IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 22, 2021, accepted March 23, 2021, date of publication March 29, 2021, date of current version April 22, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3069278

Stochastic Dynamic Programming-Based Online Algorithm for Energy Management of Integrated Energy Buildings With Electric Vehicles and Flexible Thermal Loads

RENLI CHENG[1], WEIJIE CHENG[1], JIE LI[2], ZEQI CHEN[1], JUN SHI[1], ZHENNING PAN[2], YUFENG WU[2], AND TAO YU[2], (Member, IEEE)

[1]Shenzhen Power Supply Bureau Company, Ltd., Shenzhen 518000, China
[2]College of Electric Power, South China University of Technology, Guangzhou 510640, China

Corresponding author: Zhenning Pan (scutpanzn@163.com)

This work was supported by the Science and Technology Project of China Southern Power Grid under Grant 090000KK52190072.

ABSTRACT With the rapid development of economy and technology, large-scale integrated energy buildings account for an increasing proportion of urban load. However, the randomness of EV owner behaviors, electricity price and outdoor temperature have brought challenges to the energy management of integrated energy buildings. This paper proposes a stochastic dynamic programming-based online algorithm to address the energy management of integrated energy buildings with electric vehicles and flexible thermal loads under multivariate uncertainties. First, an online energy management framework is established, which is further formulated as a multi-stage stochastic sequential decision-making problem. To address the complexities of the problem, a novel stochastic dynamic programming is employed to develop a distribution-free, computationally efficient, and scalable solution. By using extensive training samples, the algorithm is trained offline to learn how to deal with multivariate uncertainties and get the approximate optimal solution, which no longer depends on intraday forecast information. Numerical tests demonstrate the effectiveness of the proposed algorithm compared with other online algorithms in terms of optimality and computation efficiency.

INDEX TERMS Stochastic dynamic programming, online algorithm, energy management, integrated energy building, multivariate uncertainties.

I. INTRODUCTION

Strengthening the connection between various energy sources have become a necessary way for the power grids in the world today [1]–[3]. It is worth mentioning that the integrated energy system realizes integration optimization and multi-energy complementation by fusing electricity, gas, heat, cold and other energy forms. Furthermore, the integrated energy building is an important application form of the integrated energy system, which uses combined cooling, heating and power (CCHP) as the key technology, and unites advanced control, communication and management methods to build a building energy management system (BEMS). In developed cities, the load of large commercial and residential buildings exceeds 30% of the total load [4]. Therefore, exploiting the potential of energy management represented by integrated energy buildings is of great significance for improving electric power usage.

There have been numerous research on how to realize the energy management of BEMS. Reference [5] presents mathematical optimization models of residential energy hubs with the objective functions of minimizing energy consumption, total cost of electricity and gas, emissions, peak load. Reference [6] proposes a residential energy hub model and performs optimal load management, which realizes co-optimization of electricity consumption costs and carbon emissions reductions. Reference [7] describe a MILP model for determining the optimal capacity and operation of seven CCHP systems in the heating and cooling network of a residential district. Reference [8] establishes the cooling and

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

electricity coordinated MG day-ahead scheduling and real-time dispatching model.

The above research provides a variety of methods to promote the collaborative optimization of multiple energy sources in integrated energy buildings, but how to obtain the real-time optimal policy under multivariate uncertainties has not been comprehensively studied. Reference [9] uses chance constrained programming to solve the economic operation problem of the building-level integrated energy system under uncertain factors. Reference [10] verifies the effectiveness of the robust index method to handle the uncertainties of customer behavior and proves its application in load scheduling.

But most of the literature considers the influence of randomness from the perspective of day-ahead scheduling decision-making. Therefore, many literatures use model predictive control (MPC) to solve the problem of real-time scheduling of integrated energy buildings considering randomness. Reference [11] proposes an appliance scheduling scheme for residential building energy management controllers based on MPC method. Reference [12] uses MPC method to handle the uncertainties and price variations, and proposed a scenario-based branch-and-bound approach to solve MES scheduling in urban buildings. Reference [13] designs an MPC controller to reduces the operating and maintenance cost in a district heating system, considering TES and flexible loads. However, the MPC method is highly dependent on the load and electricity price forecast accuracy.

In order to make up for the shortcomings of the above-mentioned literature, this paper proposes a stochastic dynamic programming-based online algorithm to address the energy management of integrated energy buildings with electric vehicles and flexible thermal loads under multivariate uncertainties such as EV behaviors, load, and electricity price. First, we establish an economic dispatch model for integrated energy buildings, and transform the multi-period decision-making problem into a recurrence problem. To address the complexities of the problem, a novel stochastic dynamic programming is employed to develop a distribution-free, computationally efficient, and scalable solution. The approximate form of the value function is constructed to deal with the multivariate uncertainties of the building. The value function is trained based on the SPAR algorithm, and a convergent approximate value function is obtained. Put the convergent approximation function into online operation, and the real-time energy control problem of integrated energy buildings is solved time by time.

Compared with existing research, the major contributions of this paper are as follows:

- A novel stochastic dynamic programming is proposed to address the complexities of the problem, and the highly close-to-optimal online solution can be obtained by using extensive training samples.
- Multivariate uncertainties are sufficiently considered, and multiple flexible resources are jointly optimized in the energy management of integrated energy buildings, which is further formulated as a multi-stage stochastic sequential decision-making problem.
- The proposed algorithm makes operators no longer rely on intraday forecast information to make approximate optimal decisions, which reduces the impact of multivariate uncertainties on energy management. Simulation validates its better performance than MPC and Myopic algorithm.

The rest of this paper is arranged as follows. Section I introduces the integrated energy building operation framework. In Section II, the economic dispatch model of the building energy management system is established, and the objective function with the maximum expectation of total benefits considering multivariate uncertainties is constructed. Section III introduces the proposed real-time energy management algorithm for integrated energy buildings, including model conversion, processing and solution. Section IV simulates an example of a comprehensive energy building in the south. Section V gives conclusions.

II. MODELING FOR INTEGRATED ENERGY BUILDINGS

A. THE FRAMEWORK OF INTEGRATED ENERGY BUILDINGS

The operating framework of integrated energy building is shown in Figure 1. The energy input of the building includes the distribution network and natural gas station, while the energy output contains the electrical load, heating load, and cooling load. Electrical load contains basic load such as drainage, office, lighting, elevator, etc. In addition, the building also includes CCHP, EV, and energy storage system (ESS). The CCHP system in this paper consists of gas generator, heat recovery device, electric refrigerator, absorption refrigerator, water storage tank, heat exchanger and so on. The operation mode of ''determining electricity by heat'' is adopted by CCHP. The operator of the building purchases gas from the natural gas station to supply the gas turbine with the aim of producing electricity. The wasted heat can be used as a by-product to supply the heat load. ESS and EV are power storage devices, which are chosen by the operator to charge or discharge according to the current electricity price and total load. The operator integrates various resources of power supply, load and energy storage in the building, aiming to realize the coordination of the three and meet the needs of electricity, cooling and heating loads. At the same time, the operator tries to reduce operating costs and improve economic benefits.

B. ECONOMIC DISPATCH MODEL OF INTEGRATED ENERGY BUILDINGS

1) CONSTRAINTS OF POWER BALANCE

$$P_t^{\mathrm{DN}} + P_t^{\mathrm{cchp}} + \sum_{i=1}^{N_{\mathrm{ESS}}} P_{i,t}^{\mathrm{ESS}} = P_t^{\mathrm{CON}} + \sum_{i=1}^{N_{\mathrm{EV}}} P_{i,t}^{\mathrm{EV}} \tag{1}$$

where P_t^{DN} is the active power transmitted from the distribution network to the building at time t; P_t^{cchp} is the power

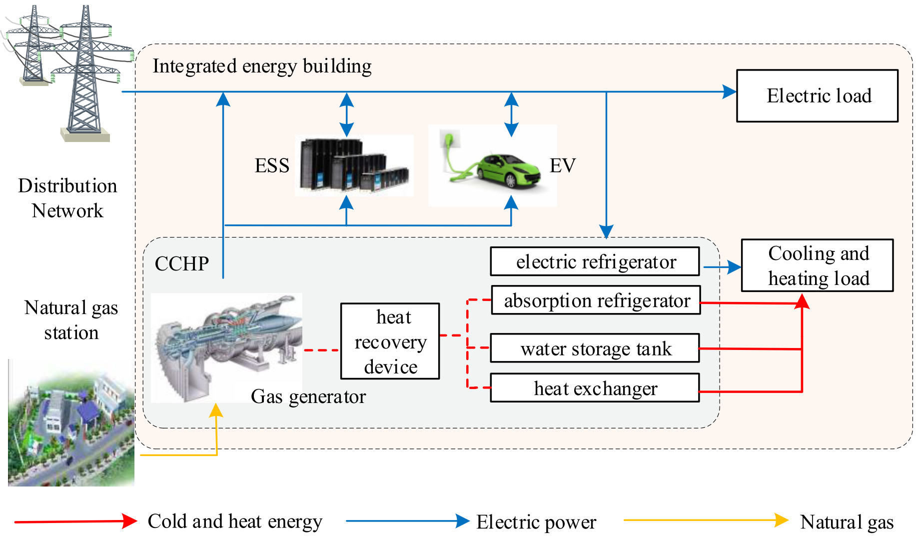


FIGURE 1. Framework of integrated energy buildings.

output of CCHP at time t; $P_{i,t}^{\mathrm{ESS}}$ is the active output of ESS at time t, which is positive when discharging and negative when charging; P_t^{CON} is the electrical load except for EV at time t; $P_{i,t}^{\mathrm{EV}}$ is the active load of EV at time t, which is positive when charging and negative when discharging; N_{EV} is the number of EV and N_{ESS} is the number of ESS.

2) CONSTRAINTS OF ESS

$$E_{i,t+1} = \begin{cases} E_{i,t} - \dfrac{1}{\eta_i} P_{i,t} & P_{i,t} \geq 0 \\ E_{i,t} - \eta_i P_{i,t} & P_{i,t} < 0 \end{cases} \tag{2}$$

$$P_{i,t,\min} \leq P_{i,t} \leq P_{i,t,\max} \tag{3}$$

$$E_{i,t,\min} \leq E_{i,t} \leq E_{i,t,\max} \tag{4}$$

where $E_{i,t}$ denotes the energy of the i-th ESS at time t. $P_{i,t}$ is the power output of the i-th ESS at time t, which is positive when discharging and negative when charging. η_i is the efficiency of charging and discharging. Equations (3) and (4) are the upper and lower limits of energy and power respectively. Since the ESS cannot be charged and discharged at the same time, equation (2) is replaced with the following equation [14]:

$$\begin{cases} P_{i,t} = P_{i,t}^{+} - P_{i,t}^{-} \\ E_{i,t+1} = E_{i,t} - \dfrac{1}{\eta_i} P_{i,t}^{+} + \eta_i P_{i,t}^{-} \\ 0 \leq P_{i,t}^{+} \leq P_{i,\max}^{+}, 0 \leq P_{i,t}^{-} \leq P_{i,\max}^{-} \end{cases} \tag{5}$$

As the mathematical proof showed in [14], (2) and (5) are equivalent only if equation (6) holds,

$$P_{i,t}^{+} \bullet P_{i,t}^{-} = 0 \tag{6}$$

That is to say, the optimal solution that satisfies the model will automatically satisfy the complementary constraints (6) that cause strong non-convexity.

3) CONSTRAINTS OF EV

As for a single EV, it is connected to the charging pile in the parking lot of the building at t_{in} and intends to leave at t_{out}.

Assuming that the EV is charged with the maximum power after it is connected to the charging pile, and is charged to the desired power at t_{limit}. Then we define the upper boundary of the energy to represent the power change curve during this period. Assuming that after the EV is plugged in, the charging is delayed until the moment of departure just reaches the user's expected power value, then we define the lower boundary of the energy as the power change curve during this period. The upper and lower boundaries of energy reflect the adjustable characteristics of the EV [15]. When $t_{\mathrm{out}} > t_{\mathrm{limit}}$, it means that this EV can be used as an adjustable load to participate in BEMS scheduling, and only needs to meet the power constraints and energy constraints requirements at each time t [15], [16]:

$$\begin{cases} E_{i,t} = \eta \sum_{k=t_{\mathrm{start}}}^{t} P_{i,k} \Delta t = E_{i,t-1} + \eta P_{i,k} \Delta t \\ E_{i,t_{\mathrm{end}},\min} = E_{i,t_{\mathrm{end}},\max} = E_{\exp} \\ E_{i,t,\min} \leq E_{i,t} \leq E_{i,t,\max} \\ P_{i,t,\max} = \min(P_{\max}, (E_{i,t,\max} - e_{t-1})/\eta/\Delta t) \\ P_{i,t,\min} = \max(0, (E_{i,t,\min} - E_{i,t-1})/\eta/\Delta t) \\ P_{i,t,\min} \leq P_{i,t} \leq P_{i,t,\max} \end{cases} \tag{7}$$

where Δt is the time interval; $E_{i,t}$ is the power of the i-th EV battery at time t; $P_{i,k}$ represents the charging power of the i-th EV at time k; $E_{i,t,max}$ and $E_{i,t,min}$ are the upper and lower boundaries of EV energy at time t respectively; $E_{\exp}$ is the expected charging capacity of the user, which generally represents the maximum capacity of the battery; $P_{i,t,max}$ and $P_{i,t,min}$ are the upper and lower limits of EV charging power at time t. P_{max} indicates the maximum charging power affected by the charging pile and the EV itself.

With the aim of avoiding the emergence of ''dimensionality disaster'', EVs are clustered once every t. Consider EVs with the similiar departure time as the same cluster. The charging model of a single EV in the cluster can be superimposed to obtain an equivalent cluster model. The correctness of this method can be proved in the reference [15], [16].

4) CONSTRAINTS OF THERMOSTATICALLY CONTROLLED LOAD

Assuming that the water tank is always in a state of full water, and ignoring the dynamic process of water flow. Therefore, the mathematical model of the water tank can be expressed by equation (8),

$$\begin{cases} T_{t+1}^{\mathrm{wt}} = \dfrac{V_t^{\mathrm{cold}}\left(T^{\mathrm{cold}} - T_t^{\mathrm{wt}}\right) + VT_t^{\mathrm{wt}}}{V} + \dfrac{h_t^{\mathrm{wt}}\Delta t}{VC_{\mathrm{w}}} \\ T_{\min}^{\mathrm{wt}} \le T_t^{\mathrm{wt}} \le T_{\max}^{\mathrm{wt}} \end{cases} \tag{8}$$

where V is the volume of the water tank; C_{w} is the specific heat capacity of water; T_t^{wt} is the temperature of the water tank at time t; V_t^{cold} and T_t^{cold} are the volume and temperature of injected cold water at time t respectively; h_t^{wt} is the thermal power at time t; $T_{\min}^{\mathrm{wt}}$ and $T_{\max}^{\mathrm{wt}}$ are the upper and lower limits of the hot water temperature acceptable to the user.

On the other hand, the model of room temperature regulation system has been studied in reference [2], [3] and [17], [18]. Some are first-order models, some are second-order models, and some consider more factors, such as the influence of adjacent rooms. In this paper, we use a simplified and common first order mathematical model of room temperature regulation system on the basis of reference [17], [18] to reduce the complexity of the problem. The model is established discretely according to the heat balance principle, the external temperature conditions and building parameters. It can be expressed by equation (9) during cooling, and can be expressed by equation (10) during heating.

$$\begin{aligned} T_{t+1}^{\mathrm{indoor}} &= T_t^{\mathrm{indoor}} \mathrm{e}^{\Delta t/\mathrm{RC}_{\mathrm{air}}} + \left(\mathrm{e}^{\Delta t/\mathrm{RC}_{\mathrm{air}}} - 1\right) \\ &\quad \times \left(T_t^{\mathrm{outdoor}} - RP_t^{\mathrm{re}}\Delta t\right) \end{aligned} \tag{9}$$

$$\begin{aligned} T_{t+1}^{\mathrm{indoor}} &= T_t^{\mathrm{indoor}} \mathrm{e}^{-\Delta t/\mathrm{RC}_{\mathrm{air}}} - \left(\mathrm{e}^{\Delta t/\mathrm{RC}_{\mathrm{air}}} - 1\right) \\ &\quad \times \left(T_t^{\mathrm{outdoor}} + Rh_t^{\mathrm{heat}}\Delta t\right) \end{aligned} \tag{10}$$

$$T_{\min}^{\mathrm{indoor}} \le T_t^{\mathrm{indoor}} \le T_{\max}^{\mathrm{indoor}} \tag{11}$$

where T_t^{indoor} and T_t^{outdoor} are the indoor temperature and outdoor temperature respectively at time t; R the thermal resistance of the house; C_{air} is the specific heat capacity of air; P_t^{re} and h_t^{heat} are the cooling and heating power respectively at time t; $T_{\min}^{\mathrm{indoor}}$ and $T_{\max}^{\mathrm{indoor}}$ are the upper and lower limits of indoor temperature acceptable to users.

5) CONSTRAINTS OF CCHP

The CCHP system works in the mode of ''heating to determine power'' [19]. That is to say, according to the heating load at each moment, the corresponding heat energy and the corresponding proportion of electricity are output. Therefore, its output electric power and thermal power should satisfy:

$$\begin{cases} P_t^{\mathrm{cchp}} = \eta_e Q_{\mathrm{gas}} F_t^{\mathrm{cchp}} \\ h_t^{\mathrm{cchp}} = \eta_h Q_{\mathrm{gas}} F_t^{\mathrm{cchp}} \end{cases} \tag{12}$$

Furthermore, it's described as follows:

$$\frac{h_t^{\mathrm{cchp}}}{\eta_h} = \frac{P_t^{\mathrm{cchp}}}{\eta_e} \tag{13}$$

$$h_{\min}^{\mathrm{cchp}} \le h_t^{\mathrm{cchp}} \le h_{\max}^{\mathrm{cchp}} \tag{14}$$

where h_t^{cchp}, P_t^{cchp} and F_t^{cchp} are the thermal power, electric power and natural gas consumption of CCHP output respectively at time t. η_e and η_h are the efficiency of electricity and heat generation of the CCHP system respectively at time t. Q_{gas} is natural gas calorific value. $h_{\min}^{\mathrm{cchp}}$ and $h_{\max}^{\mathrm{cchp}}$ are the upper and lower limits of the output power.

6) THE OBJECTIVE FUNCTION

The operator of integrated energy building aims to minimize the total cost of energy management, including fuel costs namely gas purchase costs, electricity purchase costs, uncomfortable costs of temperature control loads, and ESS operating costs.

$$\min \sum_{t=1}^{T} \left(C_t^{\mathrm{gas}} + C_t^{\mathrm{DN}} + C_t^{\mathrm{tcl}} + \sum_{i=1}^{\mathrm{ESS}} C_{i,t}^{\mathrm{ESS}} \right) \tag{15}$$

$$\begin{cases} C_t^{\mathrm{gas}} = p^{\mathrm{gas}} \cdot F_t^{\mathrm{cchp}} \\ C_t^{\mathrm{DN}} = p_t^{\mathrm{DN}} \cdot P_t^{\mathrm{DN}} \\ C_t^{\mathrm{tcl}} = p^{\mathrm{indoor}} \left(T_t^{\mathrm{indoor}} - T^{\mathrm{indoorset}}\right) \\ \qquad + p^{\mathrm{wt}} \left(T_t^{\mathrm{wt}} - T^{\mathrm{wtset}}\right) \\ C_{i,t}^{\mathrm{ESS}} = p^{\mathrm{ESS}} (P_{i,t}^{+} + P_{i,t}^{-}) \end{cases} \tag{16}$$

where T is the total number of periods in the scheduling cycle; C_t^{gas} is the fuel cost at time t; p^{gas} is the gas price; C_t^{DN} is the electricity purchase cost at time t; p_t^{DN} is the electricity price at time t; C_t^{tcl} is the uncomfortable cost of temperature control loads at time t; p^{indoor} and p^{wt} is the sensitivity coefficient of indoor temperature and hot water temperature [20]; $T^{\mathrm{indoorset}}$ and T^{wtset} is the indoor temperature and hot water temperature set by user; $C_{i,t}^{\mathrm{ESS}}$ is operating cost of i-th ESS at time t; p^{ESS} is the operating cost factor.

To deal with the multivariate uncertainties of electric vehicles' owners, electricity prices and outdoor temperature in real-time energy management, the objective function should be the maximize expected value of the total benefit in the dispatch period:

$$\max \mathbf{E} \left\{ -\sum_{t=1}^{T} C_t \left(S_t, x_t, w_t\right) \right\} \tag{17}$$

$$\begin{cases} C_t\left(S_t, x_t, w_t\right) = C_t^{\text{gas}} + C_t^{\text{DN}} + C_t^{\text{tcl}} + \sum_{i=1}^{\text{ESS}} C_t^{\text{ESS}} \\ S_t = P_t^{\text{CON}}, E_{i,t,\min}^{\text{EV}}, E_{i,t,\max}^{\text{EV}}, p^{\text{DN}} \\ x_t = P_{i,t}^{\text{ESS,+}}, P_{i,t}^{\text{ESS,-}}, P_{i,t}^{\text{EV}}, h_t^{\text{cchp}}, P_t^{\text{DN}} \\ w_t = w_t^{EV}, w_t^{T}, w_t^{p} \end{cases} \tag{18}$$

where C_t is the total cost at time t; S_t is the state of the system, including load power, the upper and lower boundaries of the EV cluster, and electricity price information; x_t is the decision variable, including the charge and discharge power of ESS, the charge power of EV, the output thermal power of CCHP, and the purchasing power; w_t is the random information, including the newly connected or leaving EVs, changes in outdoor temperature and changes in electricity prices.

III. STOCHASTIC DYNAMIC PROGRAMMING-BASED ONLINE ALGORITHM FOR INTEGRATED ENERGY BUILDINGS

A. THE CONVERSION OF MODEL

The key to real-time energy management of integrated energy buildings is how to deal with various uncertain factors, namely random information w_t, and how to solve the objective function containing the expected value. Therefore, this paper proposes a stochastic dynamic programming-based online algorithm for integrated energy buildings that can adapts to various uncertain factors based on the idea of stochastic dynamic programming.

According to Bellman's optimality principle, the multi-period optimization decision problem can be transformed into a recursive problem, that is, equation (17) can be turned into (19),

$$V_t\left(S_t\right) = \max(C_t\left(S_t, x_t, w_t\right) + \xi \mathbf{E}(V_{t+1}(S_{t+1}|S_t))) \tag{19}$$

That is, for Markov multi-step decision-making process with no aftereffect (the future has nothing to do with the past), the decision-making of each step only depends on the current state and the subsequent system evolution, and has nothing to do with the previous history [21]. In other words, when the decision made at time t makes the formula (19) take the minimum value, it is the optimal strategy. Where $V_t\ (S_t)$ is the value function of the system in state S_t; $V_{t+1}\ (S_{t+1}|\ S_t)$ is the value function at $t+1$ under the premise of system state S_t, which means the impact of the current decision on the subsequent period cost; is the future attenuation factor, whose value is between 0 and 1.

The solution of equation (19) requires the value function of the system state at each time. While solving small-scale problems, we can use the method of recursion from back to front. That is, listing various possible single-stage states and their state value functions. Since each step is screened according to the principle of optimality, the amount of calculation is greatly reduced. However, in the real-time energy management of integrated energy buildings, the coupling constraint of time period between ESS and EV decision variables, and the decision variables and state variables are continuous,

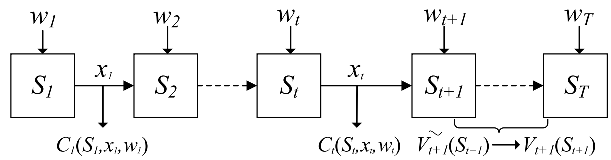


FIGURE 2. The solution diagram of stochastic dynamic programming.

which make the state space increases exponentially with the raise of dimension, so it is not practical to list all states and their value functions. Therefore, we need to use effective methods to estimate the value function in order to get the optimal scheduling strategy as close as possible.

B. THE PROCESSING OF MODEL

Stochastic dynamic programming solves the problem of "Curse of dimensionality" by using approximate value functions trained in a large number of simulation scenarios to replace the value function in equation (19) and adopts the way of recursive solution from the forward to the back. The thought of stochastic dynamic programming can be shown in Figure 2.

When the system is in the state S_t, the approximate value function $\tilde{v}_{t+1}(S_{t+1})$ obtained by the approximate dynamic programming algorithm can replace the value function $v_{t+1}(S_{t+1})$, and the decision variables x_t can be obtained by solving the equation (19) to generate real-time revenue C_t, and the next state S_{t+1} is obtained through the state transition equation and the observed random information.

We define the state transition equation as shown in equation (20),

$$S_t = f\left(S_{t-1}, x_{t-1}, W_t\right) \tag{20}$$

That is, the system state S_t at time t is determined by the state S_{t-1}, the decision made at $t-1$, and the random variable w_t at time t in common. In order to solve the problem, we divide the system state into two stages, the pre-decision state S_t^{x-} and the post-decision state S_t^x. Before making a decision, the system observes changes in random variables and changes its state,

$$S_{\text{t}}^{X-} = S_{\text{t-1}}^{X} + f^{X-}\left(W_t\right) \tag{21}$$

After making the decision, the system, the state of the system changes further,

$$S_{\text{t}}^{X} = S_{\text{t}}^{X-} + f^{X}\left(x_t\right) \tag{22}$$

Take electric vehicles as an example. Before the decision is made, the battery level of each electric vehicle remains unchanged. But after observing the random access of new electric vehicles, the upper and lower boundaries of the energy, and the constraints of the electric vehicles change, and the adjustable capacity of the electric vehicle also changes. After the decision is made, the battery power of each electric vehicle changes according to the decision made, which realizes the transfer of the system state.

After dividing the system state into the pre-decision and post-decision states, equation (19) can be divided into two parts. Therefore, the pre-decision state value function and the post-decision state value function can be obtained,

$$V_t^{X-}\left(S_t^{X-}\right) = \max_{x_t \in X}(C(S_t, x_t) + \xi V_t^X(S_t^X|S_t^{X-})) \quad (23)$$

$$V_t^X\left(S_t^X\right) = E(V_{t+1}^{X-}(S_{t+1}^{X-}|S_t^X)) \quad (24)$$

While the approximate value function is designed, the value function corresponding to each state is updated through an iterative process to obtain the optimal strategy. The idea of iteration is to calculate the optimal decision based on the initial estimate of the value function, and update the value function with the information obtained from the decision. Apply the updated value function to the next iteration process, so that the approximate value function continuously approaches the exact value function. The higher the approximate precision, the closer the decision is to the optimal solution.

$$x_t^n = \arg\max_{x_t \in X}(C(S_t^n, x_t^n) + \tilde{V}_t^{n-1}(S_t^{n-1}|S_{t-1}^{n-1})) \quad (25)$$

where n represents the number of iterations. It can be seen that how to obtain an approximate value function to replace the accurate value function is the key to the problem. There exist two difficulties. One is the design of the approximate value function for real-time scheduling problems, and the other is the iterative update method of the approximate value function.

C. THE SOLUTION OF MODEL

The design methods of approximate value function include linear model method based on basis function, piecewise linear function method and hierarchical clustering method [21]. In this paper, we use the piecewise linear method to construct the approximate value function, and the estimation parameters are updated with the marginal revenue of each piece of virtual storage. At last, the approximation function is obtained by the SPAR method (Successive Projective Approximation Routine).

Use the piecewise linear method to construct the approximate value function, we have

$$x_t^n = \arg\max_{x_t \in X}(C(S_t^n, x_t^n) + \xi \sum_{r=1}^{\beta} v_t^n\,(r, W_t) y_{tr}) \quad (26)$$

Which must satisfy

$$\sum_{r=1}^{\beta} y_{tr} = f^X\,(R_t, x_t)\,, y_{tr} \in [0, \rho]\,, r \in \{1, \cdots, \beta\} \quad (27)$$

where r represents the number of segments, ρ is the length of each segment, and y_{tr} is the amount of resources in each segment. Equation (27) ensures that all resources are added to the resource value after decision. The model assumes that all segments are uniform.

The steps of using the SPAR method to obtain the approximate value function are as follows:

Step 1: Initialize $v_t^n(r, W_t)$, and make $v_t(1, W_t) \geq v_t(2, W_t), \geq \ldots \geq v_t(\beta, W_t)$ to ensure that the slope is monotonically decreasing so that they can meet the concavity; then we use Monte Carlo method to generate N training samples, and each training sample contains the changes of various random quantities in a comprehensive energy building in a day. Make n = 1 and t = 1;

Step 2: Update the system state according to the latest random variables, and use the slope of each segment after the last iteration to solve equation (28), that is

$$x_t^n = \arg\max_{x_t \in X}(C(S_t^n, x_t^n) + \xi \sum_{r=1}^{\beta} v_t^{n-1}\,(r, W_t) y_{tr}) \quad (28)$$

Therefore, we can obtain each decision variable x_t^n, the system state after the decision $S_t^{n,X}$, including the adjustable capacity $R_t^{n,X}$, after the decision, etc.

Step 3: Calculate the temporary value of the slope from the updated sample,

$$g(y) = \begin{cases} (1-\alpha)v_{t-1}^{n-1}(\mathrm{y}) + \alpha \hat{v}_t^n & \mathrm{y} = r \\ v_{t-1}^{n-1}\,(\mathrm{y}) & y \neq r \end{cases} \quad (29)$$

where g is a temporary vector, a is the step size, $\hat{v}_t^n$(y) is the marginal benefit, V_{t-1}^{n-1}(y) is approximate slope.

Step 4: Perform projection operation on the temporary vector to get the approximate slope component of the nth iteration

$$\begin{cases} \min \parallel v_t^n\text{-g} \parallel \\ s.t.\ v_t^n\,(\text{r+1}) \leq v_t^n\,(\text{r}) \end{cases} \quad (30)$$

Step 5: Make $t = t + 1$, return to step 2. when $t > T$, return to step 6.

Step 6: Make $n = n + 1$, return to step 2. The loop ends when $n >$ N.

IV. CASE STUDIES

In order to verify the effectiveness of the proposed model and algorithm, we select an integrated energy building in the south for simulation analysis. Reference [22] can see the electricity price data and CCHP related parameters. The temperature sensitivity coefficients of hot water and rooms are based on [23]. As for uncertainty factors, we mainly consider the impact of EV owner's access and departure time, electricity price forecast, outdoor temperature deviation on the dispatch result. The calculation example is modelled by MATLAB R2018b and GAMS on a computer with an intel(R) Core (TM) i7-7700 CPU, a main frequency of 3.40 GHz and a memory of 8 G.

A. THE SETTING OF SCENE

This paper sets up four different scenarios as shown in Table 1. As for the uncertainty of EV, we mainly consider two cases. One is typical working days, the peaks of

TABLE 1. The setting of scene.

Scene	The law of EV	The law of electricity price
1	typical working day	normal situation
2	typical working day	abnormal situation
3	atypical working day	normal situation
4	atypical working day	abnormal situation

access and departure are mainly concentrated in the day and night. The second is atypical working days, the distribution of access and departure times is more random. As for the uncertainty of the electricity price, one is normal situation where the high peak appears during the day and the low peak appears at night. The second is abnormal situation where the electricity price is distributed according to the sine function and changes more randomly. The calculation example in this paper only simulates and analyses the uncertainty of EV and electricity price. As for other uncertainties, our algorithm is also applicable, but it will not be repeated due to the limitation of the length of the literature.

B. THE RESULT OF SIMULATION

Assuming that at each time t, all the random information is known in advance, so the theoretically optimal solution can be obtained. Compare the optimal solution with the approximate solution obtained by the algorithm proposed in this paper, we can evaluate the performance of the proposed algorithm.

Use Monte Carlo method for sampling to generate 100 training scenarios to train the value function. Each training scenario contains different random information. It can be seen from Figure 3 that in the four scenarios, the algorithm proposed in this paper can quickly approach the optimal solution after a certain number of trainings. Because in each different training scenario, the value function is constantly learning, so as to adapt to the learned environment. Besides, it can use the learned knowledge to make an approximate optimal strategy in the unknown environment in the future. It can be seen from the figure 3 that in the 15th to 100th training scenarios, the error between the approximate solution obtained by the algorithm proposed in this paper and the optimal solution is small, so the algorithm can be considered as convergent. In scenario 4 where the law of EV and electricity prices is more random, the fluctuation of the objective function is greater, because the uncertainty factors change more drastically. But it can be seen that the algorithm in this paper can still maintain the approximate optimal, which means it has better stability.

After the value function converges, put it into online operation. That is, according to the current state of the system and random information, use the approximate value function to solve the single-period optimization problem. Take scene 1 as an example to analyse the output of each unit in the integrated energy building during real-time operation on a certain day, as shown in Figures 4 and 5. The total electrical load of the building is composed of EV load, basic load and cooling load,

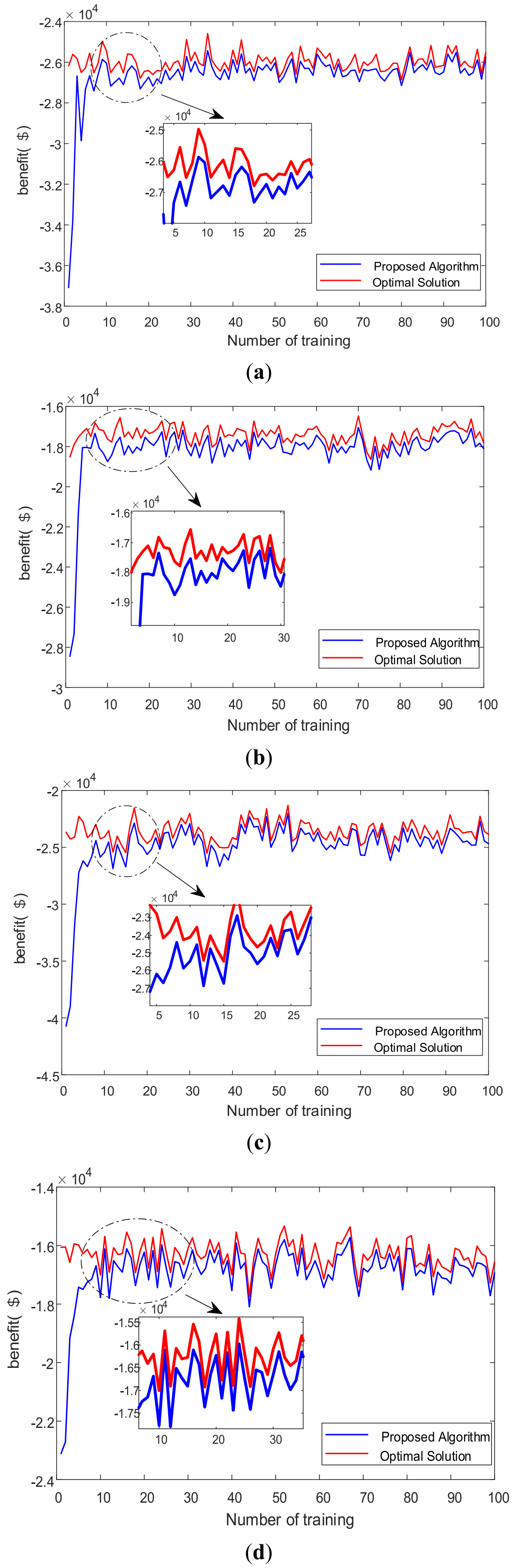


FIGURE 3. Training process: (a) Scene 1 (b) Scene 2 (c) Scene 3 (d) Scene 4.

and it is supplied by the output of ESS, purchased power, and CCHP output. The power output of ESS depends on the level

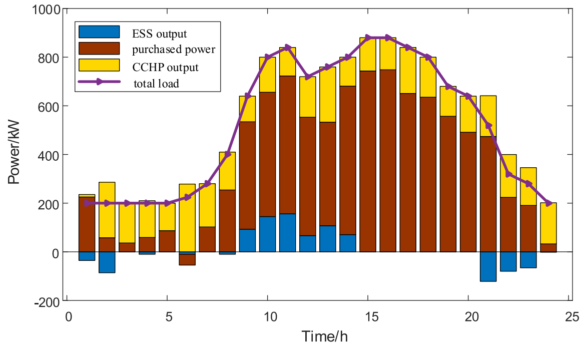


FIGURE 4. The output of each unit.

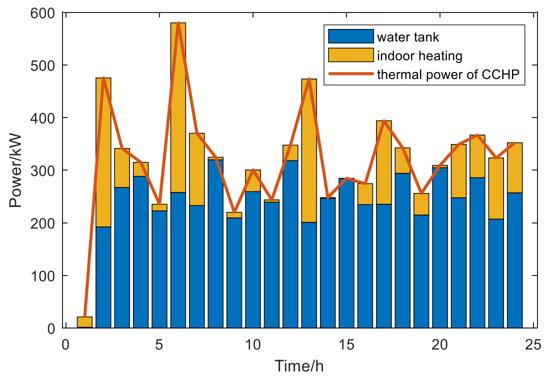


FIGURE 5. The demand of heat load.

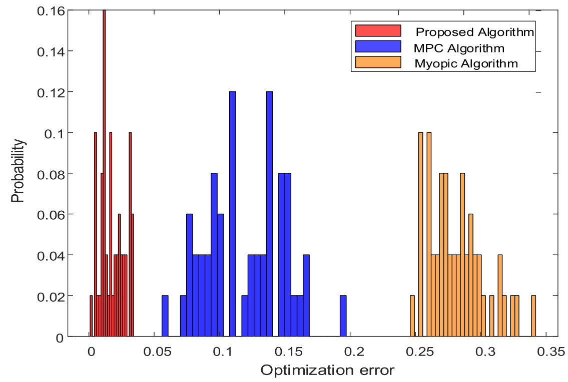


FIGURE 6. Errors of different algorithms.

TABLE 2. Effects of different algorithms.

Algorithm	Average operating cost/ $	Average calculation time/s
Proposed algorithm	$1.69*10^4$	0.192
MPC algorithm	$1.91*10^4$	0.235
Myopic algorithm	$2.13*10^4$	0.293

of electricity price and the size of the load. Since scene 1 is a situation where the electricity price appears high peak during the day and low peak in the morning and evening, it can be seen that the ESS is charged at 22:00∼8:00 when the electricity price is low, and discharged at 9:00∼14:00 when the electricity price is high. Since CCHP works in the ''fixed electricity by heat'' mode, we can see that from 6:00 to 7:00, the heating demand of integrated energy buildings increases, and the thermal power output of CCHP increases, thereby its electrical power output raises. On the basis of meeting its own load demand, the building has sold surplus electricity to the grid, which has improved the operating economy of the building.

C. THE COMPARISON OF ALGORITHMS

In order to verify the effectiveness of the algorithm proposed in this paper, we extract another 100 simulation scenarios to conduct online simulations of the convergent approximation function, and compare them with the MPC algorithm and myopic algorithm. Figure 6 shows the algorithm comparison effect diagram under scene 4. The optimization error is obtained by comparing the result of the algorithm with the theoretical optimum. It can be seen from the figure 6 that the optimization effect of the algorithm proposed in this paper is better than the other two algorithms. Due to the influence of random information, the algorithm in this paper cannot reach the global optimum. However, in various scenarios, the optimization error of this algorithm can be kept within 0.04. Since the MPC algorithm and Myopic algorithm are optimized in a short time window, they cannot use the approximate value function to consider the impact of all subsequent periods, which leads to poor effect. The optimization error of MPC algorithm is between 0.06 and 0.2, while the error of Myopic algorithm is between 0.25 and 0.35.

It can be seen from the table 2 that the algorithm in this paper is more economical in simulation scenarios, which means it has lower average operating cost (the opposite of the objective function of this paper). After the value function is trained to convergence offline, it's put into real-time operation, which only takes 0.192 s on average to get the approximately optimal solution for each period, meeting the demand of real-time operation. In contrast, although MPC algorithm and Myopic algorithm do the calculation in a shorter time window, they take more time in each calculation.

V. CONCLUSION

There are multivariate uncertainties inside integrated energy buildings, including deviations from EV owner behaviors, electricity price forecasts, and outdoor temperature. To address this problem, this paper proposes an algorithm based on stochastic dynamic programming to deal with the impact of uncertainties on online energy management of the integrated energy building, which improve the economics of building operation. The results of the calculation example show:

1) The proposed algorithm makes operators no longer rely on intraday forecast information to make approximate optimal decisions, which greatly reduces the impact of multivariate uncertainties on energy management.

2) The proposed algorithm performs obviously better than MPC and Myopic algorithm in optimization accuracy, which improves the economic benefits of operators.

3) After the algorithm is trained offline by using extensive training samples, it's put into online operation, and the average calculation time can reach 0.192s, which meets the needs of online operation.

How to realize the coordinated dispatch of distribution network and integrated energy buildings at multiple time scales is the focus of the next step.

REFERENCES

[1] L. Cheng, T. Yu, X. Zhang, and B. Yang, "Parallel cyber-physical-social systems based smart energy robotic dispatcher and knowledge automation: Concepts, architectures, and challenges," *IEEE Intell. Syst.*, vol. 34, no. 2, pp. 54–64, Mar. 2019.

[2] C. D. Korkas, S. Baldi, I. Michailidis, and E. B. Kosmatopoulos, "Occupancy-based demand response and thermal comfort optimization in microgrids with renewable energy sources and energy storage," *Appl. Energy*, vol. 163, pp. 93–104, Feb. 2016.

[3] C. D. Korkas, S. Baldi, I. Michailidis, and E. B. Kosmatopoulos, "Grid-connected microgrids: Demand management via distributed control and human-in-the-loop optimization," *Adv. Renew. Energies Power Technol.*, vol. 2, pp. 315–344, Mar. 2018.

[4] P. Wang, H. Xie, and W. Fu, "Design and implementation of the intelligent building energy management and control system," *Process Autom. Instrum.*, vol. 35, pp. 64–68, Mar. 2014.

[5] M. C. Bozchalui, S. A. Hashmi, H. Hassen, C. A. Canizares, and K. Bhattacharya, "Optimal operation of residential energy hubs in smart grids," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1755–1766, Dec. 2012.

[6] D. Setlhaolo, S. Sichilalu, and J. Zhang, "Residential load management in an energy hub with heat pump water heater," *Appl. Energy*, vol. 208, pp. 551–560, Dec. 2017.

[7] M. Ameri and Z. Besharati, "Optimal design and operation of district heating and cooling networks with CCHP systems in a residential complex," *Energy Buildings*, vol. 110, pp. 149–158, Nov. 2015.

[8] Z. Bao, Q. Zhou, Z. Yang, Q. Yang, L. Xu, and T. Wu, "A multi time-scale and multi energy-type coordinated microgrid scheduling solution—Part I: Model and methodology," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2257–2266, Sep. 2015.

[9] Y. Zhang, Q. Ai, and R. Hao, "Economic dispatch of integrated energy system at building level based on chance constrained programming," *Power Syst. Technol.*, vol. 43, no. 1, pp. 108–116, May 2019.

[10] C. Wang, Y. Zhou, J. Wu, J. Wang, Y. Zhang, and D. Wang, "Robust-index method for household load scheduling considering uncertainties of customer behavior," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1806–1818, Jul. 2015.

[11] C. Chen, J. Wang, Y. Heo, and S. Kishore, "MPC-based appliance scheduling for residential building energy management controller," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1401–1410, Sep. 2013.

[12] M. Dan, S. Srinivasan, S. Sundaram, A. Easwaran, and L. Glielmo, "A scenario-based branch-and-bound approach for MES scheduling in urban buildings," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7510–7520, Dec. 2020.

[13] F. Verrilli, S. Srinivasan, G. Gambino, M. Canelli, M. Himanka, C. Del Vecchio, M. Sasso, and L. Glielmo, "Model predictive control-based optimal operations of district heating system with thermal energy storage and flexible loads," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 2, pp. 547–557, Apr. 2017.

[14] P. Yang and A. Nehorai, "Joint optimization of hybrid energy storage and generation capacity with renewable energy," *IEEE Trans. Smart Grid*, vol. 5, no. 4, pp. 1566–1574, Jul. 2014.

[15] J. Li, T. Yu, and Z. Pan, "Real-time stochastic dispatch method for incremental distribution network based on reinforcement learning," *Power Syst. Technol.*, vol. 44, no. 9, pp. 3321–3330, Aug. 2020.

[16] Z. Pan, T. Yu, J. Li, K. Qu, L. Chen, B. Yang, and W. Guo, "Stochastic transactive control for electric vehicle aggregators coordination: A decentralized approximate dynamic programming approach," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4261–4277, Sep. 2020.

[17] N. Liu, Q. Chen, J. Liu, X. Lu, P. Li, J. Lei, and J. Zhang, "A heuristic operation strategy for commercial building microgrids containing EVs and PV system," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2560–2570, Apr. 2015.

[18] Y. Lu, S. Wang, Y. Sun, and C. Yan, "Optimal scheduling of buildings with energy generation and thermal energy storage under dynamic electricity pricing using mixed-integer nonlinear programming," *Appl. Energy*, vol. 147, pp. 49–58, Jun. 2015.

[19] H. Zhang, F. Wen, C. Zhang, J. Meng, G. Lin, and S. Dang, "Operation optimization model of home energy hubs considering comfort level of customers," *Autom. Electr. Power Syst.*, vol. 40, pp. 32–39, Oct. 2016.

[20] Z. Yang and X. Ai, "Distributed optimal scheduling for integrated energy building clusters considering energy sharing," *Power Syst. Technol.*, vol. 44, no. 10, pp. 3769–3778, Jun. 2020.

[21] H. Shuai, J. Fang, X. Ai, Y. Tang, J. Wen, and H. He, "Stochastic optimization of economic dispatch for microgrid based on approximate dynamic programming," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2440–2452, May 2019.

[22] C. Xu, D. Wang, C. Ma, R. Xu, J. Wu, T. Yu, and B. Liu, "Optimization for commercial building energy management of multi-energy fusion," in *Proc. IEEE 8th Annu. Int. Conf. Cyber Technol. Autom., Control, Intell. Syst. (CYBER)*, Jul. 2018, pp. 492–497.

[23] J. Wang and N. Liu, "Distributed optimal operation method of integrated energy system with multi-agents," *Southern Power Syst. Technol.*, vol. 12, no. 3, pp. 98–104 and 115, May 2018.

RENLI CHENG received the B.Eng. degree in electrical power system from the Guangdong University of Technology, Guangzhou, China, in 1993. She is currently working as an Engineer with Shenzhen Power Supply Bureau Company Ltd., China. Her research interests include optimal operation and control of power systems.

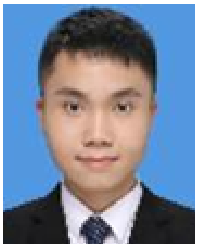
WEIJIE CHENG received the B.Eng. degree in electrical power system from the Huazhong University of Science and Technology, Wuhan, China, in 2005. He is currently working as an Engineer with Shenzhen Power Supply Bureau Company Ltd., China. His research interests include optimal operation and artificial intelligence techniques in planning.

JIE LI received the B.Eng. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2019, where he is currently pursuing the M.Eng. degree in electrical engineering with the College of Electric Power. His research interests include reinforcement learning power systems operation and stochastic optimization.

ZEQI CHEN received the B.Eng. degree in electrical power system from Xi'an Jiaotong University, Xi'an, China, in 2008, and the M.Eng. degree in electrical engineering from Wuhan University, Kunming, China, in 2014. He is currently working as an Engineer with Shenzhen Power Supply Bureau Company Ltd., China. His research interests include optimal operation and artificial intelligence techniques in planning.

JUN SHI received the B.Eng. degree in electrical power system and the M.Eng. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2005 and 2008, respectively. He is currently working as an Engineer with Shenzhen Power Supply Bureau Company Ltd., China. His research interests include optimal operation and grid planning.

ZHENNING PAN received the B.Eng. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2016, where he is currently pursuing the Ph.D. degree in electrical engineering with the College of Electric Power. His major research interests include intelligent optimization and machine learning in operation of power systems.



YUFENG WU received the B.Eng. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2020, where he is currently pursuing the M.Eng. degree in electrical engineering with the College of Electric Power. His research interests include reinforcement learning power systems operation and stochastic optimization.

TAO YU (Member, IEEE) received the B.Eng. degree in electrical power system from Zhejiang University, Hangzhou, China, in 1996, the M.Eng. degree in hydroelectric engineering from Yunnan Polytechnic University, Kunming, China, in 1999, and the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2003. He is currently a Professor with the College of Electric Power, South China University of Technology, Guangzhou, China. His special fields of research interests include nonlinear and coordinated control theory, artificial intelligence techniques in planning, and operation of power systems.