies $\boldsymbol{P}^*\in N$ and $\forall\boldsymbol{P}\in N$.

$$\frac{\mathrm{d}\boldsymbol{P}}{\mathrm{d}t}=-\omega\left(\nabla FO(\boldsymbol{P})+\nabla C(\boldsymbol{P})\,\boldsymbol{y}\right) \tag{C9}$$

The Jacobian linearization result of the system can be described as $\mathrm{d}\boldsymbol{P}/\mathrm{d}t=\boldsymbol{A}\boldsymbol{P}$, where $\boldsymbol{A}=-\omega\nabla_P^2L\left(\boldsymbol{J}^*\right)$. It should have one or more negative eigenvalues if $\nabla_P^2L\left(\boldsymbol{J}^*\right)$ is not positive semidefinite, so $\boldsymbol{A}$ has one or more positive eigenvalues correspondingly. Thus, the system (C9) is unstable at $\boldsymbol{J}^*$ through the Lyapunov indirect theorem, and the same is true for the RNN in (11), the following function is defined:

$$V^*(\boldsymbol{J})=-\boldsymbol{F}(\boldsymbol{J})^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{J})-\frac{1}{2\omega}\left\|\boldsymbol{R}(\boldsymbol{J})\right\|^2+\frac{1}{2\omega}\left\|\boldsymbol{J}-\boldsymbol{J}^*\right\|^2 \tag{C10}$$

where $\boldsymbol{F}(\boldsymbol{J})=\left(\nabla FO(\boldsymbol{P})+\nabla C(\boldsymbol{P})\,\boldsymbol{y},-C(\boldsymbol{P})-P_L\right)^{\mathrm{T}}$; $\boldsymbol{R}(\boldsymbol{J})=P_{\Omega}\left(\boldsymbol{J}-\omega\boldsymbol{F}(\boldsymbol{J})\right)-\boldsymbol{J}$; and $\Omega=\boldsymbol{J}\times\left(\boldsymbol{R}^{+}\right)^m$. Similar to the proof of Theorem 1 [29], the following formulas can be obtained:

$$\frac{\mathrm{d}V^*}{\mathrm{d}t}\le-\boldsymbol{F}(\boldsymbol{J})^{\mathrm{T}}\left(\boldsymbol{J}-\boldsymbol{J}^*\right)-\boldsymbol{R}(\boldsymbol{J})^{\mathrm{T}}\nabla\boldsymbol{F}(\boldsymbol{J})\,\boldsymbol{R}(\boldsymbol{J})\le\boldsymbol{0} \tag{C11}$$

$$V^*(\boldsymbol{J})\ge\frac{1}{2\omega}\left\|\boldsymbol{J}-\boldsymbol{J}^*\right\|^2 \tag{C12}$$

where $\nabla\boldsymbol{F}(\boldsymbol{J})$ can be described as:

$$\nabla\boldsymbol{F}(\boldsymbol{J})=\begin{bmatrix}\nabla^2FO(\boldsymbol{J})+\sum_{i=1}^{m}y\nabla^2C_i(\boldsymbol{J})+P_L & \nabla C(\boldsymbol{J})\\ -\nabla C(\boldsymbol{J})^{\mathrm{T}} & \boldsymbol{0}\end{bmatrix} \tag{C13}$$

$\nabla^2FO(\boldsymbol{P})+y_i\nabla^2C_i(\boldsymbol{P})$ is positive semidefinite, so $\nabla\boldsymbol{F}(\boldsymbol{J})$ is positive semidefinite on N. For $t>t_0$, $\boldsymbol{R}(\boldsymbol{J})^{\mathrm{T}}\nabla\boldsymbol{F}(\boldsymbol{J})\,\boldsymbol{R}(\boldsymbol{J})\ge0$, and $\boldsymbol{F}(\boldsymbol{J})^{\mathrm{T}}\left(\boldsymbol{J}-\boldsymbol{J}^*\right)\ge\left(\boldsymbol{F}(\boldsymbol{J})-\boldsymbol{F}\left(\boldsymbol{J}^*\right)\right)^{\mathrm{T}}\left(\boldsymbol{J}-\boldsymbol{J}^*\right)\ge0$. Because $\nabla_P^2L(\boldsymbol{J})$ is positive definite, there exists a constant $\xi>0$ leading to $\boldsymbol{h}^{\mathrm{T}}\nabla_P^2L(\boldsymbol{J})\,\boldsymbol{h}\ge\xi\left\|\boldsymbol{h}\right\|^2$, $\boldsymbol{h}\in\mathbf{R}^{n+m}$.

For any $\boldsymbol{X}^1$, $\boldsymbol{X}^2$, the mean-value theorem is applied to $\boldsymbol{F}(\boldsymbol{X})$ such that:

$$\boldsymbol{F}\left(\boldsymbol{X}^2\right)-\boldsymbol{F}\left(\boldsymbol{X}^1\right)=\int_0^1\nabla\boldsymbol{F}\left(\boldsymbol{X}^1+s\left(\boldsymbol{X}^2-\boldsymbol{X}^1\right)\right)\left(\boldsymbol{X}^2-\boldsymbol{X}^1\right)\mathrm{d}s \tag{C14}$$

where $\boldsymbol{X}^1=\boldsymbol{J}^*=\left[\left(\boldsymbol{P}_G^*\right)^{\mathrm{T}},\left(\boldsymbol{Y}_W^*\right)^{\mathrm{T}},\left(\boldsymbol{y}^*\right)^{\mathrm{T}}\right]^{\mathrm{T}}$; $\boldsymbol{X}^2=\boldsymbol{J}=\left[\boldsymbol{P}_G^{\mathrm{T}},\boldsymbol{Y}_W^{\mathrm{T}},\boldsymbol{y}^{\mathrm{T}}\right]^{\mathrm{T}}$. Then note that:

$$\begin{aligned}&\left(\boldsymbol{F}(\boldsymbol{J})-\boldsymbol{F}\left(\boldsymbol{J}^*\right)\right)^{\mathrm{T}}\left(\boldsymbol{J}-\boldsymbol{J}^*\right)=\\&\int_0^1\left(\boldsymbol{J}-\boldsymbol{J}^*\right)^{\mathrm{T}}\nabla\boldsymbol{F}\left(\boldsymbol{J}+s\left(\boldsymbol{J}-\boldsymbol{J}^*\right)\right)\left(\boldsymbol{J}-\boldsymbol{J}^*\right)\mathrm{d}s=\\&\int_0^1\Bigg[\left(\boldsymbol{P}-\boldsymbol{P}^*\right)^{T}\left(\nabla_P^2FO(\boldsymbol{P})+\sum_{i=1}^{m}y_i\nabla^2C_i(\boldsymbol{P})\right)\left(\boldsymbol{P}-\boldsymbol{P}^*\right)-\\&\left(\boldsymbol{y}-\boldsymbol{y}^*\right)^{\mathrm{T}}\nabla C\left(\hat{\boldsymbol{P}}\right)^{\mathrm{T}}\left(\boldsymbol{P}-\boldsymbol{P}^*\right)-\left(\boldsymbol{P}-\boldsymbol{P}^*\right)^{\mathrm{T}}\nabla\,C\left(\hat{\boldsymbol{P}}\right)^{\mathrm{T}}\left(\boldsymbol{y}-\boldsymbol{y}^*\right)\Bigg]\mathrm{d}s=\\&\int_0^1\left(\boldsymbol{P}-\boldsymbol{P}^*\right)^{\mathrm{T}}\left(\nabla_P^2FO\left(\hat{\boldsymbol{P}}\right)+\sum_{i=1}^{m}\hat{y}_i\nabla^2C_i\left(\hat{\boldsymbol{P}}\right)\right)\left(\boldsymbol{P}-\boldsymbol{P}^*\right)\mathrm{d}s\end{aligned} \tag{C15}$$

where $\hat{\boldsymbol{P}}=\boldsymbol{P}+s\left(\boldsymbol{P}-\boldsymbol{P}^*\right)$; and $\hat{\boldsymbol{y}}=\boldsymbol{y}+s\left(\boldsymbol{y}-\boldsymbol{y}^*\right)$. Since N is convex, $\left(\boldsymbol{F}(\boldsymbol{J})-\boldsymbol{F}\left(\boldsymbol{J}^*\right)\right)^{\mathrm{T}}\left(\boldsymbol{J}-\boldsymbol{J}^*\right)\ge\xi\left\|\boldsymbol{P}-\boldsymbol{P}^*\right\|$ and $\left(\hat{\boldsymbol{P}},\hat{\boldsymbol{y}}\right)\in N$, $\forall t\ge t_0$. Formula (C11) can be rewritten as:

$$\frac{\mathrm{d}V^*}{\mathrm{d}t}\le-\boldsymbol{F}(\boldsymbol{J})^{\mathrm{T}}\left(\boldsymbol{J}-\boldsymbol{J}^*\right) \tag{C16}$$

By integrating this formula, the following formulas are obtained:

$$V^*\left(\boldsymbol{J}(t)\right)\le V^*\left(\boldsymbol{J}\left(t_0\right)\right)-\int_{t_0}^{t}\left(\boldsymbol{J}(s)-\boldsymbol{J}^*\right)^{\mathrm{T}}\boldsymbol{F}\left(\boldsymbol{J}(s)\right)\mathrm{d}s \tag{C17}$$

$$\begin{aligned}&\left\|\boldsymbol{P}(t)-\boldsymbol{P}^*\right\|\le2\omega V^*\left(\boldsymbol{J}\left(t_0\right)\right)-2\omega\int_{t_0}^{t}\left(\boldsymbol{J}(s)-\boldsymbol{J}^*\right)^{\mathrm{T}}\boldsymbol{F}\left(\boldsymbol{J}(s)\right)\mathrm{d}s\le\\&-2\omega\int_{t_0}^{t}\left(\boldsymbol{J}(s)-\boldsymbol{J}^*\right)^{\mathrm{T}}\left(\boldsymbol{F}\left(\boldsymbol{J}(s)\right)-\boldsymbol{F}\left(\boldsymbol{J}^*\right)\right)\mathrm{d}s+2\omega V^*\left(\boldsymbol{J}\left(t_0\right)\right)\le\\&-2\eta\int_{t_0}^{\mathrm{t}}\left\|\boldsymbol{P}(s)-\boldsymbol{P}^*\right\|\mathrm{d}s+2\omega V^*\left(\boldsymbol{J}\left(t_0\right)\right)\end{aligned} \tag{C18}$$

where $\eta=\omega\xi>0$. According to the Gronwall inequality proposed in [30], we can obtain:

$$\left\|\boldsymbol{P}-\boldsymbol{P}^*\right\|\le\sqrt{2\omega V^*\left(\boldsymbol{J}\left(t_0\right)\right)}\exp\left(-\eta\left(t-t_0\right)\right)\quad\forall t\ge t_0 \tag{C19}$$

where $\boldsymbol{P}^*$ is the set of minimum. Based on the above proof, the neural network exponentially converges to $\boldsymbol{P}^*$. Thus, the output trajectory of the proposed neural network is convergent to the optimal solution of problem (10).

REFERENCES

[1] V. K. Jadoun, N. Gupta, K. R. Niazi *et al*., "Modulated particle swarm optimization for economic emission dispatch," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 80-88, Dec. 2015.

[2] V. K. Jadoun, N. Gupta, K. R. Niazi *et al*. "Nonconvex economic dispatch using particle swarm optimization with time varying operators," *Advances in Electrical Engineering*, vol. 2014, pp. 1-13, Oct. 2014.

[3] N. Sinha and B. Purkayastha, "PSO embedded evolutionary programming technique for nonconvex economic load dispatch," in *Proceedings of IEEE PES Power Systems Conference and Exposition*, New York, USA, Oct. 2004, pp. 66-71.

[4] S. Jiang, Z. Ji, and Y. Wang, "A novel gravitational acceleration enhanced particle swarm optimization algorithm for wind-thermal economic emission dispatch problem considering wind power availability," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 1035-1050, Dec. 2015.

[5] M. Younes, K. Fouad, and B. Bagdad, "A new technique for solving the multi-objective optimization problem using hybrid approach," *Frontiers in Energy*, vol. 8, no. 4, pp. 499-503, Dec. 2014.

[6] S. Kumar and R. Naresh, "Nonconvex economic load dispatch using an efficient real-coded genetic algorithm," *Applied Soft Computing*, vol. 9, no. 1, pp. 321-329, Jan. 2009.

[7] Y. Labbi, D. B. Attous, and B. Mahdad, "Artificial bee colony optimization for economic dispatch with valve point effect," *Frontiers in Energy*, vol. 8, no. 4, pp. 449-458, Dec. 2014.

[8] P. S. Kulkarni, A. G. Kothari, and D. P. Kothari, "Combined economic and emission dispatch using improved backpropagation neural network," *Electric Machines and Power Systems*, vol. 28, no. 1, pp. 31-44, Nov. 2000.

[9] S. Boudab and N. Goléa, "Combined economic-emission dispatch problem: dynamic neural networks solution approach," *Journal of Renewable and Sustainable Energy*, vol. 9, no. 3, p. 35503, Jun. 2017.

[10] S. Balakrishnana, P. S. Kannanb, and C. Aravindan, "On-line emission and economic load dispatch using adaptive Hopfield neural network," *Applied Soft Computing*, vol. 2, no. 4, pp. 297-205, Feb. 2003.

[11] C. M. Huang and Y. C. Huang, "A novel approach to real-time economic emission power dispatch," *IEEE Transactions on Power Systems*, vol. 18, vol. 1, pp.288-294, Feb. 2003.

[12] Y. Du, W. Pei, N. Chen *et al*., "Real-time microgrid economic dispatch based on model predictive control strategy," *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 5, pp. 787-796, Sept. 2017.

[13] P. Hou, J. Zhu, K. Ma *et al*., "A review of offshore wind farm layout optimization and electrical system design methods," *Journal of Mod-*

ern Power Systems and Clean Energy, vol. 7, no. 5, pp. 975-986, Sept. 2019.
[14] R. A. Fernandes and N. Rajan, "Power optimization of linear feedback shift register (LFSR) using power gating," *Power*, vol. 5, no. 5, pp. 1-8, Sept. 2018.
[15] Y. Li, J. Wang, C. Gu *et al.*, "Investment optimization of grid-scale energy storage for supporting different wind power utilization levels," *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 6, pp. 1721-1734, Nov. 2019.
[16] K. G. Boroojeni, M. H. Amini, S. S. Iyengar *et al*., "An economic dispatch algorithm for congestion management of smart power networks," *Energy Systems*, vol. 8, no. 3, pp. 643-667, Aug. 2017.
[17] S. Bahrami and V. W. S. Wong, "Security-constrained unit commitment for AC-DC grids with generation and load uncertainty," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2171-2732, May 2018.
[18] N. Amjady, A. Attarha, S. Dehghan *et al*., "Adaptive robust expansion planning for a distribution network with DERs," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1698-1715, Aug. 2017.
[19] E. Samani and F. Aminifar, "Tri-level robust investment planning of DERs in distribution networks with AC constraints," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3749-3757, Sept. 2019.
[20] X. Hu and J. Wang, "A recurrent neural network for solving nonconvex optimization problems," in *Proceedings of IEEE International Joint Conference on Neural Network*, Vancouver, Canada, Jul. 2006, pp. 4522-4528.
[21] H. Bouchekara, A. E. Chaib, and M. A. Abido, "Optimal power flow using GA with a new multi-parent crossover considering: prohibited zones, valve-point effect, multi-fuels and emission," *Electrical Engineering*, vol. 100, no. 1, pp. 151-165, Mar. 2018.
[22] Y. Elsheakh, S. Zou, Z. Ma *et al.*, "Decentralised gradient projection method for economic dispatch problem with valve point effect," *IET Generation, Transmission & Distribution*, vol. 12, no. 16, pp. 3844-3851, Sept. 2018.
[23] W. Meng and X. Wang, "Distributed energy management in smart grid with wind power and temporally coupled constraints," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 8, pp. 6052-6062, Aug. 2017.
[24] F. Guo, C. Wen, J. Mao *et al*., "Distributed economic dispatch for smart grids with random wind power," *IEEE Transactions on Smart Grid*, vol. 7 , no. 3, pp. 1572-1583, May 2016.
[25] X. Liu and W. Xu, "Minimum emission dispatch constrained by stochastic wind power availability and cost," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1705-1713, Aug. 2010.
[26] T. Wang, X. He, and T. Huang, "Collective neurodynamic optimization for economic emission dispatch problem considering valve point effect in microgrid," *Neural Networks*, vol. 93, pp. 126-136, Sept. 2017.
[27] S. A. H. Soliman and A. A. H. Mantawy, *Modern Optimization Techniques with Applications in Electric Power Systems.* New York: Springer, 2011.
[28] Y. Xia and J. Wang, "A recurrent neural network for nonlinear convex optimization subject to nonlinear inequality constraints," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 7, pp. 1385-1394, Jul. 2004.
[29] D. Kinderlehrer and G. Stampacchia, "An introduction to variational inequalities and their applications," *SIAM Review*, vol. 23, no. 4, pp. 539-543, Jan. 1981.
[30] H. K Khalil, *Nonlinear Systems.* Upper Saddle River: Prentice Hall, 2002.
[31] Y. Shi and R. C. Eberhart. "Empirical study of particle swarm optimization," in *Proceedings of the 1999 Congress on Evolutionary Computation (CEC99)*, Washington DC, USA, Aug. 2002, pp. 1945-1950.
[32] S. Chatterjee, S. Sarkar, S. Hore *et al*., "Particle swarm optimization trained neural network for structural failure prediction of multistoried RC buildings," *Neural Computing and Applications*, vol. 28, no. 8, pp. 2005-2016, Jan. 2016.
[33] O. Kulkarni, N. Kulkarni, A. J. Kulkarni *et al*., "Constrained cohort intelligence using static and dynamic penalty function approach for mechanical components design," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 33, no. 6, pp. 570-588, Nov. 2016.
[34] S. Katayama, Y. Satoh, and M. Doi, "Nonlinear model predictive control for systems with state-dependent switches and state jumps using a penalty function method," in *Proceedings of 2018 IEEE Conference on Control Technology and Applications (CCTA)*, Copenhagen, Denmark, Aug. 2018, pp. 312-317.
[35] J. Zhou, P. E. D. Love, and K. L. Teo, "An exact penalty function method for optimising QAP formulation in facility layout problem," *International Journal of Production Research*, vol. 55, no. 10, pp. 2913-2929, Sept. 2016.
[36] J. B. Clempner and A. S. Poznyak, "A Tikhonov regularized penalty function approach for solving polylinear programming problems," *Journal of Computational and Applied Mathematics*, vol. 328, pp. 267-286, Jan. 2018.
[37] Z. Meng, C. Dang, and M. Jiang, "Exactness and algorithm of an objective penalty function," *Journal of Global Optimization*, vol. 56, no. 2, pp. 691-711, Apr. 2012.
[38] Y. Bai and K. Wang, "An exact penalty function of objective parameter," *Numerical Methods and Computer Applications*, vol. 35, no. 1, pp. 35-45, Mar. 2014.
[39] K. L. Du and M. N. S. Swamy, "Particle swarm optimization," in *Search and Optimization by Metaheuristics*, Birkhäuser, Cham: Springer, Jul. 2016, pp. 153-173.
[40] Z. Yan, J. Wang, and G. Li, "A collective neurodynamic optimization approach to bound-constrained nonconvex optimization," *Neural networks*, vol. 55, pp. 20-29, Jul. 2014.
[41] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proceedings of 1998 IEEE International Conference on Evolutionary Computation*, Anchorage, USA, May 1998, pp. 69-73.

Jiayu Wang is currently working towards the B.S. degree in communication engineering from Southwest University, Chongqing, China. Her current research interests include neural networks and smart grid.

Xing He received the B.S. degree in mathematics and applied mathematics from the Department of Mathematics, Guizhou University, Guiyang, China, in 2009, and the Ph. D. degree in computer science and technology from Chongqing University, Chongqing, China, in 2013. Since 2018, he has been a Professor at the School of Electronics and Information Engineering, Southwest University, Chongqing, China. From November 2012 to October 2013, he was a Research Assistant with the Texas A&M University at Qatar, Doha, Qatar. His research interests include neural networks, bifurcation theory, optimization method, smart grid, and nonlinear dynamical system.

Junjian Huang received the B.S. degree from Chongqing Communication Institute, Chongqing, China, in 2002, and the Ph.D. degree in computer science and technology at Chongqing University, Chongqing, China, in 2014. He has been a Professor at Chongqing University of Education, Chongqing, China, since 2014. His current research interests include neural networks, memristive systems, intermittent control and synchronization.

Guo Chen received the Ph.D. degree in electrical engineering from the University of Queensland, Brisbane, Australia, in 2010. He is currently an Australian Research Council Discovery Early Career Researcher Award (ARC DECRA) Fellow with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia. He has held academic positions with the Australian National University, Canberra, Australia, the University of Sydney, Sydney, Australia, and the University of Newcastle, Callaghan, Australia. His current research interests include power system security assessment, optimization and control, complex network, intelligent algorithms and their applications in smart grid.