IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 22, 2020, accepted December 7, 2020, date of publication December 18, 2020, date of current version December 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3045754

Application Conditions of Bounded Rationality and a Microgrid Energy Management Control Strategy Combining Real-Time Power Price and Demand-Side Response

NAN WU[1,2], HONGLEI WANG[1], LINFEI YIN[3], XUFENG YUAN[4], (Member, IEEE), AND XIAOXIA LENG[1]

[1]School of Management, Guizhou University, Guiyang 550025, China
[2]School of Mechanical and Electrical Engineering, Guizhou Normal University, Guiyang 550025, China
[3]College of Electrical Engineering, Guangxi University, Nanning 530004, China
[4]College of Electrical Engineering, Guizhou University, Guiyang 550025, China

Corresponding author: Honglei Wang (gzdxhlwang@163.com)

This work was supported in part by the Key Science and Technology Projects of China Southern Power Grid Company Ltd., under Grant GZKJQQ00000417, in part by the Science and Technology Plan Project of Guizhou Province under Grant [2016]5103, and in part by the Doctoral Research Project of Guizhou Normal University under Grant GZNUD[2019]21.

ABSTRACT Microgrid energy management is a typical discrete non-linear optimization problem that is usually solved by off-line optimization, day-ahead demand-side management, and long-term dynamic optimization scheduling strategy. However, due to the intermittent distributed generation and time-varying load in microgrids, more attention should be paid to the real-time optimal scheduling of the overall operation of energy to ensure the dynamic balance of supply and demand in microgrids. Combining demand-side response with real-time power price, this paper applies the strategy to microgrid energy management and proposes a distributed energy real-time management model of microgrid based on demand-side response function. A deep adaptive dynamic programming optimization algorithm is also proposed for the model. The real-time interaction between microgrid operators and users is realized. The closed-loop feedback control structure of the proposed model ensures the real-time optimization control strategy. Therefore, the proposed energy management model and control strategy can realize intra-day dispatching in microgrids. The real-time performance and effectiveness of the proposed energy management model and control strategy are also verified by numerical simulation. Finally, since the proposed model is approximate, whether the solution obtained by the algorithm is the optimal or satisfactory solution of the optimization strategy set is a lack of theoretical support. Therefore, according to the approximation theorem of bounded rationality, the application conditions of the model in power markets are proposed. It is proved that the proposed model meets the application conditions, and is a specific application of bounded rationality approaching complete rationality in the power market. It is also proved that the best solution is involved in the satisfactory solution set of the model. Thus, the control strategy is a rational and feasible optimal management control strategy, which provides a theoretical basis for its further implementation.

INDEX TERMS Microgrid, real-time power price, demand-side response, deep adaptive dynamic programming optimization algorithm, energy management, bounded rationality.

I. INTRODUCTION

Power system plays a major role in fossil energy consumption and is an important source of air pollution. To lower air pollution caused by thermal power generation, the power system is investing in distributed clean energy such as solar and wind power so that use of traditional fossil energy can be reduced effectively [1]. Establishing a regional microgrid is the best way to manage distributed energy. The microgrid can facilitate the local consumption of generated

The associate editor coordinating the review of this manuscript and approving it for publication was Bin Zhou.

power, effectively save losses due to long-distance transmission and reduce thermal power generation, contributing to lowering carbon emissions. Energy management strategies (EMS) for microgrids aim to optimize the operational performance of microgrids. For example, under the premise of stable operation and making full use of renewable energy, long-term operational costs minimum can be realized by optimizing dispatch of electric energy in the grid.

Since microgrid consists of various types of distributed power supplies and energy storage systems, generally, the typical discrete nonlinear optimization method is used to establish the energy management model [2]–[6]. As an optimization problem, the aim of energy managing for the microgrid is to realize local energy consumption, the balance between supply and demand, stable operation and optimal dispatch through establishing objective functions and constraints. For energy managing of microgrid, the method of energy managing for distributions networks are usually employed. Based on predicting the future supply and demand in the microgrid, the balance between supply and demand in the microgrid can be reached through the day-ahead EMS [7]–[10]. Demand-side management or combining it with dynamic prices is normally used to obtain the energy management and control strategies (EMCS) for microgrids, for which the offline or static optimization algorithms adopted mainly include mixed-integer quadratic optimization, sequence planning or sequential quadratic programming. Some research, based on dynamic prices under demand-side management, treated the dispatch optimizing for microgrid as a hybrid integer quadratic programming problem and used costs of generation and dynamic pricing for achieving optimized dispatch in microgrid [11]. In [12], energy managing for the microgrid is based on the generation-side and demand-side management and is fulfilled through representing the energy managing and dispatching problem as a linear hybrid integer programming case with multi-constraints and multi-objectives. In [13], through sequential optimization of solutions obtained by discrete algorithms, real-time EMS under a certain time scale is achieved for the microgrid. Under the background of the potential demand of day-ahead dispatching and combining demand-side management with real-time power price (RTPP), RTPP is adopted to optimize the power demand of users, to achieve peak shaving and valley filling, supply-demand balance, and stable operation of the power grid. This is a hot research direction of energy optimization management strategy at present [14]–[17].

Predicting supply and demand in advance is required for day-ahead dispatching. Changing supply and demand caused by uncertainties of power generation and distribution (PGD) in microgrids can make day-ahead dispatching difficult. Therefore, dynamic EMS has been proposed, using time dividing [18], [19] or on-line optimization to solve dynamic dispatching problems caused by changing supply and demand in microgrid [20]–[23]. In [24], energy storage systems are managed using approximate dynamic programming, enabling energy managing for an isolated microgrid. To reduce the calculation time required by the optimization, the literature [25] uses dynamic programming to perform energy management for the microgrid. Improved results are obtained but the optimization is still not real-time. In [26], energy managing for the microgrid is achieved by creating a rule-based aggregate energy managing model. In [27], an on-line energy managing system is proposed through switching between charging and discharging modes of batteries to reach the optimized operation of the microgrid. The above methods can solve the optimization problem of managing long-term costs for microgrid influenced by renewable energy resources subject to environmental changes, but because of the long calculation time and optimization objectives not adjusted, applying these methods to practical control strategies can have limitations.

The abovementioned research has made significant contributions to energy management for the microgrid. However, for a microgrid with intermittent distributed generation (DG) resources and changing loads, day-ahead dispatching or non-real-time dispatching strategies can influence the balance between supply and demand and the backup plan of supplying power if there is a significant difference between the prediction and the actual amount of supply and demand. Therefore, to reach a real-time balance between supply and demand, peak shaving and valley filling, and long-term stable economic operating, it is essential to implement real-time EMCS for the microgrid. Thus, this paper proposed a microgrid energy management and control model based on demand-side response (DSR) under RTPP, and also proposed an adaptive dynamic programming optimization algorithm based on deep learning, which realizes intra-day scheduling and dynamic supply-demand balance of microgrid energy. The potential of a microgrid for optimal dispatch between distributed energy and load is fully exploited. Since the proposed control model is approximate, to ensure that its control strategy is optimal, the application conditions of the strategy in the power market are proposed according to the approximation theorem of bounded rationality, which proves that the control strategy is rational and feasible optimal energy management control strategy.

The paper is organized as follows. In Section II, the distributed generation cost model and RTPP model are given, and the microgrid cost objective optimization function is proposed; in Section III, the RTPP microgrid energy management and control strategy is based on the deep adaptive dynamic programming optimization algorithm is proposed. In Section IV, the effectiveness and feasibility of the proposed model and algorithm are verified by numerical simulation, and the analysis conclusions are given. In Section V, the application conditions of bounded rationality in the power market are proposed, which proves that the proposed control model meets the application requirements. Summaries and conclusions are given in Section VI.

II. MICROGRID SYSTEM MODEL

The power supply in a microgrid is a variety of distributed power sources, whose power supply cost is different from each other. Microgrid operators usually make the lowest cost of power generation and supply utilizing optimal scheduling, on the premise of ensuring the quality of power supply and making full use of renewable energy. Therefore, the cost model of DG in a microgrid is established firstly, then the cost model of the power storage device and the RTPP model are given. Finally, the optimal scheduling cost model of the microgrid is proposed.

A. INTRODUCTION TO THE MICROGRID SYSTEM

There are various kinds of DG in the microgrid, such as wind, solar, gas, hydro, etc. The DG resources are represented as $R = \{r_1, r_2, r_3 \ldots r_G\}$. Since solar power generation and wind power generation are intermittent, it is usually necessary to install energy storage devices in the microgrid to store and convert this part of electric energy, and then supply it to the users in the grid to ensure the continuous and stable power supply, which can be expressed as $D = \{\mathrm{d}_1, d_2, d_3 \ldots d_S\}$. The users in a microgrid can be divided into two types: one is the non-flexible load whose load demand must meet, and the other is the flexible load that the microgrid operators can cut / parallel power supply flexibly according to the optimal scheduling of the system. In the microgrid, operators dispatch the distributed power supply, power storage device and flexible load in the microgrid operation and energy management center (MGOEMC) through the two-way traffic network in real-time. The framework is shown in FIGURE 1.

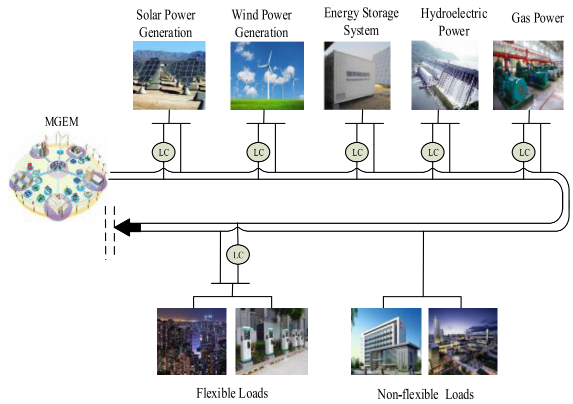


FIGURE 1. Structure of the microgrid system.

B. MODEL OF DG

In the microgrid, distributed power generation is mainly divided into two categories: one is intermittent power supply represented by wind power generation and solar power generation, whose power generation will change intermittently due to the impact of climate. Therefore, the power generation of this kind of power source can't be controlled and dispatched artificially; the other kind is the distributed power source which can be controlled and dispatched by a human, which is mainly composed of hydroelectric power generation and gas power generation.

1) INTERMITTENT DISTRIBUTED ENERGY

For the intermittent distributed power generation mainly composed of wind power generation and solar power generation, since the power generation energy comes from the natural environment and does not need the human energy conversion before power generation, it can be approximately considered that they only have the operation and maintenance cost and the power generation cost can be approximately zero. To provide a stable power supply for users in the microgrid, it is usually necessary to store the power generated by intermittent power supplies into energy storage devices, and then the energy storage devices will supply stable power to users.

2) MODEL OF TRADITIONAL ENERGY

The schedulable power supplies in the microgrid are composed of hydroelectric power and gas power, which can be adjusted according to the change of power demand. The hydropower is expressed as $h \in R_h$ whose operating constraints are: $\forall t$

$$0 \le p_h(t) \le p_{h\max} \tag{1}$$
$$|p_h(t) - p_h(t-1)| \le p_{h\max} \tag{2}$$

where $p_{h\max}$ is the maximum output of hydropower. The generation cost model at the time t is [28]:

$$C_h(p_h(t)) = a_h p_h(t)^2 + \beta_h p_h(t) + \kappa_h \tag{3}$$

where a_h, β_h, κ_h are constants. For a gas distributed power supply, its generating unit can be expressed as $b \in R_b$, and its constraints and generation cost model at the time t are: $\forall t$

$$0 \le p_b(t) \le p_{b\max} \tag{4}$$
$$|p_b(t) - p_b(t-1)| \le p_{b\max} \tag{5}$$
$$C_b(p_b(t)) = a_b p_b(t) \tag{6}$$

where $p_{\mathrm{bmax}}(t)$ is the maximum output of gas power and a_b is a constant.

C. MODEL OF ENERGY STORAGE SYSTEM

In this paper, the basic unit of the power storage device of the microgrid is the battery, and all the storage units constitute the battery energy storage system (BESS). The microgrid stores and converts the generation of intermittent distributed power through BESS. For a storage battery in a microgrid $d \in D$, if the intermittent power supply charges the BESS at the time t, the active power $p_d(t)$ is positive; on the contrary, if the BESS is discharged, the active power $p_d(t)$ is negative. At this time, the intermittent distributed power supply and battery supply power to the microgrid load together. $E_d(t)$ is the capacity of the storage battery at time t, and its operational constraints are as follows: $\forall t$

$$p_{f\max} \le p_d(t) \le p_{c\max} \tag{7}$$
$$E_d(t+1) = E_d(t) + p_d(t)\Delta t \tag{8}$$

$$E_{d\min} \leq E_d(t) \leq E_{d\max} \tag{9}$$

where $-p_{f\max}$ is the maximum allowable discharging rate, $p_{c\max}$ is the maximum allowable charging rate, $E_{d\min}$ is the minimum capacity that the battery needs to retain, and $E_{d\max}$ is the maximum capacity that the battery can store.

Since the life of a battery will be affected by fast charging and releasing, this behavior of the battery must be controlled. This function is realized by establishing a cost function model, that is, for a given storage battery d at the time t, we have: [29]:

$$C_d(p_d(t)) = a_d p_d(t)^2 + \kappa_d \tag{10}$$

where a_d is a fast charging and discharging penalty factor and κ_d is a basic cost factor. Both are positive constants. The cost function $C_d(p_d(t))$ is convex and is non-increasing in $[P_{d\max}(t), 0]$ and non-decreasing in $[0, P_{c\max}(t)]$. Moreover, the capacity of the BESS is set to meet fully the storing and scheduling requirements of intermittent DG resources.

D. MODEL OF RTPP

The RTPP is an important measure to adjust the power demand of the flexible loads when the DG in the microgrid cannot accommodate the flexible loads largely. The relationship between the RTPP and the DSR based power prices in the microgrid is:

$$\rho_r(t) = e^{\frac{Kp_m(t)}{P_{l\max}(t)-P_{l\min}(t)}} \cdot \rho_l \tag{11}$$

where $p_{l\min}(t)$ is the non-flexible loads that must be supplied, and $p_{l\max}(t)$ is the maximum power demand in the microgrid. $\rho_r(t)$ is the RTPP for the flexible loads when the DG resources in the microgrid are insufficient. $p_m(t)$ is the amount of thermal power that the microgrid operator buys from distribution networks. K is a constant ratio. $p_{l\max}(t) - p_{l\min}(t)$ is the power demand of flexible loads. ρ_l is the power price of the distribution networks. Equation (11) is a function of the DSR based real-time power prices in the microgrid, where it can be seen that when the power demand of the flexible loads is much higher than the power output of the DG in the microgrid, the RTPP will increase quickly with the increasing power demand, thereby achieving the fast and real-time energy management for the microgrid. Because power services are public products, power prices cannot be completely determined by the free market. Moreover, when the price of a product increases to higher than 3 – 5 times its normal prices, there will be several consumers who will rather not buy or look for alternatives to the product, which is also a kind of consumer rationality. Therefore, the following constraints of the RTPP are introduced to account for the principle:

$$0 < \rho_r(t) \leq \xi \cdot \rho_l \tag{12}$$

where ζ is the rational factor. On the one hand, the setting of the rational factor reflects that, as a social resource, the price of electric power consumption can't be completely liberalized and has certain policy constraints. On the other hand, it can also well reflect the commodity price consumption psychology of flexible load users. For determining the amount of power to buy $p_m(t)$, given that the distributed energy is fully used, two relevant conditions are to be met:

(1). When $p_{l\min} \geq p_d(t) + p_h(t) + p_b(t)$, i.e., the DG cannot supply enough power to the non-flexible loads in the microgrid, the amount of power to buy from the distribution networks can be divided into two parts. One part is for supplying the non-flexible loads and the RTPP is not applied for this part, which means supplying at cost. The other part is, on top of the amount of power bought for the non-flexible loads, the extra amount of power bought for supplying the flexible loads, and this extra amount of power is charged with the RTPP. Therefore, the range $p_m(t)$ is

$$p_m \in [0, p_{l\max}(t) - p_{l\min}(t)] \tag{13}$$

(2). When $p_{l\min} \leq p_d(t) + p_h(t) + p_b(t)$, i.e., the DG can not only supply the non-flexible loads in the microgrid but also get some leftover power to supply the flexible loads, but the leftover power is not enough for the flexible loads in the microgrid. To supply the flexible loads not covered by the DG, some power needs to be bought from the distribution networks and this part of power will be charged with the RTPP. Therefore, the range of the amount of power to buy $p_m(t)$ is:

$$p_m \in [0, p_{l\max}(t) - (p_d(t) + p_h(t) + p_b(t))] \tag{14}$$

E. OBJECTIVE OPTIMIZATION FUNCTION

The goal of microgrid operation is to achieve the lowest operating cost in the whole long-term power supply process, which includes the cost function and profit function of the following parts:

(1). Costs of DG and energy storage:

$$F_1(p_h(t), p_b(t), p_d(t)) = (\sum_{h\in G_h} C_h(p_h(t)) + \sum_{b\in G_b} C_b(p_b(t)) + \sum_{b\in D} C_d(p_d(t))) \tag{15}$$

The microgrid operator is supposed to lower the cost of PGD as much as possible, i.e.,

$$\min F_1(p_h(t), p_b(t), p_d(t)) \tag{16}$$

where the constraints are (1), (2), (4), (5), (7)-(9).

(2). Costs of buying power from the distribution networks

$$F_2(p_m(t)) = \rho_l \cdot p_m(t) \tag{17}$$

To reduce the cost of PGD, the amount of power to buy is the smaller the better, i.e.,

$$\min F_2(p_m(t)) \tag{18}$$

where the constraints are (13)-(14).

(3). Profits of RTPP

$$F_3(p_m(t)) = (\rho_r(p_m(t)) - \rho_l) \cdot p_m(t) \tag{19}$$

The RTPP in the microgrid is higher than the pricing of the distribution networks. Therefore, the higher the RTPP, the more the profits gained by the microgrid operator, and furthermore buying more power from the distributions networks means more profits, i.e.,

$$\max F_3(p_m(t)) \tag{20}$$

where the constraint is (12).

According to the cost function and the profit function above, to minimize the overall generation and supply cost of microgrid, its objective function can be written as

$$\min(F_1(p_h(t), p_b(t), p_d(t)), F_2(p_m(t)), F_3(P_m(t))) \tag{21}$$

For a convenient solution, weighting is applied to transform the multi-objective function above to a single objective function. The obtained objective optimization function for the minimum cost of PGD in the microgrid is:

$$\begin{aligned}\min(\phi_h \sum_{h\in G_h} C_h(p_h(t))+\phi_b \sum_{b\in G_b} C_b(p_b(t))+\phi_d \sum_{d\in D} C_d(p_d(t)) \\ + \phi_m(\rho_r(t) - \rho_l) \cdot p_m(t))\end{aligned} \tag{22}$$

where the following constraints are satisfied:

$$st \begin{cases} 0 \le p_h(t) \le p_{h\max},\ |p_h(t) - p_h(t-1)| \le p_{h\max} \\ 0 \le p_b(t) \le p_{b\max},\ |p_b(t) - p_b(t-1)| \le p_{b\max} \\ p_{f\max} \le p_d(t) \le p_{c\max},\ E_{d\min} \le E_d(t) \le E_{d\max} \\ E_d(t+1) = E_d(t) + p_d(t)\Delta t \\ 0 < \rho_r(t) \le \xi \cdot \rho_l \\ p_m(t) \in [0,\ p_{l\max}(t) - p_{l\min}(t)]\ or \\ p_m(t) \in [0,\ p_{l\max}(t) - (p_d(t) + p_h(t) + p_b(t))] \end{cases} \tag{23}$$

The establishment of the objective function can realize the real-time interaction between the microgrid operators and the users through the price response function on the demand side, which enables the flexible load users to make reasonable power consumption plans or strategies according to their own needs. When the capacity of the distributed generation in the microgrid fails to meet the total power consumption, the control strategy can use RTPP to keep the supply-demand balance in the microgrid. This can not only effectively reduce the thermal power purchased by the microgrid from the distribution network, cutting power generation and operation cost of the microgrid, but also make the microgrid operate in a safe and stable environment.

Besides, using the rational consumption behavior of most of the flexible loads, the RTPP-based energy management and control strategy will not completely restrain the power demand of the flexible loads in the network. Some flexible loads which are indeed in a need of power and are not sensitive to the power prices can continue to use power. The amount of power these flexible loads can use is not limited, but the costs will be increased significantly.

However, concerning the total amount of consumed power, the share of thermal power will be reduced effectively, lowering the demand for thermal power and the times the microgrid connects to the power grid while exploiting as many DG resources as possible. To ensure the real-time performance and stability of the microgrid RTPP energy management control strategy, a deep adaptive dynamic programming optimization algorithm which combines the optimization algorithm with the control algorithm is proposed by using the method of system control theory, and the algorithm is applied to solve the multi-objective optimization problem of equation (22), to obtain optimal control strategy for real-time energy management of dual variable microgrid based on the RTPP and the flexible load.

III. DEEP ADAPTIVE DYNAMIC PROGRAMMING OPTIMIZATION ALGORITHM BASED EMCS

A. PRINCIPLES OF DYNAMIC PROGRAMMING AND ADAPTIVE DYNAMIC PROGRAMMING

Dynamic systems are ubiquitous in nature. An important branch of the dynamic theory is optimal control. Optimal control has been widely applied to fields such as system engineering, economy, management, decision making, etc. In 1957, Bellman proposed an effective tool to solve optimal control problems: dynamic programming [30]. The basic idea of this method is Bellman's principle of optimality which says that if there is a nonlinear system whose dynamic equation is:

$$x(k+1) = F(x(k), u(k), k) \tag{24}$$

where $x(k)$ is the state of the system and the initial state $x(k) = x_k$ is given. $u(k)$ is the control input of the system and F is the utility function of the system. The function of the system performance index can be defined as:

$$J(x(i), i) = \sum_{k=1}^{\infty} \lambda^{k-i} F(x(k), u(k), k) \tag{25}$$

The target of control is to solve the sequence of admission control (or decision making) $u(k)$, $k = 1, 2, 3, \ldots$, getting the lowest cost function (25). According to Bellman's principle, for time k the lowest cost of any state includes two parts. One is the lowest cost at time k, and the other is a sum of the lowest costs from time $k+1$ to infinite time, i.e.:

$$J^*(x(k)) = \min_{u(k)}\{F(x(k), u(k), k) + \lambda J^*(x(k+1))\} \tag{26}$$

At this point, the corresponding control strategy $u(k)$ at time k can also be optimized, i.e.:

$$u^*(k) = \arg\min_{u(k)}\{F(x(k), u(k), k) + \lambda J^*(x(k+1))\} \tag{27}$$

Therefore, the dynamic programming method is a powerful tool for solving optimal control problems [31]. However, it is difficult to apply dynamic programming to practical applications directly. The reason is that optimal control needs to work on a system following time evolution and needs to give the optimal control index of the control sequence. Nevertheless, the function of the overall performance index is completely unknown before the sequence is completed, i.e., the problem of "Curse of dimensionality". Webers in 1977 first proposed

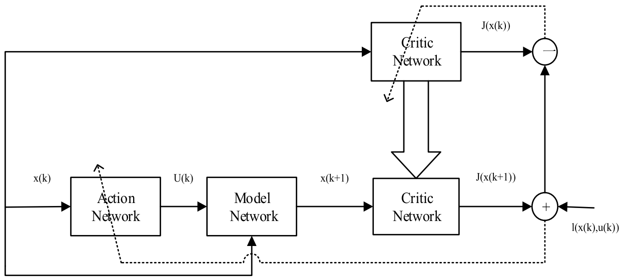


FIGURE 2. Structure of adaptive dynamic programming.

adaptive dynamic programming (ADP) to solve the problem. The ADP method is essentially using the structure of function approximation to fit cost functions and control strategies of dynamic programming, deriving solutions for optimal control of nonlinear systems [32]. A typical ADP structure is shown in FIGURE 2.

In FIGURE 2, the functions of the performance index can be expressed as:

$$J(x(k)) = l(x(k), u(x(k))) + J(x(k+1)) \tag{28}$$

where $u(x(k))$ is the variable of feedback control and the functions of the performance index $J(x(k))$ and $J(x(k+1))$ are the output of the critic network. If the weight of the critic network is w, the right side of (28) can be written as:

$$d(x(k), w) = l(x(k), u(x(k))) + J(x(k+1), w) \tag{29}$$

The left side of (28) can be written as $J(x(k), w)$. By changing the weight w of the critic network the mean square error function can be minimized to obtain the function of the optimal performance index:

$$w^* = \arg\min_{w} \left\{ |J(x(k), w) - d(x(k), w)|^2 \right\} \tag{30}$$

According to the principle of optimality, optimal control is required to meet the necessary condition of the first-order differential:

$$\begin{aligned}\frac{\partial J^*(x(k))}{\partial u(k)} &= \frac{\partial l(x(k), u(k))}{\partial u(k)} + \frac{\partial J^*(x(k+1))}{\partial u(k)} \\ &= \frac{\partial l(x(k), u(k))}{\partial u(k)} + \frac{\partial J^*(x(k+1))}{\partial u(k+1)} \frac{\partial f(x(k), u(k))}{\partial u(k)}\end{aligned} \tag{31}$$

Therefore, optimal control is derived as:

$$u^* = \arg\min_{u} \left(\left| \begin{array}{l} \dfrac{\partial J(x(k))}{\partial u(k)} - \dfrac{\partial l(x(k), u(k))}{\partial u(k)} - \\ \dfrac{\partial J^*(x(k+1))}{\partial x(k+1)} \dfrac{\partial f(x(k), u(k))}{\partial u(k)} \end{array} \right| \right) \tag{32}$$

In recent years, the theory of ADP has been evolving continuously. Online ADP methods have been proposed to solve optimal control problems of nonlinear analog systems, and optimal stabilization and optimal tracking problems of nonlinear discrete systems [33], [34]. Therefore, ADP has become an extremely important method for solving both scientific and engineering problems of modern complex systems.

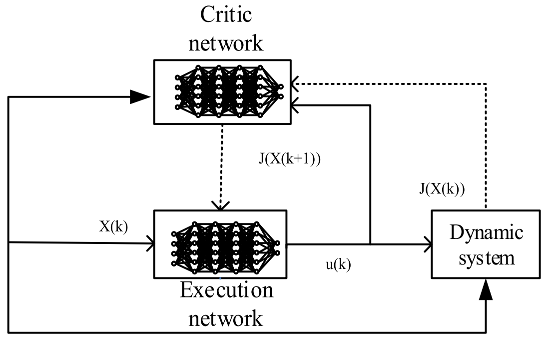


FIGURE 3. Model framework of deep adaptive dynamic programming optimization algorithm.

B. DEEP ADAPTIVE DYNAMIC PROGRAMMING OPTIMIZATION ALGORITHM BASED REAL-TIME EMCS

Based on the basic framework of adaptive dynamic programming, this paper proposes a deep adaptive dynamic programming optimization algorithm to solve the microgrid energy management problem under the RTPP by combining the deep learning neural network with the ADP simplified framework. Due to the real-time optimal control characteristics of the model of the algorithm, its basic structure only includes the critic network and execution network, which can reduce the dependence of the controller on the system model. At the same time, since the algorithm architecture has the characteristics of closed-loop feedback control system in the control theory, it has the function of real-time correction and ensures the real-time and accuracy of the control strategy. Its basic structure is shown in FIGURE 3.

In the framework of the deep adaptive dynamic programming optimization algorithm, the inputs of the execution network are: ① $X(k)$ which is the real-time power generated by each distributed power source in the microgrid measured by the power sensor, i.e., $p_h(t)$; $p_b(t)$; $p_d(t)$, ② the total load demand, and ③ non-flexible demand in the grid at the same time. The output u(k) of the execution network is the control strategy of the system, i.e., the RTPP and the corresponding purchase amount which are required by the microgrid operators trying to meet the objective function of the optimal operation. The objective of the RTPP based EMCS is scheduling power services in the microgrid through the RTPP. Using the proposed strategy, the microgrid operator can keep the balance between supply and demand by adjusting the prices.

For the critic network, its input consists of the state of the systemic controlled objects and the output of the execution network, i.e., the control strategy. Its output is a cost function for adjusting the control strategy of the executive network. The objective function to be studied in this paper is to make the operation cost of microgrid generation and power supply lowest under the RTPP. To make the objective function have

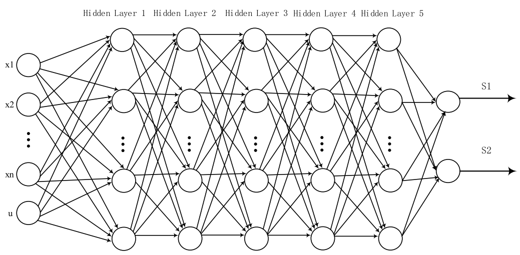


FIGURE 4. **Topology of deep learning multilayer neural networks.**

the optimal real-time approximation value and obtain the optimal real-time control strategy, it is necessary to control the output of the execution network with the minimum error. To achieve the objectives above and meet the rapidity of system optimization, the execution network and critic network in FIGURE 3 both use the neural network framework of multiple hidden layers, multiple inputs and dual outputs based on deep learning, as shown in FIGURE 4.

Since the generation capacity of DG in the microgrid changes intermittently with the climate change, and the load demand in the microgrid is also time-varying, the microgrid operators need to optimize the grid energy scheduling in real-time. To realize the multi-objective optimization of the system under the RTPP energy management control strategy, and based on the model structure above, the implementation process of the microgrid energy management control optimization strategy is shown in FIGURE 5. The specific implementation steps are as follows:

(1). Set parameters for the structure of the system, including the relevant parameters of the executive network and the critic network, as well as the learning rate, the number of input layers, the number of hidden layers, the number of output layers and the maximum number of iterations of the neural network;

(2). The real-time generation capacity of DG and the maximum demand in the network are collected by the power sensor and supplied to the execution network. At the same time, the parameters needed for the controlled state output are set and the control objectives of the system are given;

(3). Initializing the execution network and critic network;

(4). According to equation (15), calculate the RTPP and the corresponding purchased electricity, and then calculate the objective function to see if it has reached the minimum. If not, train the network in the next step. If it is the minimum, keep the current value;

(5). Input the data into the execution network, train the execution network, update its weights, and finally output the calculated RTPP and purchase electricity, which are the control strategies of the system;

(6). Input the collected data and the control strategy obtained from the execution network to the critic network for training, update its weight, and then get the corresponding cost function;

(7). Maintain the current control strategy, and calculate the modified objective function;

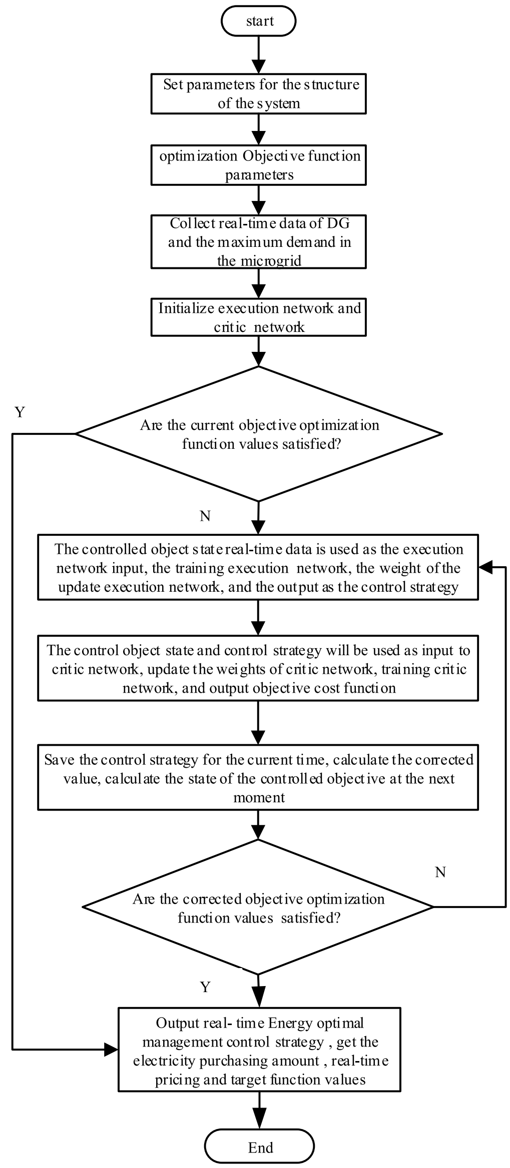


FIGURE 5. **Flowchart of microgrid energy management strategy based on deep adaptive dynamic programming optimization algorithm.**

(8). Repeat steps (4) - (7) during the observation time until the end of the optimization process, and obtain the corresponding optimization results.

After each training of the execution network, an RTPP $\rho_r(t)$ and its corresponding amount of purchasing power $p_m(t)$ can be obtained. If the corresponding constraints are met and the objective function is optimized, the training of the system will stop. If not, the execution network and critic network will continue to be trained until the optimal RTPP and real-time energy optimal management control strategy corresponding to the purchased electricity are found.

IV. SIMULATION AND ANALYSIS

To analyze the effectiveness and feasibility of the proposed control strategy, and according to the standards of the China Southern Power Grid Co. Ltd, a standard 15-minutes sam-

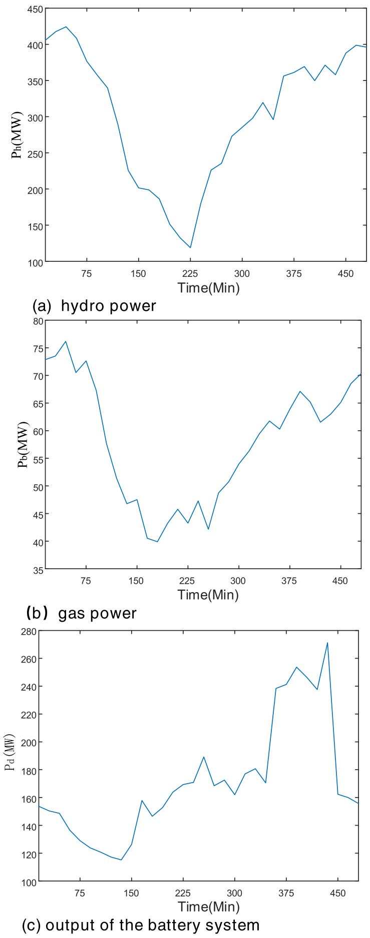


FIGURE 6. **Amount of power from hydro, gas and storage equipment.**

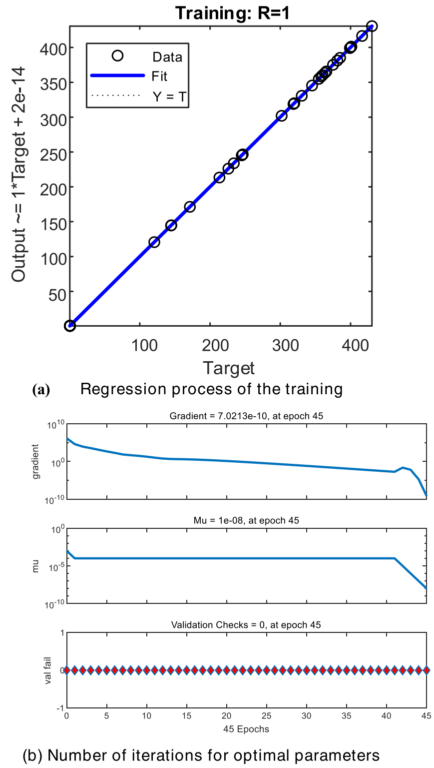


FIGURE 7. **Training process of deep learning.**

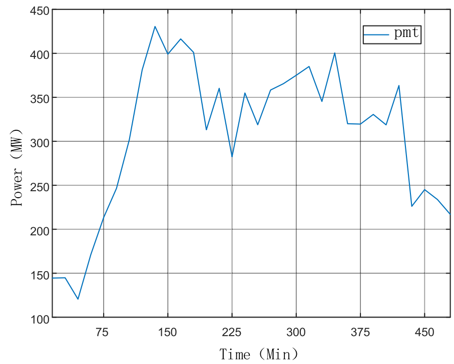
pling of the DG was implemented in a microgrid. The total sampling time is 480 minutes, and the proposed control strategy is verified by simulation. The power output of the DG in the microgrid including wind, solar, hydro, and gas resources is plotted against time in FIGURE 6 where the horizontal axis is the time of data acquisition and the vertical axis is the amount of power from the DG. The parameters in the model of hydropower and gas power generation are: $\alpha_h = 5$, $\beta_h = 12$, $\kappa_h = 0$, the cost coefficient of gas power $\alpha_b = 0.2$. The rational factor of RTPP $\xi = 3.5$, the proportional coefficient of the RTPP $k = 2.5$, the power price of distribution network $\rho = 0.4$ CNY/KW.H, and the coefficients in the minimum target function of generation and supply cost of the microgrid is set as $\phi_h = \phi_b = \phi_d = \phi_m = 1$. The parameters in the storage battery model are set as: $\alpha_d = 1$, $\kappa_d = 0$, $E_{d\max}(t) = 1600$ MW, $E_{d\min}(t) = 350$ MW, original battery capacity $E_d(0) = 800$ MW.

The examples in this paper were all calculated using a PC with a CPU of 3.4 GHz, RAM of 8 GB, an operating system of Win10, and a MATLAB release of R2018b V9.5.0.

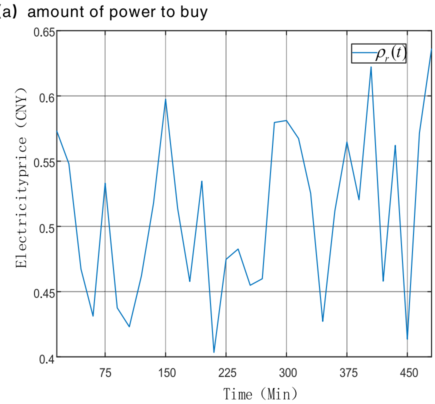
The generation capacity of hydropower, gas power, and intermittent distributed energy as well as the established models and parameters are brought into the algorithm proposed in this paper. In the deep learning neural network, three hidden layers are used, and each layer has 40 neurons to carry out optimized iterative training for the execution network and critic network. The training process and relevant data are shown in FIGURE 7. It can be seen from FIGURE 7 that the output after training closely approaches the target value. After 45 iterations, the best approximation means square deviation value is 5.1606×10^{-23}, and the whole process takes 41 seconds. It can be seen that the real-time and accuracy of the proposed energy management control strategy can meet the real-time scheduling requirements of the microgrid.

By running the deep adaptive dynamic programming optimal control algorithm proposed in the frame of FIGURE 3, the optimized RTPP, and the corresponding amount of power to buy can be obtained. FIGURE 8 shows that under the RTPP based EMCS the real-time amount of power to buy considering the total power demand in the microgrid, and the corresponding RTPP subject to the current price of power. FIGURE 8 suggests that the supply and demand for power can be balanced using the RTPP when the DG resources in the microgrid cannot provide enough power. FIGURE 8 also indicates that when the DG resources are insufficient, it can take costs higher than the normal power price of the distribution networks for the flexible loads to consume more power, which helps implement the rational adjustment of energy management for the microgrid.

FIGURE 9 shows the cost of the minimum target function of RTPP corresponding to the purchasing power after the control strategy is implemented in the system.



(a) amount of power to buy



(b) RTPP balancing supply and demand

FIGURE 8. RTPP of the microgrid and the corresponding amount of power to buy.

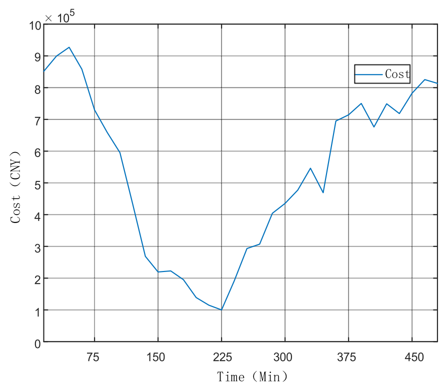
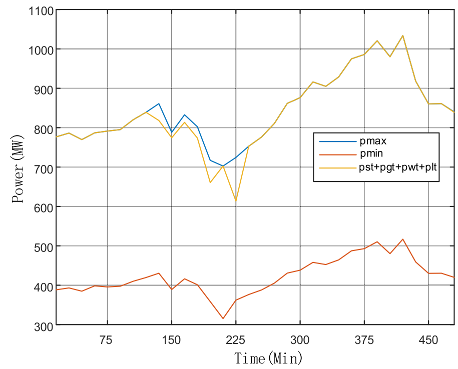
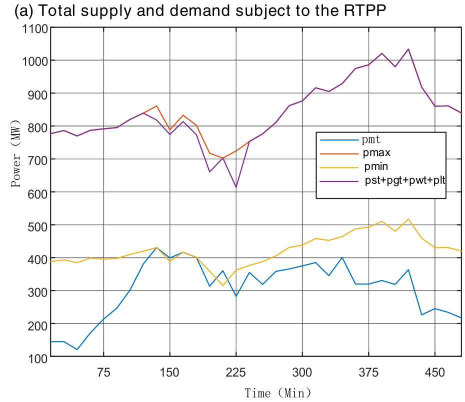


FIGURE 9. The lowest cost of PGD in the microgrid with the RTPP.

Following the RTPP based control strategy discussed in this section and considering that the DG in the microgrid cannot satisfy all the users' demand, FIGURE 10(a) shows relationships between the amount of supplied power, the maximum power demand and the non-flexible loads, and FIGURE 10(b) shows relationships between the power output of the DG in the microgrid, the amount of power to buy and the total power demand in the microgrid. It can be seen from FIGURE 10, that the RTPP based control strategy can achieve the balance between the power supply and demand of microgrid through economic leverage, and can guide the flexible load to reasonably adjust the power demand, to achieve the peak shaving and valley filling of the microgrid.



(a) Total supply and demand subject to the RTPP



(b) DG output, non-flexible loads and the maximum power demand

FIGURE 10. Relationships between the amount of power and the supply and demand in the microgrid with the RTPP.

FIGURE 10 suggests that when there is a large gap between the distributed energy and the total demand in the microgrid, under the optimized strategy there are still some adjustments for the balance between supply and demand in a small number of intervals. During these intervals, the users' power consumption is not completely limited. It is just that the amount of power bought by the microgrid operator for the optimal operating cost is certain. If some users indeed have extra power demand, the microgrid operator can increase the amount of supplied power and the increased RTPP (up to the defined highest rational RTPP) can apply to the extra power supplied, which can keep the balance between supply and demand. In terms of rationality, we assume that most of the flexible loads in the microgrid are rational or relatively rational so when the power prices increase to 50% higher than the normal prices, the rational flexible loads will reduce their power demand or adjust their plans of using power. In this case, the microgrid operator will buy some power from the distribution networks when the DG resources in the microgrid cannot provide enough power, and higher RTPP will apply to these customers. Therefore, the RTPP based EMCS will not suppress power consumption. Moreover, FIGURE 8(b) shows that the RTPP is not always much higher. Sometimes the RTPP is only slightly higher than the normal prices and it can increase with the widening gap between the total power demand and distributed energy in the microgrid. FIGURE 8(b) also indicates that the RTPP fluctuates within a reasonable range which is completely acceptable to part of the flexible loads. It should be noticed that it is not when the distributed energy cannot satisfy all the loads in the microgrid, immediately, the flexible loads will be charged a very high RTPP, making the power demand of most flexible loads not satisfied. Therefore, it also shows that the microgrid energy management and control strategy based on RTPP is more rational and feasible.

V. APPLICATIONS OF BOUNDED RATIONALITY TO THE REAL-TIME EMCS FOR MICROGRID

Decision making in human economic behavior is usually reflected in the process of seeking the optimum for objective function variables. In this process, optimal function variables are approximated in a feasible region until a balance is reached. This kind of decision making can be characterized as using bounded rationality to approximate perfect rationality. Because solving the objective optimization function can hardly find a solution that agrees with the defined target, with bounded rationality approximate solutions of objective function variables are usually obtained, which can satisfy the decision making of the optimization target [35]. Just like Simon's definition of bounded rationality, he thought that bounded rationality behavior is a kind of behavior which is expecting reasonableness subjectively but is limited to objectively [36]. The theory of bounded rationality put forward by Simon is based on a satisfaction principle, i.e., the principle of satisfying decision-makers. Simon thought problems are similar and solutions are also similar. If an approximate solution gives a decision that can satisfy decision-makers, the decision given by the approximate solution is a good enough solution. For the decision-making model of realistic optimization, in many cases, it is difficult to obtain an accurate optimal solution. Therefore, compromise is made by using experimental data from optimization control to check and correct the systematic error in solving the given model, thereby finding a satisfactory solution for the mathematic model of the optimization problem and moreover fitting the approximate solution to the accurate optimal solution of the model. Generally, if an approximate solution does not have an appreciable influence or impact on the target of decision making, then the results obtained based on the assumption of perfect rationality are still reasonable and acceptable in most cases. Therefore, according to the approximation theorem for optimization problems with bounded rationality, for mathematic models of objective functions, solutions satisfying bounded rationality are in fact approximating to perfect rationality under the bounded rationality condition [37]. The results of Yu are theoretically and practically important.

Theorem 1 R is the space of real numbers, X is a bounded closed interval in $R. f: X \to R$ is a semi-continuous function. Assume:

(1) $\{f_n\}$ is a real-valued function defined in X, satisfying that

$\underset{x\in X}{Sup}\,|f_n(x) - f(x)| \to 0 \ \ (n \to \infty)$;

(2) A is a nonempty bounded closed set in X.

(3) $\forall n = 1, 2, 3\cdots, x_n \in X$, satisfying $d(x_n, A) \to 0 \ \ (n \to \infty)$, and

$$f_n(x_n) \le \min_{x\in A} f_n(x) + \varepsilon_n$$

where $\varepsilon_n \ge 0$, $\varepsilon_n \to 0$ $(n \to \infty)$. There are

(1) Sequence $\{x_n\}$ must have a convergent subsequence $\left\{x_{n_k}\right\}$, making $x_{n_k} \to x^* \in A$

(2) $f(x^*) = \min_{x\in A} f(x)$

(3) If $f(x)$ at $x \in A$ has a minimum point set $\{x^*\}$ and it is a single point set, there must be $x_n \to x^*$.

Whether the RTPP based EMCS for the microgrid mentioned in the previous section is a bounded rationality approximation of perfect rationality, moreover proving that the obtained EMCS is optimal. Yu's model of bounded rationality is used to discuss this question as below. For the convenience of solving, the simulations in the previous section are completed based on transforming multiple objective functions to single-objective functions as follows,

$$\min(\phi_h \sum_{h\in G_h} C_h(p_h(t)) + \phi_b \sum_{b\in G_b} C_b(p_b(t)) + \phi_d \sum_{d\in D} C_d(p_d(t)) + \phi_m(\rho_r(t) - \rho_l)\cdot p_m(t))$$

Costs of hydropower, gas power and battery charging and discharging rely on the amount of generated or stored power. When smart meters in the microgrid record the amount of power from the DG and have the data transmitted to the center of energy managing and scheduling, the specific cost will be obtained. Therefore, the costs can be treated as an overall cost function of time, and the above objective cost function can be written as

$$\min(\phi_C C(t) + \phi_m(\rho_r(t) - \rho_l)\cdot p_m(t)) \tag{33}$$

where $\rho_r(t)$ is an exponential function about the amount of power to buy and it is also the demand side price response function in the energy management control model, and its expression is as follows:

$$\rho_r(t) = e^{\frac{Kp_m(t)}{P_{l\max}(t) - P_{l\min}(t)}}\cdot \rho_l$$

Since exponential functions are continuous, $\rho_r(t)$ is a continuous function. If the amount of power to buy is represented using x, and the objective function is expressed as an objective function of x, i.e., $f(x(t))$, then equation (22) can be written as

$$f(x(t)) = \min(\phi_C C(t) + \phi_m(\rho_r(x(t)) - \rho_l)\cdot x(t)) \tag{34}$$

Because the amount of power to buy can only be a limited range of real numbers, we can assume that within the range of two constraints the maximum amount of power to buy is M_x. Therefore, the range of x is $x \in [0, M_x] \subset R$, and equation (34) and the above discussion can confirm that $f(x(t))$ is continuous, thereby satisfying one of the application criteria on variables and objective functions required by the above approximation theorem. Because $x \in [0, M_x] \subset R$, x can be found in a bounded closed set and the bounded closed set can be written as $A = [0, M_x]$. For solving the above objective function, the deep adaptive dynamic programming optimization algorithm is used. Being adaptive is a process of gradually approximating. In searching for the optimum using the deep adaptive dynamic programming optimization algorithm will give a group of optimized approximate solution sequence interval A_n first, where a solution in the approximate solution sequence is x_n. x_n is an approximate value of ε_n and is obtained by evaluating $f_n(x(t))$ in an extremely small range $[0, \delta]$, and $x_n \in A_n$. Substituting each x_n in the approximate solution sequence interval A_n into the objective function will get an approximate solution sequence $f_n(x_n(t))$ of the objective function. If the difference between the obtained approximate solution of the objective function and the required minimum of the objective function is smaller than a minimal MSE, then the objective function value to be found is the one obtained using the approximate solution. This can be expressed as

$$|f_n(x_n(t)) - f(x_n(t))| < \varepsilon_n \tag{35}$$

where ε_n represents the MSE. The minimal MSE found in the case study of the previous section is 5.1606×10^{-23} which is a very small number. Therefore, the minimal MSE can be considered as sufficiently close to zero, and equation (35) can be written as:

$$|f_n(x_n(t)) - f(x_n(t))| \to \varepsilon_n \tag{36}$$

Since the minimal MSE is small enough, ε_n in equation (36) meets the condition of $\varepsilon_n \geq 0$, $\varepsilon_n \to 0$. The following summaries can be drawn according to the theorem, where x^* is the optimal value satisfying the solution of the objective function:

(1) Sequence $\{x_n\}$ must have a convergent subsequence $\{x_{n_k}\}$, making $x_{n_k} \to x^* \in A$

(2) $f(x^*) = \min_{x\in A} f(x)$

(3) If $f(x)$ at $x \in A$ has a minimum point set $\{x^*\}$ and it is a single point set, there must be $x_n \to x^*$.

Therefore, the solution sequence obtained after optimizing iterations must have $\{x_n\}$ which is approximating the optimal solution x^*, making $f(x(t))$ reach the cost of the minimum objective function $f(x^*(t))$ corresponding to the amount of power to buy. This is the other criterion for applying the approximation theorem of bounded rationality, which can reach the objective of the lowest cost of PGD in the microgrid. Therefore, the satisfactory solution of the control strategy obtained by the deep adaptive dynamic programming optimization algorithm is the optimal solution, and the strategy itself is the optimal strategy based on the objective function cost model. The calculation and simulation of the control strategy is a practical application of the finite rationality to complete rationality approximation theory.

In fact, according to the discussion above, there remains the following normal framework and summaries in terms of topology:

Assume (X, d) is a metric space,

$$Y = \left\{ \begin{array}{c} y=(f,A) : f \text{ is continuous at} X,\ \sup_{x\in X} |f(x)| < +\infty, \\ A \text{ is a nonempty compact set of } X \end{array} \right\}$$

$$\forall y_1 = (f_1, A_1),\ y_2 = (f_2, A_2) \in Y, \text{ define :}$$

$$\rho(y_1, y_2) = \sup_{x\in X} |f(x_1) - f(x_2)| + h(A_1, A_2)$$

where h is the Hausdorff distance at X. It is easy to prove that (Y, ρ) is a complete metric space.

$\forall y = (f, A) \in Y$, select the function sequence $\{f^n\}$, making $\sup_{x\in X} |f^n(x) - f(x)| \to 0$ $(n \to \infty)$ choose the subset sequence $\{A_n\}$ from X, and satisfying $h(A_n, A) \to 0$ $(n \to \infty)$. Select $x_n \in X$, making $d(x_n, A_n) \to 0$ $(n \to \infty)$, and

$$f^n(x_n) \leq \inf_{u\in A_n} f^n(u) + \varepsilon_n$$

where $\varepsilon_n \geq 0$, $\varepsilon_n \to 0$ $(n \to \infty)$

Theorem 2 There exists a dense residual set Q of Y, making $\forall y = (f, A) \in Q$, there must be $x_n \to x$, and $f(x) = \min_{u\in A} f(u)$.

The objective function of the microgrid is approximate. The set of feasible solutions is approximate. The parameters are approximate. The accuracy of solving is also approximate. Therefore, an approximating sequence $\{x_n\}$ is obtained, and there must be a convergent subset $\{x_{n_k}\}$ of $\{x_n\}$, i.e., $x_{n_k} \to x \in A$ and x must be the solution of the optimization problem, which reflects bounded rationality approximating perfect rationality. Since Q belongs to the second class, it can be considered from the topological sense that most satisfactory solutions of the RTPP based control strategy are convergent, rational, and acceptable. They can all converge to the optimal solution, and the counter example of their non-convergence belongs to the first class. In other words, in most cases the satisfactory solution under bounded rationality can replace the accurate solution under perfect rationality, and the counter cases belonging to the first class are extremely few. Explanations to some of the above concepts are as follows:

(1) Assume A is a nonempty subset of the metric space X, if $\bar{A} = X$, then A is a dense set of X. Assume A is a nonempty subset of the metric space X, if $\mathrm{int}(\bar{A}) = \phi$, then A is a nowhere dense set of X. A union of countable numbers of nowhere dense sets is called the first class, otherwise, it is called the second class.

(2) Assume X is a metric space, if Q, a subset of X, contains an intersection of dense open sets of X, then Q is called a residual set of X. Assume X is a complete metric space, if Q is a residual set of X, then Q must be a dense set of the second class.

(3) Assume X is a complete metric space, then a residual set Q of X must be dense and belong to the second class. If $\forall x \in Q$ and property P dependent on x holds, then P is called a generic property of X or property P holds generically for X. In terms of the Baire classification, nonlinear analysis or topology, the first class is considered as a "small set". Therefore, it can be said that for most $x \in X$ the property P dependent on x holds.

VI. CONCLUSION

In this paper, the distributed energy management model of the microgrid under the RTPP of DSR is proposed by using the method of system control theory, and a deep adaptive dynamic programming optimization algorithm is proposed for the framework, forming the real-time optimal control strategy of microgrid energy management. The real-time two-way interaction between power consumers in the microgrid and the microgrid operator is realized, based on which the mechanism of the DSR and RTPP based energy management for the microgrid is created. The RTPP based EMCS can not only get the dynamic balance between supply and demand but also through economic leverage motivate the flexible loads to consume power rationally, fulfilling peak shaving and valley filling for the intraday loads in the microgrid. This makes the control strategy more reasonable and ensures the stable and safe operation of the microgrid. Because the EMCS is real-time, the microgrid operator can combine the RTPP based EMCS with relevant rules of management and operation to perform intraday scheduling. At the same time, the simulation results show that the proposed control strategy is feasible and effective. Finally, according to the theory of bounded rational approximation, the application conditions of the control strategy in the power market are proposed. It is proved that the control strategy meets the application conditions, which is a specific application of the theory of bounded rational approximation in the power market. Since the simulation of the control strategy satisfies the application condition that the price response function is continuous and the approximation iteration process is continuous, the control strategy can be accepted. Because the set Q of the bounded rational solution contains the completely rational optimal solution, the multi-objective control strategy derived from the deep adaptive dynamic programming optimization algorithm is a rational, feasible, and acceptable optimal control strategy. This strategy can help microgrid operators make real-time online multi-objective decisions, optimize operating costs, improve the utilization rate of renewable energy, and effectively reduce the carbon emissions and environmental pollution caused by the massive use of fossil energy.

REFERENCES

[1] C. Li, X. Yu, W. Yu, T. Huang, and Z.-W. Liu, "Distributed event-triggered scheme for economic dispatch in smart grids," *IEEE Trans. Ind. Informat.*, vol. 12, no. 5, pp. 1775–1785, Oct. 2016.

[2] S. Lee, B. Do Chung, H. W. Jeon, and J. Chang, "A dynamic control approach for energy-efficient production scheduling on a single machine under time-varying electricity pricing," *J. Cleaner Prod.*, vol. 165, pp. 552–563, Nov. 2017.

[3] C. Luo, Y.-F. Huang, and V. Gupta, "Stochastic dynamic pricing for EV charging stations with renewable integration and energy storage," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1494–1505, Mar. 2018.

[4] M. Marzband, E. Yousefnejad, A. Sumper, and J. L. Domínguez-García, "Real time experimental implementation of optimum energy management system in standalone microgrid by using multi-layer ant colony optimization," *Int. J. Electr. Power Energy Syst.*, vol. 75, pp. 265–274, Feb. 2016.

[5] A. V. Jayawardena, L. G. Meegahapola, D. A. Robinson, and S. Perera, "Representation of a grid-tied microgrid as a reduced order entity for distribution system dynamics and stability studies," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 591–600, Dec. 2015.

[6] J. C. Pena-Aguirre, A.-I. Barranco-Gutierrez, J. A. Padilla-Medina, A. Espinosa-Calderon, and F. J. Perez-Pinal, "Fuzzy logic power management strategy for a residential DC-microgrid," *IEEE Access*, vol. 8, pp. 116733–116743, 2020.

[7] A. Ghasemi and M. Enayatzare, "Optimal energy management of a renewable-based isolated microgrid with pumped-storage unit and demand response," *Renew. Energy*, vol. 123, pp. 460–474, Aug. 2018.

[8] M. Nemati, M. Braun, and S. Tenbohlen, "Optimization of unit commitment and economic dispatch in microgrids based on genetic algorithm and mixed integer linear programming," *Appl. Energy*, vol. 210, pp. 944–963, Jan. 2018.

[9] T. A. Nguyen and M. L. Crow, "Stochastic optimization of renewable-based microgrid operation incorporating battery operating cost," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2289–2296, May 2016.

[10] M. Zare, T. Niknam, R. Azizipanah-Abarghooee, and A. Ostadi, "New stochastic bi-objective optimal cost and chance of operation management approach for smart microgrid," *IEEE Trans. Ind. Informat.*, vol. 12, no. 6, pp. 2031–2040, Dec. 2016.

[11] M. Jin, W. Feng, C. Marnay, and C. Spanos, "Microgrid to enable optimal distributed energy retail and end-user demand response," *Appl. Energy*, vol. 210, pp. 1321–1335, Jan. 2018.

[12] V. S. Tabar, M. A. Jirdehi, and R. Hemmati, "Energy management in microgrid based on the multi objective stochastic programming incorporating portable renewable energy resource as demand response option," *Energy*, vol. 118, pp. 827–839, Jan. 2017.

[13] K. Rahbar, J. Xu, and R. Zhang, "Real-time energy storage management for renewable integration in microgrid: An off-line optimization approach," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 124–134, Jan. 2015.

[14] J. Campillo, E. Dahlquist, F. Wallin, and I. Vassileva, "Is real-time electricity pricing suitable for residential users without demand-side management?" *Energy*, vol. 109, pp. 310–325, Aug. 2016.

[15] X. Yang, Y. Zhang, B. Zhao, F. Huang, Y. Chen, and S. Ren, "Optimal energy flow control strategy for a residential energy local network combined with demand-side management and real-time pricing," *Energy Buildings*, vol. 150, pp. 177–188, Sep. 2017.

[16] K. Ma, T. Yao, J. Yang, and X. Guan, "Residential power scheduling for demand response in smart grid," *Int. J. Electr. Power Energy Syst.*, vol. 78, pp. 320–325, Jun. 2016.

[17] D. Xu, Q. Wu, B. Zhou, C. Li, L. Bai, and S. Huang, "Distributed multi-energy operation of coupled electricity, heating, and natural gas networks," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2457–2469, Oct. 2020.

[18] P. Siano, C. Cecati, H. Yu, and J. Kolbusz, "Real time operation of smart grids via FCN networks and optimal power flow," *IEEE Trans. Ind. Informat.*, vol. 8, no. 4, pp. 944–952, Nov. 2012.

[19] S. A. Pourmousavi, M. H. Nehrir, C. M. Colson, and C. Wang, "Real-time energy management of a stand-alone hybrid wind-microturbine energy system using particle swarm optimization," *IEEE Trans. Sustain. Energy*, vol. 1, no. 3, pp. 193–201, Oct. 2010.

[20] J. Xiao, L. Setyawan, P. Wang, and C. Jin, "Power-capacity-based bus-voltage region partition and online droop coefficient tuning for real-time operation of DC microgrids," *IEEE Trans. Energy Convers.*, vol. 30, no. 4, pp. 1338–1347, Jul. 2015.

[21] V. Boscaino, J. M. Guerrero, I. Ciornei, L. Meng, E. R. Sanseverino, and G. Zizzo, "Online optimization of a multi-conversion-level DC home microgrid for system efficiency enhancement," *Sustain. Cities Soc.*, vol. 35, pp. 417–429, Nov. 2017.

[22] B. Cheng and W. Powell, "Co-optimizing battery storage for the frequency regulation and energy arbitrage using multi-scale dynamic programming," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1997–2005, May 2018.

[23] P. Zeng, H. Li, H. He, and S. Li, "Dynamic energy management of a microgrid using approximate dynamic programming and deep recurrent neural network learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4435–4445, Jul. 2019.

[24] A. Das and Z. Ni, "A computationally efficient optimization approach for battery systems in islanded microgrid," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6489–6499, Nov. 2018.

[25] B. Heymann, J. F. Bonnans, P. Martinon, F. J. Silva, F. Lanas, and G. Jiménez-Estévez, "Continuous optimal control approaches to microgrid energy management," *Energy Syst.*, vol. 9, no. 1, pp. 59–77, Feb. 2018.

[26] J. B. Almada, R. P. S. Leão, R. F. Sampaio, and G. C. Barroso, "A centralized and heuristic approach for energy management of an AC microgrid," *Renew. Sustain. Energy Rev.*, vol. 60, pp. 1396–1404, Jul. 2016.
[27] A. Merabet, K. Tawfique Ahmed, H. Ibrahim, R. Beguenane, and A. M. Y. M. Ghias, "Energy management and control system for laboratory scale microgrid based wind-PV-battery," *IEEE Trans. Sustain. Energy*, vol. 8, no. 1, pp. 145–154, Jan. 2017.
[28] Y. Liu, C.-S. Tan, and K.-L. Lo, "Effects of locational marginal price on electricity market with distributed generation," *Proc. CSEE*, vol. 27, no. 31, pp. 89–97, 2007.
[29] S. Sun, M. Dong, and B. Liang, "Joint supply, demand, and energy storage management towards microgrid cost minimization," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2014, pp. 109–114.
[30] R. Bellman, "Dynamic programming," *Science*, vol. 153, nos. 37–31, pp. 34–37, 1966.
[31] H.-G. Zhang, X. Zhang, Y.-H. Luo, and J. Yang, "An overview of research on adaptive dynamic programming," *Acta Automatica Sinica*, vol. 39, no. 4, pp. 303–311, Mar. 2014.
[32] W. B. Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Hoboken, NJ, USA: Wiley, 2007.
[33] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, May 2010.
[34] K. G. Vamvoudakis and F. L. Lewis, "Online solution of nonlinear two-player zero-sum games using synchronous policy iteration," *Int. J. Robust Nonlinear Control*, vol. 22, no. 13, pp. 1460–1483, 2012.
[35] H. Wang and J. Yu, "Bounded rationality and stability of solutions of multiobjective optimization problems," *Oper. Res. Trans.*, vol. 12, no. 1, pp. 104–108, 2008.
[36] H. A. Simon, *Administrative Behavior*. Noida, India: Simon and Schuster, 2013.
[37] J. Yu, *Bounded Rationality and the Stability of the Set of Equilibrium Points in Game Theory*, Beijing, China: Science Press, 2017, pp. 53–125.

NAN WU was born in Guiyang, Guizhou, China. He received the B.S. degree from the Guizhou University of Technology, Guiyang, in 1999, and the master's and Ph.D. degrees from Guizhou University, China, in 2005 and 2019, respectively. He joined the School of Mechanical and Electrical Engineering, Guizhou Normal University, in 2008. From December 2010 to March 2011, he was a Visiting Fellow with the Department of Electronic Engineering, City University of Hong Kong. His research interests include optimal control theory and antenna designs.

HONGLEI WANG was born in Beijing, China, in 1959. He received the B.S. degree from the Guizhou Institute of Technology, Guiyang, China, in 1982, and the Ph.D. degree from CAS, Beijing, in 2008. From 1982 to 1986, he was a Lecturer with the Department of Electrical Engineering, Guizhou University of Technology, Guiyang. From 1991 to 1993, he was a Visiting Fellow with the Institute of Modeling and Control, Mediterranean Institute of Technology, France. From 1996 to 1998, he was an Associate Professor with the Guizhou University of Technology. He was a Professor with the Guizhou University of Technology, in 2002. Since 2004, he has been a Professor with Guizhou University, Guiyang. His current research interests include application of network technology in control systems, management system engineering (optimization and decision direction), optimization control of complex system, and so on.

LINFEI YIN was born in Jiujiang, Jiangxi, China, in 1990. He received the B.S. and M.S. degrees in information engineering from Nanchang Hangkong University, Nanchang, China, in 2012 and 2015, respectively, and the Ph.D. degree from the College of Electric Power Engineering, South China University of Technology, Guangzhou, China, in 2018. He is currently an Assistant Professor with the College of Electrical Engineering, Guangxi University, Nanning, Guangxi, China. His research interest includes artificial intelligence techniques in operation of power systems.

XUFENG YUAN (Member, IEEE) was born in Guizhou, China, in 1976. He received the Ph.D. degree from the Huazhong University of Science and Technology, Wuhan, in 2007. He is currently a Professor with the Department of Electrical Engineering, Guizhou University. His current research interests include electronic power and distribution networks.

XIAOXIA LENG was born in Guizhou, China, in 1962. She received the bachelor's degree from Southwest University, Chongqing, China, in 1984, and the master's degree from Wuhan University, Wuhan, China, in 2003. She is currently an Associate Professor with the School of Management, Guizhou University. Her current research interests include human resources, social security, complex system optimization and management, and so on.