Imitation Learning Based Real-time Decision-making of Microgrid Economic Dispatch Under Multiple Uncertainties

Wei Dong, Fan Zhang, Meng Li, Xiaolun Fang, and Qiang Yang, *Senior Member, IEEE*

Abstract—The intermittency of renewable energy generation, variability of load demand, and stochasticity of market price bring about direct challenges to optimal energy management of microgrids. To cope with these different forms of operation uncertainties, an imitation learning based real-time decision-making solution for microgrid economic dispatch is proposed. In this solution, the optimal dispatch trajectories obtained by solving the optimal problem using historical deterministic operation patterns are demonstrated as the expert samples for imitation learning. To improve the generalization performance of imitation learning and the expressive ability of uncertain variables, a hybrid model combining the unsupervised and supervised learning is utilized. The denoising autoencoder based unsupervised learning model is adopted to enhance the feature extraction of operation patterns. Furthermore, the long short-term memory network based supervised learning model is used to efficiently characterize the mapping between the input space composed of the extracted operation patterns and system state variables and the output space composed of the optimal dispatch trajectories. The numerical simulation results demonstrate that under various operation uncertainties, the operation cost achieved by the proposed solution is close to the minimum theoretical value. Compared with the traditional model predictive control method and basic clone imitation learning method, the operation cost of the proposed solution is reduced by 6.3% and 2.8%, respectively, over a test period of three months.

Index Terms—Energy management, imitation learning, data-driven decision, economic dispatch.

I. Introduction

THERE is a broad consensus that the high proportion of renewable energy generation is the key technology for achieving low-carbon energy supply and solving environmental problems [1]. The microgrid is considered as an energy prosumer for integrating and utilizing different forms of renewable energy sources (RESs), e.g., wind and solar, to meet local demand. However, the internal uncertainties due to the random and intermittent nature of renewable energy generation and the variability of power load pose various power balancing challenges, e.g., voltage rise and frequency deviation, to the reliable and safe operation of microgrids. In addition, real-time pricing mechanisms have been implemented in numerous demonstration areas, introducing significant uncertainty to the external trading environment [2].

In practice, energy storage systems (ESSs) including the vehicle-to-everything mobile ESS [3] are employed to mitigate the impact of such uncertainties and to enhance the system stability margin [4]. In addition, ESSs can be flexibly controlled to shift the output of RESs so as to arbitrage revenues in the energy market while maintaining the power supply-demand balance [5]. Therefore, an efficient energy management system (EMS) considering long-term optimization objectives in microgrid is demanded to realize the multiple economic dispatch decisions in the face of various uncertainties introduced by the renewable energy generation, load demand, and electricity pricing.

In the literature, a number of microgrid economic dispatch solutions have been proposed to cope with the aforementioned multiple uncertainties, covering day-ahead scheduling, intra-day optimization, and real-time dispatch.

The optimal day-ahead scheduling solutions generally require precise predictions of system power production and demand [6]. It is well recognized that the accurate prediction is difficult to accomplish, or even impossible in practice, which is why the stochastic, fuzzy, and robust based methods are used. The uncertainties in stochastic optimization are often considered as random variables with certain standard probabilistic distributions [7], [8], which is difficult to fully describe the uncertain properties of variables. In [9], a scenario-based analysis method is utilized to model the uncertainties in stochastic processes of optimal scheduling in multi-energy microgrids. In [10], the uncertainties in RESs and loads are modeled by the Taguchi method to minimize the ESS life cycle cost, loss of power supply probability, and carbon emissions. The fuzzy optimization takes the uncertain inputs as fuzzy variables to mitigate the effects of prediction inaccuracy [11], [12]. The robust optimization addresses the system uncertainties by evaluating the worse-case scenarios [13], [14]. These day-ahead scheduling solu-

W. Dong, F. Zhang, and M. Li are with the School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China (e-mail: wdong@hdu.edu.cn; zhangfan@hdu.edu.cn; lm_auto@hdu.edu.cn).

X. Fang and Q. Yang (corresponding author) are with the College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: xlf@zju.edu.cn; qyang@zju.edu.cn).

tions can be solved by optimization techniques (e.g., Lagrange multipliers [15], strong duality theory [16], and heuristic algorithms [17]) in an offline fashion. However, it should be highlighted that the day-ahead energy scheduling solutions in the presence of inaccurate prediction may lead to a significant deviation from the optimum of economic dispatch.

Based on the day-ahead energy scheduling, the intra-day optimization strategy can be implemented according to the real-time operation and the latest forecast information for the energy scheduling in the coming hours. These rolling/receding horizon optimization solutions combine offline and online methods to mitigate the problem of variability and uncertainty with their predictive and self-correcting capabilities. Since the latest system states can be updated with more accurate information on intra-day stage, the two-stage optimization or closed-loop model predictive control (MPC) framework has been investigated. In [18], a two-stage robust stochastic programming model for commercial microgrids was applied to maximize the anticipated profits optimally in the day-ahead market, minimizing the imbalance cost by the real-time adjustment of RESs. In [19], a two-stage approach was proposed to integrate day-ahead unit commitment and real-time economic dispatch to deal with high uncertainties in the load demands and renewable generation. In [20], a closed-loop distributed MPC was designed to reduce potential variations of intra-day economic dispatch considering the optimization of behaviors of several participants. In [21], a two-stage stochastic programming approach combined with MPC was suggested for microgrid planning considering economic and environmental objectives. In [22], a two-layer energy dispatch strategy was proposed, in which the pre-layer optimal results can be adjusted by solving a boundary value problem to improve the robustness of prediction errors. However, these intra-day optimization solutions are highly reliant on an explicit forecast or an appropriate modified strategy against the future uncertainties, which may be influenced by incorrect models, online regulations, or prediction horizons.

To further reduce uncertainties and eliminate the effect of prediction errors, the real-time dispatch has received more and more attention. These online solutions could not rely on cumbersome predictions of multiple random variables. This stochastic sequential decision problem is often considered and modeled as a Markov decision process (MDP), which uses Bellman's equation to decompose the temporal dependency and partition the large-scale optimization. However, such a high-dimensional decision space may lead to the "curse of dimensionality" to MDP methodologies. To address these challenges, approximate dynamic programming (ADP) and reinforcement learning (RL) are developed to solve Bellman's equation through value function approximation (VFA) [23] or policy function approximation (PFA) [24]. In [25], an RL method was developed that enables generation resources, distributed storage, and customers to identify the best strategies for energy dispatch and load scheduling without any prior knowledge. In [26], an RL-based online optimal control method was proposed to smooth the charging and discharging profile and suppress the distur-

bance of the hybrid energy storage. In [27], a cooperative RL algorithm with a diffusion approach was proposed for the distributed economic dispatch in real-time to deal with the vast and continuous state spaces. For the dynamic dispatch of battery energy storage system (BESS), [28] proposed an RL method supplemented with Monte-Carlo tree search and domain knowledge expressed as dispatching rules. However, these solutions require sophisticated designs of learning strategies and approximation functions to avoid dimensionality problems arising from the continuous state/action space, complex constraints, and sluggish training. It is challenging to balance the exploration and the exploitation of the reward function so that these solutions typically fall into the local optimization.

To overcome the above limitations of RL, imitation learning (IL) based economic dispatch methods have attracted more and more attention. IL can greatly enhance the efficiency of RL in decision-making by learning from demonstration samples with expert knowledge. IL methods generally include two categories: behavior clone learning (BCL) and inverse reinforcement learning (IRL). BCL methods simulate the expert suggested demonstrations through supervised learning to realize the action decision under the corresponding state. IRL methods adopt a similar structure to RL, but the reward function in IRL is unknown. IRL simulates the optimal reward function by matching it with expert demonstrations. IRL tries to find the underlying intent of the expert policy so that it can provide a better generalization policy for unseen states or environments with slightly different dynamics. Whereas the model parameters of BCL without policy learning process are easier to train and optimize, and BCL is more convenient to deploy and reproduce.

Reference [29] proposed an IRL approach to address building energy management by learning the objective of controller agents with domain experts. In [30], a novel IRL-based IL framework is used to identify the bidding decision objective function of ESSs in coupled multi-market to gain more profits. Reference [31] proposed a behavior cloning based method to address the online optimal power scheduling by mimicking a mixed-integer linear programming (MILP) solver. In [32], a BCL-based IL method was proposed for the pricing strategy of an electricity retailer, which consists of a self-generated expert knowledge mechanism that instructs the agent to simulate given expert policy with generated expert knowledge. Reference [33] proposed an economic dispatch solution for islanding microgrid based on clone IL. This solution used random regression forest model to directly learn the optimal operation trajectories to make the decision-making model have enough intelligence and experience. This solution of EMS can be efficiently deployed in a cloud-edge control architecture to enable real-time dispatch. The main techniques of the literature under different economic dispatch frameworks for microgrids are summarized in Table I.

Compared with RL methods commonly used in a real-time fashion, IL-based economic dispatch methods offer the advantage of fully exploring the pattern distribution in historical data and making more efficient use of high-quality dem-

onstration samples derived from expert experience. Following the IL process, the intelligent decision-making model can be deployed on edge computing platforms for extended periods of time, while minimizing latency and bandwidth requirements. Unlike existing MPC-based solutions, this direct inference approach without the need for iterative optimization saves significant field computing resources and reduces communication delays and congestion [34]. The IL-based economic dispatch framework is designed for future RES-dominated power systems equipped with ubiquitous sensing, thereby necessitating further follow-up research.

TABLE I
SUMMARY OF LITERATURE ON ECONOMIC DISPATCH FRAMEWORKS FOR MICROGRIDS

Framework	Main technique	Reference
Day-ahead scheduling	Stochastic optimization	[7]-[10]
	Fuzzy optimization	[11], [12]
	Robust optimization	[13], [14]
Intra-day optimization	Two-stage approach	[18], [19], [22]
	Rolling optimization	[20], [21]
Real-time dispatch	RL	[23]-[28]
	IL	[29]-[33]

A successful clone IL model that can make an accurate inference to discriminate between different situations requires plentiful labeled training samples with sufficient diversification to maximize the pattern information in the data. However, the high cost of demonstration labeling or the inaccessibility of labeled samples is always the main reason for the "over-fitting" phenomenon in machine learning [35]. An unsupervised learning model, with the ability to automatically learn signatures and dependencies from the raw data without labels, plays a more important role in conjunction with supervised learning.

Thus, based on BCL, a hybrid model combining unsupervised learning and supervised learning is developed in this study to further enhance the learning accuracy and generalization capability of the decision-making solution for economic dispatch. The main ideas of this solution are as follows: the vast amount of historical data on the cloud platform are leveraged to analyze stochastic variables with inherent uncertainties of wind, photovoltaic (PV), load, and real-time price (RTP) through unsupervised learning to obtain the latent representation of the system operation patterns. Then, the decision-making sequences of the economic dispatch for certain historical days are recalculated by modeling the offline optimization problem. Since the optimization problem is solved after the fact and the conditions that have occurred are already known, there are no uncertainties involved, thus allowing for the attainment of an optimal dispatch. Afterward, the supervised learning model is applied to learn, remember, and understand the complex mapping between the optimal dispatch and the corresponding operation patterns. Finally, by utilizing sensing devices to obtain the latest system information, the well-trained model can be deployed to the edge and perform real-time economic dispatch based on

actual operation conditions.

The main contributions can be summarized in two-fold.

1) An IL-based decision-making solution is developed to realize real-time economic dispatch, which substantially reduces the need for the precise forecasting of multiple stochastic variables and the development of sophisticated policies.

2) A hybrid model combining unsupervised learning and supervised learning is utilized to learn the optimal dispatch of different operation patterns using expert demonstrations, which improves the generalization ability of the proposed solution under multiple operation uncertainties.

The remainder of this paper is organized as follows. The system modeling and the formulation of the economic dispatch problem are presented in Section II. Section III presents the proposed IL-based real-time decision-making solution for microgrid economic dispatch. Section IV extensively evaluates the proposed solution and analyzes the numerical findings. Finally, conclusions are drawn in Section V.

II. SYSTEM MODELING AND FORMULATION OF ECONOMIC DISPATCH PROBLEM

A grid-connected microgrid with cloud-edge architecture is examined in this study through the point of common coupling (PCC) with various types of RESs, i.e., PV sources, micro wind turbines (WTs), and BESS, as illustrated in Fig. 1. It is assumed that the RTP is available from the aggregator/retailer and the power can be transferred between the microgrid and power utility. The real-time monitoring information can be obtained by sensors based on Internet of Things (IoT) technology. The system states collected from IoT sensors can be made available to the cloud platform. The economic dispatch model can be trained in the cloud platform using the historical data and then accessible to field edge computing devices for real-time decision-making and local control actions.
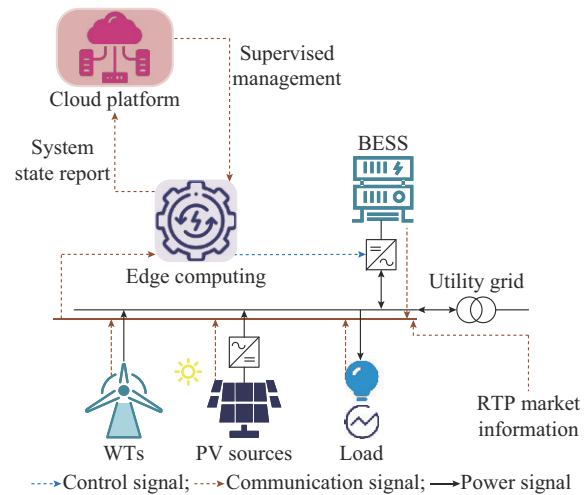

Fig. 1. Illustration of grid-connected microgrid with cloud-edge architecture.

A. System Modeling

In this study, the PV sources, WTs, and load demands are considered non-dispatchable. The ESS is a dispatchable unit

that coordinates the renewable energy generation and demand during the economic dispatch. The decision-making of economic dispatch in the microgrid aims to optimize the use of RESs to reduce imbalances between the power generation and demand, while minimizing operation costs in a real-time pricing environment and improving the lifespan of storage devices. The economic dispatch of a microgrid can be formulated as an optimization problem that considers long-term economic objectives and operation constraints.

Since regional microgrids are often located within a limited geographical area, the power loss is negligible. The power balance constraint is formulated as:

$$p_w(t) + p_p(t) + p_b(t) + p_g(t) = p_l(t) \tag{1}$$

where $p_w(t)$, $p_p(t)$, and $p_l(t)$ are the power of WTs, PV sources, and loads in time slot t, respectively, which are the non-dispatchable variables; $p_g(t)$ is the exchanged power absorbed/injected by/to the utility grid, and when $p_g(t)$ is positive, it means purchasing electricity from the utility grid; and $p_b(t)$ is the dispatched power of the BESS. When BESS is discharging, $p_b(t)$ is positive, and when BESS is charging, $p_b(t)$ is negative.

The total operation cost of a microgrid mainly includes two components: the electricity purchasing cost from the utility grid ($C_g(\cdot)$) and the BESS deterioration cost due to charging and discharging ($C_b(\cdot)$). The utility grid with sufficient capacity can enable the power of the microgrid to be fed back at the same electricity price. Thus, the objective of economic dispatch in a long-term optimization horizon is:

$$\min \sum_{t \in T} \left(C_g(p_g(t)) + C_b(p_b(t), SOH(t)) \right) \tag{2}$$

where $SOH(\cdot)$ is the state of health (SOH) of the BESS; and T is the set of time slots for the long-term objective (always one day) in the global optimization. The itemized costs are shown as:

$$C_g(p_g(t)) = e_g(t) p_g(t) \tag{3}$$

$$C_b(p_b(t), SOH(t)) = \rho_b(SOH(t) - SOH(t+1)) \tag{4}$$

where $e_g(t)$ is the RTP of the electricity of the power grid in time slot t; and ρ_b is the degradation coefficient of the BESS.

The SOH degradation iteration of the BESS caused by charging and discharging cycles is formulated as [27]:

$$SOH(t+1) = SOH(t) - \Delta h(t) \cdot SOH(t) \tag{5}$$

where $\Delta h(t)$ is the degradation factor associated to the change in the state of charge (SOC) of the BESS, which can be calculated as [27]:

$$\Delta h(t) = \alpha_h \left((\Delta SOC(t))^{\beta_h} + \eta_h \right)^{-1} \tag{6}$$

where α_h, β_h, and η_h are the degradation parameters determined by the BESS characteristics from empirical tests.

The SOC change of the BESS, i.e., $\Delta SOC(t)$, is determined by the charging or discharging power:

$$\Delta SOC(t) = \begin{cases} c_b p_b(t) \Delta t / E_b & p_b(t) \le 0 \\ p_b(t) \Delta t / (d_b E_b) & p_b(t) > 0 \end{cases} \tag{7}$$

where c_b and d_b are the efficiency coefficients of charging and discharging, respectively; Δt is the time interval; and E_b is the capacity of the BESS. SOC is restricted to [0.2, 0.8] to prevent BESS deterioration caused by deep charging and discharging.

The dispatched power output constraint of the BESS satisfies:

$$p_b^{\min} < p_b(t) < p_b^{\max} \tag{8}$$

where $p_b^{\min}$ and $p_b^{\max}$ are the lower and upper limits of the dispatched power of the BESS, respectively.

B. Formulation of Economic Dispatch Problem

Once the renewable energy generation, demand, and electricity price are known, the economic dispatch can be formulated as a deterministic optimization problem. By solving this problem, the optimal dispatch over a day in different operation patterns from the historical data can be obtained. Different types of optimization tools can be used to solve this problem, such as commercial solvers Gurobi and CPLEX. For the optimization problem established in Section II-A, heuristic algorithms can be a powerful approach to address such non-convex optimization problem. The solved results can be evaluated by the expert experience to obtain the optimal dispatch decision as close as possible to the optimal solution. Among the heuristic algorithms, particle swarm optimization (PSO) algorithm is considered efficient with a minimal implementation complexity [36], [37], which is extensively utilized to solve different engineering optimization problems. In this work, the PSO algorithm [38] is utilized to solve this optimal economic dispatch problem with continuous decision variables.

III. PROPOSED IL-BASED REAL-TIME DECISION-MAKING SOLUTION FOR MICROGRID ECONOMIC DISPATCH

In this paper, a hybrid model combining unsupervised learning and supervised learning is proposed to construct the mapping relationship between complex operation patterns and optimal dispatch decisions considering multiple uncertain inputs and a real-time operation environment.

First, through the unsupervised learning model, the hidden representations of the time-series observations are extracted to reveal the potential knowledge in a variety of operation patterns. These observations are uncontrollable stochastic variables over a period, including wind power, PV power, load demand, and RTP. The matrix of the observed stochastic variables is formulated as:

$$X_o = \begin{bmatrix} p_w(t-\tau+1) & p_w(t-\tau+2) & \dots & p_w(t) \\ p_p(t-\tau+1) & p_p(t-\tau+2) & \dots & p_p(t) \\ p_l(t-\tau+1) & p_l(t-\tau+2) & \dots & p_l(t) \\ e_g(t-\tau+1) & e_g(t-\tau+2) & \dots & e_g(t) \end{bmatrix} \tag{9}$$

where τ is the length of the time series, which indicates the perception range of operation patterns.

Next, the supervised learning model is applied to memorize and learn the sophisticated inference from input space (constructed by the extracted features through the unsuper-

vised learning model and the matrix of system state variables shown in (10)) to output space (labeled by the optimal dispatch decision at the corresponding time).

$$X_s = \begin{bmatrix} SOC(t-\tau+1) & SOC(t-\tau+2) & \dots & SOC(t) \\ p_g(t-\tau+1) & p_g(t-\tau+2) & \dots & p_g(t) \end{bmatrix} \quad (10)$$

The framework of the proposed solution is shown in Fig. 2. In the offline training process, labels of the supervised learning samples are given by the results of the deterministic offline optimization problem solving by the heuristic algorithm in Section II-B. After the unsupervised and supervised training processes are completed in the cloud platform, the proposed hybrid model can be implemented in field edge computing devices to make the economic dispatch in a real-time environment. Because the model parameters have been learned and trained, the computational complexity requirements in real time can be met during the inference process. The proposed hybrid model has a good generalization capability, which can guarantee the economic and effective dispatch decisions under the new operation pattern in a real-time fashion.
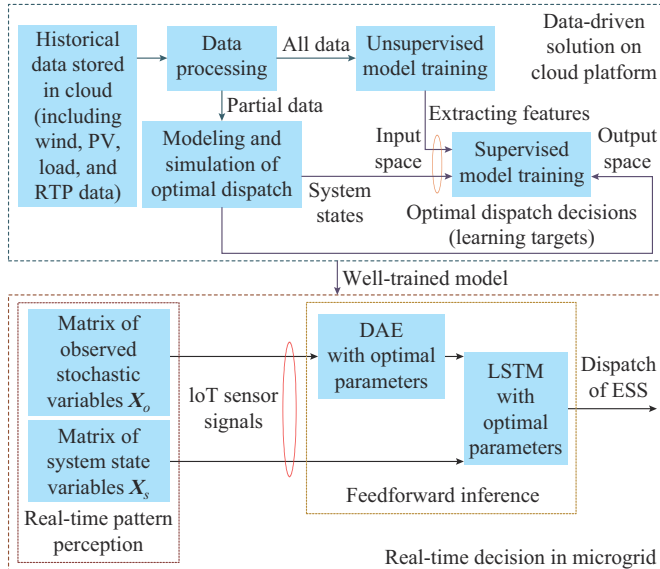

Fig. 2.   Framework of proposed solution.

It is worth noting that the training set can be supplemented by the scenario generation method when the sample data are insufficient. For example, [39] introduced a data-driven technique with interpretability for the scenario generation using controllable generative adversarial networks, which can generate novel and distinct scenarios by capturing the intrinsic characteristics of historical data. In addition, the proposed hybrid model is learnable and self-adaptive, allowing the networks to be updated with novel operation patterns through a rolling training process.

A. Denoising Autoencoder (DAE) Based Unsupervised Learning Model

The autoencoder is an unsupervised learning method that can explicitly learn the important hidden representations on the manifold [40]. The autoencoder model can automatically learn the parameters and extract the compact and robust fea-

tures by minimizing the reconstruction error through the encoder-decoder route. In this study, the DAE network [41] is adopted as the unsupervised learning feature extraction network. The DAE network can learn the observations with both respective time-series characteristics and the correlation characteristics among different variables. The matrix of extracted features can be expressed as:

$$\hat{X}_o = DAE(X_o) \quad (11)$$

where $DAE(\cdot)$ represents the encoding-decoding process using the DAE network.

The DAE can recover the original data $x \in \mathbf{R}^D$ from an encoded representation on the manifold $h = f_\theta(\tilde{x}) \in \mathbf{R}^N$ of the corrupted input data $\tilde{x} \in \mathbf{R}^D$ via a decoding function $g_{\theta'}(h)$. D is the original space dimension, and N is the encoding space dimension. The DAE learns the reconstruction distribution from the training data pairs $(x, \tilde{x})$ through the following process [41].

1) Perturbation process $q(x|\tilde{x})$ adds stochastic noise into the original data x to generate a corrupted input data $\tilde{x}$.

2) Encoding function $f_\theta(\tilde{x}): \tilde{x} \in \mathbf{R}^D \mapsto h \in \mathbf{R}^N$ generates a hidden representation of the input data.

3) Decoding function $g_{\theta'}(h): h \in \mathbf{R}^N \mapsto \hat{x} \in \mathbf{R}^D$ reconstructs the input data from the encoded representation h.

4) Loss metric $L(x, \hat{x})$ can measure the dissimilarity between the original data and the reconstructed output.

The encoded representation h is generated from a corrupted input data $\tilde{x}$ with perturbations, which necessitates learning a sufficiently clever mapping on the manifold to extract useful features for denoising.

Generally, a conditional probabilistic distribution $q(x|\tilde{x})$ is considered to independently perturb each dimension of the input data, i.e., $q(x|\tilde{x}) = \prod_{i=1}^{D} q(x_i|\tilde{x}_i)$.

In the encoding process, the corrupted input data $\tilde{x} \in \mathbf{R}^D$ are transformed to a encoded representation $h \in \mathbf{R}^N$ as [41]:

$$h = f_\theta(\tilde{x}) = s(W\tilde{x} + b) \quad (12)$$

where $W \in \mathbf{R}^{N \times D}$ is the weight coefficient matrix; $b \in \mathbf{R}^N$ is the hidden bias vector; $\theta = (W, b)$; and $s(\cdot)$ is the non-linear activation function.

Then, the hidden representation h is reconstructed to $\hat{x} \in \mathbf{R}^D$ by decoding function as [41]:

$$\hat{x} = g_{\theta'}(h) = s'(W^T h + c) \quad (13)$$

where $c \in \mathbf{R}^D$ is the input bias vector; $\theta' = (W^T, c)$; and $s'(\cdot)$ is the non-linear mapping function at the decoder. The parameters θ and θ' of the encoder and decoder functions are trained by minimizing the reconstruction error, measured by the loss metric $L(x, \hat{x})$.

B. Long Short-term Memory (LSTM) Based Supervised Learning Model

In this study, the supervised learning model is adopted to identify the complicated mapping between the input features from time sequences and the output dispatch decisions. The LSTM neural network [42] is an advanced approach to learn the long-term dependency and relevancy via the exquisite

gate structure in the "memory" blocks. It can use simultaneously the current environment information and the inherent system operation tendency. As a result, the LSTM is well suited for this dynamic decision inference of the non-linear time-varying system by analyzing and understanding the hidden patterns of operation conditions with multi-attribute time-series inputs. In this task, the input space X_{input} consists of the matrix of extracted features $\hat{X}_o$ through the unsupervised learning model and the matrix of system state variables X_s, as shown in (14).

$$X_{input} = \begin{bmatrix} \hat{X}_o \\ X_s \end{bmatrix} = \begin{bmatrix} X_{t-\tau+1} & X_{t-\tau+2} & \cdots & X_t \end{bmatrix} \quad (14)$$

where X_t denotes the input matrix of the LSTM neural network in time slot t, which includes time-series variables with different features.

The structure of the LSTM neural network is shown in Fig. 3, which includes the input layer, hidden layer, and output layer. Each "memory" block in the hidden layer consists of three gates, i.e., an input gate, a forget gate, and an output gate, which comprise a full connection (FC) layer with a sigmoid function and an element-wise product [43]. These gates are designed to regulate the information flow into or out of their "memory" blocks. The information flow of long-term memory c_t and short-term memory h_t controlled by these gates is shown in Fig. 3.
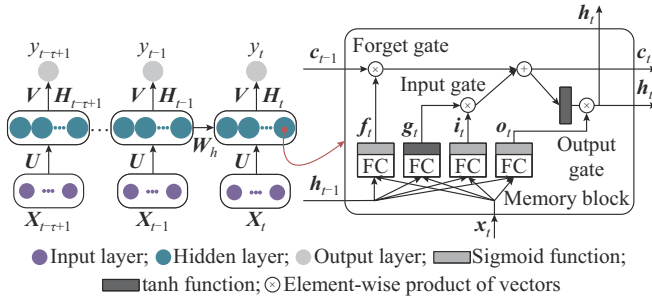

Fig. 3. Structure of LSTM neural network.

The process of forwarding propagation can be expressed as [42]:

$$H_t = f\left(UX_t + W_h H_{t-1}\right) \quad (15)$$

$$y_t = VH_t \quad (16)$$

where H_t is the state of hidden layer in time slot t; $f(\cdot)$ is the ReLU activation function; U and V are the weights between input/hidden layer and the hidden/output layer, respectively; W_h is the weight between the current hidden layer and the hidden layer in the next time slot; and y_t is the inference output of the LSTM neural network in time slot t. In this study, y_t represents the decision variable of the BESS dispatch $p_b(t+1)$ in the next time slot.

The input gate controls which parts of the new information are added and stored in the long-term memory state. The value of the input gate in time slot t i_t can be expressed as [43]:

$$i_t = \text{sig}\left(W_{xi}^T x_t + W_{hi}^T h_{t-1} + b_i\right) \quad (17)$$

where x_t is the input information of "memory" block in time slot t; W_{xi} is the weight between the input layer and the input gate; W_{hi} is the weight between the state of short-term memory in the previous time slot and the input gate; b_i is the bias vector of the input gate; and $\text{sig}(\cdot)$ is the sigmoid activation function.

The forget gate controls which long-term memory state should be dropped. The value of forget gate in time slot t f_t can be expressed as [43]:

$$f_t = \text{sig}\left(W_{xf}^T x_t + W_{hf}^T h_{t-1} + b_f\right) \quad (18)$$

where W_{xf} is the weight between the input layer and the forget gate; W_{hf} is the weight between the state of short-term memory in the previous time slot and the forget gate; and b_f is the bias vector of the forget gate.

The output gate controls which long-term memory state should be read and output in this time slot. The value of output gate in time slot t o_t can be expressed as [43]:

$$o_t = \text{sig}\left(W_{xo}^T x_t + W_{ho}^T h_{t-1} + b_o\right) \quad (19)$$

where W_{xo} is the weight between the input layer and the output gate; W_{ho} is the weight between the state of a short-term memory in the previous time slot and the output gate; and b_o is the bias vector of the output gate.

The output of c_t and h_t can be expressed as [43]:

$$g_t = \tanh\left(W_{xg}^T x_t + W_{hg}^T h_{t-1} + b_g\right) \quad (20)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes g_t \quad (21)$$

$$h_t = o_t \otimes \tanh\left(c_t\right) \quad (22)$$

where W_{xg} is the weight between the input layer and the main layer of the memory block; W_{hg} is the state of a short-term memory in the previous time slot and the main layer of the memory block; b_g is the bias vector; and $\otimes$ is the element-wise product of the vectors.

IV. PERFORMANCE ASSESSMENT AND NUMERICAL RESULT

A. Simulation Setup

The microgrid shown in Fig. 1 is used in this study to evaluate the performance of the proposed solution. WTs, PV sources, and loads have rated capacities of 0.6 MW, 1 MW, and 1.2 MW, respectively. The capacity of BESS E_b is 5 MWh. The charging and discharging efficiencies are both 0.9. SOC is confined to a range of 20% to 80%. The degradation coefficients of BESS are $\rho_b = 100$, $\alpha_h = 0.001$, $\beta_h = -2$, and $\eta_h = 0$ based on the empirical curve-fitting results suggested in [27].

For different microgrids, the optimal dispatching trajectories for IL is derived from the historical operation patterns of corresponding microgrid. Therefore, the proposed solution can be applied to the economic dispatch of various grid-connected microgrids in real-time electricity market environment.

The renewable energy generation and load profiles used in the simulations are taken from a microgrid testbed in 2015 [33]. WTs, PV sources, loads, and RTPs are simulated using real-time data with a 15-min resolution. The data gathered

from January 1 to August 31 (8 months) are used to construct training samples for the unsupervised learning model. The data from June 1 to August 31 (3 months) are substituted into the economic dispatch model (described in Section II) to obtain the optimal dispatch trajectories as the labels of training samples for supervised learning model. The data from September 1 to November 30 (3 months) are utilized as test data to evaluate performance. Figure 4 shows the pattern variation profiles of renewable energy generations, loads, and RTPs in the training set, which encompasses a broad variety of operation conditions with uncertainties. μ represents the average value of the profiles, and σ represents the standard deviation.
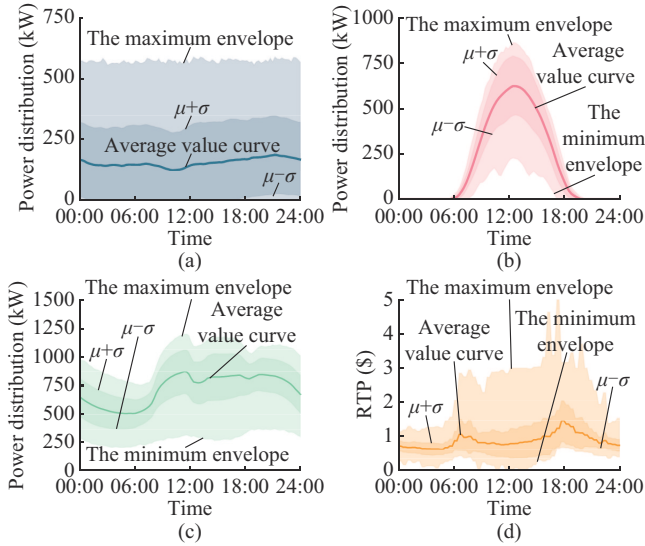

Fig. 4. Pattern variation profiles of renewable energy generations, and RTPs in training dataset. (a) WTs. (b) PV sources. (c) Loads. (d) RTPs.

In this study, the number of hidden layers of the LSTM neural network is set to be 2, and each hidden layer consists of 50 cell blocks. The number of layers of the DAE network is set to be 3. Then, the two networks are both trained using the backpropagation through the Adam algorithm with the loss function of *RMSE* [44].

In the simulation experiment, this study considers that the sensors used to monitor real-time status of the system in practice are reliable enough, so the impact of the field monitoring errors on the solutions is not considered.

B. Performance of Real-time Dispatch

To illustrate the performance of the proposed solution, two operation scenarios are selected: scenario 1 represents the normal operation scenario; while scenario 2 represents the worst operation scenario. In each scenario, the results calculated by the proposed solution and the optimal dispatch without uncertainties are compared, which are shown in Figs. 5 and 6.

The result of Fig. 5(a) intuitively supports that the proposed solution can economically dispatch the BESS. During periods characterized by lower electricity prices, the purchasing power from the utility grid is strategically allocated. When electricity prices increase, the microgrid adjusts its op-

erations intelligently, enabling power delivery to the utility grid. This dynamic behavior showcases the capacity of the microgrid to generate additional revenue by exporting surplus power. Overall, the results demonstrate that in an RTP environment, the proposed solution allows for the full utilization of RESs in the microgrid to meet the demand while effectively reducing electricity purchasing costs.
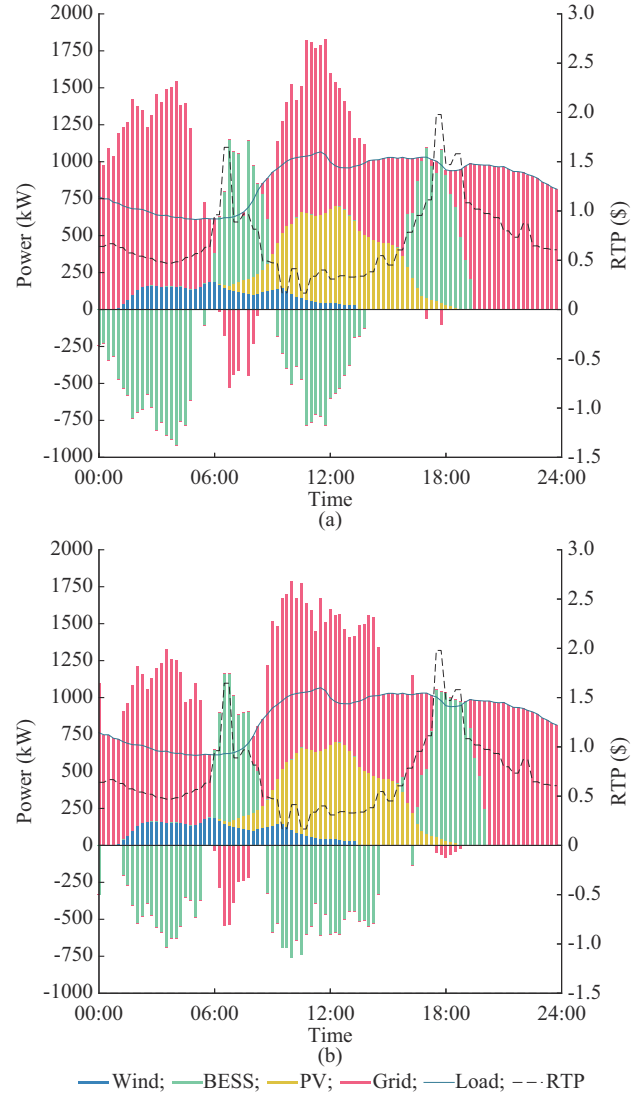

Fig. 5. Results calculated by proposed solution and optimal dispatch without uncertainties in scenario 1. (a) Results of proposed solution. (b) Results of optimal dispatch without uncertainties.

Compared with the results of the optimal dispatch without uncertainties given in Fig. 5(b), the results of the proposed solution considering operation uncertainties show a similar dispatch arrangement on the same test day. This finding demonstrates that the proposed solution, characterized by its strong generalization capability, is capable of making near-optimal decisions when confronted with new operation patterns. By using IL from the expert experiences, the proposed solution can make real-time decisions based on varying operation conditions, thus achieving a dispatching trajectory with a global optimization mindset.
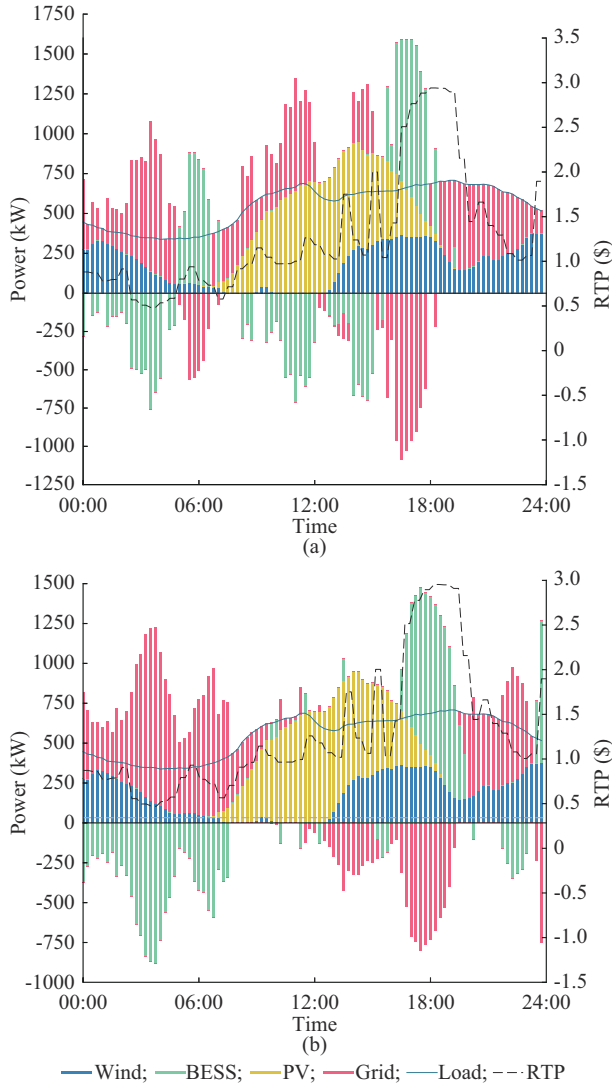
Fig. 6. Results calculated by proposed solution and optimal dispatch without uncertainties in scenario 2. (a) Results of proposed solution. (b) Results of optimal dispatch without uncertainties.

In scenario 2, as illustrated by Fig. 6, the inference trajectory significantly deviates from the optimal dispatch results due to limitations in the generalization ability of the machine learning model. In this scenario, the BESS charging and discharging time sequence decision made by the proposed solution does not match the real-time electricity market, resulting in the need to purchase power from the utility grid to meet the load demand even when the electricity price is high. Furthermore, the frequent charging and discharging of the BESS contribute to additional battery degradation costs. Despite the overall poor economic performance observed in this scenario, the proposed solution still enables the complete utilization of RESs, ensuring the stable operation of the system. These results demonstrate the adaptability and stability of the proposed solution, which enables microgrid economic dispatch without relying on predictive information.

C. Benchmarks and Computational Complexity

To assess and compare the computational complexity and

the economic performance of the proposed solution, three different solutions are used as the benchmarks. The benchmark solutions are described as follows.

1) Solution 1: day-ahead stochastic scheduling [7], which is a typical stochastic optimization scheduling using the prediction information to optimize the scheduling of one day ahead. Various stochastic variables are predicted by the LSTM neural network, where the training data set covers the period from January to August. The off-line optimization problem is also solved by PSO.

2) Solution 2: MPC-based rolling optimization, which can be summarized as follows [45]. The system operation data are forecasted over a time slot (i.e., predictive window) and the optimization can be carried out based on the forecasted data. Then, the obtained optimization result of the first time slot in rolling window is adopted. In such a manner, the predictive window moves forward with one time slot and the previous steps are repeated in the next time slot.

3) Solution 3: basic clone IL-based dispatch, which only uses the supervised learning model to verify the improvement of the unsupervised learning model on the proposed solution, in which the input of the LSTM neural network is the original feature not extracted by the DAE network.

Each solution involves either an offline process or an online process during execution. The computation of the offline process can be computed in the cloud platform. Online process is generally performed on the edge device and generally needs to be performed once per decision period. The regulation resolution of the simulation is 15 min, resulting in an online execution frequency of 96 times per day. All solutions are implemented using Python 3.7 on a computer with a 3.00 GHz Intel Core i5-7400U CPU, Nvidia GTX 1650 GPU and 8 GB RAM. The computational complexity analysis of different solutions is shown in Table II.

TABLE II
COMPUTATIONAL COMPLEXITY ANALYSIS OF DIFFERENT SOLUTIONS

Solution	Computational process	Execution time (s)
Solution 1	Stochastic variable forecasting (offline)	1.0
	Optimization solving (offline)	60.0
Solution 2	Stochastic variable forecasting (online)	1.0
	Rolling optimization (online)	60.0
Solution 3	Optimal scheduling sample optimization solving (offline)	5400.0
	Supervised learning training (offline)	1200.0
	Real-time dispatch (online)	<0.1
Proposed	Optimal scheduling sample optimization solving (offline)	5400.0
	Unsupervised learning training (offline)	720.0
	Supervised learning training (offline)	1200.0
	Real-time dispatch (online)	<0.1

The main computational complexity of the proposed solution lies in the acquisition of expert samples, which needs to optimize the optimal scheduling trajectories of historical days. Another part that requires computational cost is the training of the learning model. These complex calculations can be performed through the offline process. Economic dispatch decisions are made using the real-time information ob-

tained through the online process. At this stage, the learning model only needs to perform forward inference, which incurs relatively low computational costs. This makes it highly suitable for meeting real-time computing requirements.

Compared with solution 1, the proposed solution needs to solve more optimization problems during the offline process to obtain IL sample trajectories. The computational cost of the proposed solution increases linearly with the number of samples. In contrast, solution 1 only needs to solve one optimization problem and does not require additional online computational processes during the day. The computational cost of the proposed solution during a single offline process is much higher than that of solution 1. However, the proposed solution can deploy the decision model in a long term (in this simulation, for three months) using well-trained model parameters. Therefore, the frequency of offline process for updating model parameters in the proposed solution can be very low, whereas the solution 1 needs to execute the offline process every day.

Compared with solution 2, the proposed solution requires both offline optimization and online decision-making. Solution 2 requires continuous optimization for each control interval, with the computational cost primarily incurred during the online process. Each optimization task needs to be completed within a short period, which necessitates edge devices to possess adequate computational capability. In contrast, the proposed solution only requires forward inference during the online process of the decision model, which incurs relatively low computational costs. This makes it easier to meet real-time computing requirements.

Compared with solution 3, the proposed solution introduces the training process of an unsupervised learning model. Since the unsupervised learning model does not need to obtain expert trajectories through optimization as supervised learning samples, the offline computational cost only slightly increases during the training of unsupervised learning model.

D. Economic Evaluation and Comparison

The operation costs in each testing month obtained from the proposed solution and three benchmark solutions are shown in Fig. 7. The figure shows the theoretical lower (from optimal dispatch without uncertainties) and upper (without dispatch strategy of BESS) boundaries of the operation costs, which can reflect the economic cost-saving potential of the system.

The result presented in Fig. 7 clearly demonstrates that both the proposed solution and solution 3 consistently outperform solutions 1 and 2 throughout the three-month test period. This indicates the effectiveness of the IM technology in achieving optimal decisions for economic dispatch based on real-time input information. Moreover, the results achieved by the proposed solution exhibit the highest level of performance, closely resembling the ideal situation. This suggests that incorporating a combined unsupervised learning model can significantly enhance the system performance by capturing a deeper understanding of unlabeled patterns. These findings highlight the potential of leveraging advanced machine learning techniques to improve the decision-making process for economic dispatch, leading to more efficient operations in real-time fashion.
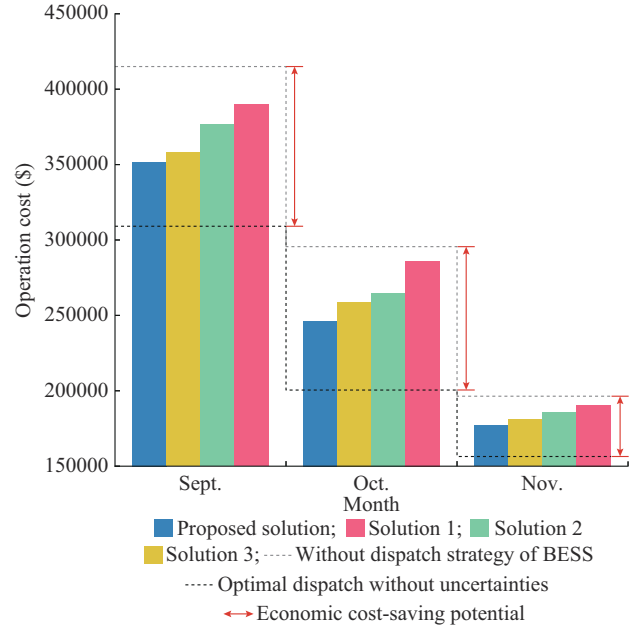

Fig. 7. Operation costs of different solutions over tested months.

As the time progresses from September to November, a noticeable trend emerges in which the cost savings for all solutions consistently decrease. The reason for this could be the consideration of a longer time period between the training data and the test month, as renewable energy generation patterns are more similar within the same season. This finding suggests that while the models may perform well initially, their effectiveness may gradually decline over time due to factors such as seasonal variations and evolving patterns of energy generations and loads. Thus, by rolling update of the model training and incorporating newly collected pattern data, the performance of the solution can be maintained and ensured.

The performance evaluation is carried out for the proposed solution, and the numerical results in terms of the average cost against the benchmark solutions are presented in Table III. Among them, inevitable cost refers to the part exceeding the cost of optimal dispatch. For the total cost that includes the BESS deterioration cost C_b and the electricity purchasing cost C_g, solution 1 performs the worst as it depends on the accuracy of prediction. Solution 2 performs better than solution 1. However, its optimization process in each rolling window requires the controller to have sufficient computing ability. To improve the prediction accuracy during the MPC process, additional costs for purchasing additional numerical weather prediction (NWP) information are needed in practice.

According to the economic evaluation results, IL-based solutions are more competitive under conditions of the same available data. The proposed solution can achieve the greatest cost savings compared with other solutions during all test months. Table IV shows the percentage of operation cost savings of the proposed solution compared with the other three benchmark solutions during each test month. Compared with solutions 1, 2, and 3, the total operation cost of the proposed

solution is reduced by 10.5%, 6.3%, and 2.8%, respectively, during all test months.

TABLE III
COMPARISON OF AVERAGE COSTS OF DIFFERENT SOLUTIONS DURING TEST MONTHS

Test month	Solution	C_b ($)	C_g ($)	$C_b + C_g$ ($)	Inevitable cost ($)
Sept.	Solution 1	341.1	12653.5	12994.6	2691.0
	Solution 2	263.7	12290.0	12553.7	2250.1
	Solution 3	215.5	11717.3	11932.8	1629.2
	Proposed	247.8	11477.4	11725.2	1421.6
Oct.	Solution 1	328.8	8902.7	9231.5	2764.1
	Solution 2	222.7	8318.6	8541.3	2073.9
	Solution 3	165.5	8181.9	8347.4	1880.0
	Proposed	194.7	7748.7	7943.4	1476.0
Nov.	Solution 1	306.8	6045.1	6351.9	1137.7
	Solution 2	207.5	5988.7	6196.2	982.0
	Solution 3	164.6	5866.0	6030.6	816.4
	Proposed	173.5	5731.5	5905.0	690.8

TABLE IV
PERCENTAGE OF OPERATION COST SAVINGS OF PROPOSED SOLUTION COMPARED WITH BENCHMARK SOLUTIONS DURING TEST MONTHS

Test month	Percentage of operation cost savings (%)		
	Solution 1	Solution 2	Solution 3
Sept.	9.8	6.6	1.7
Oct.	14.0	7.0	4.8
Nov.	7.0	4.7	2.1
Average	10.5	6.3	2.8

For the proposed solution, the optimization results of economic dispatch under deterministic conditions are computed first, and then the machine learning model is used to learn the non-linear complex mapping between input patterns and optimal dispatch results in high-dimensional space. The generalization errors are manifested as deviations of inference results from the optimal dispatch trajectories in the new patterns. For the stochastic optimization framework based on prediction results, the cumulative prediction errors of multiple random variables cause the day-ahead stochastic optimization in day-ahead and the deterministic optimization under actual conditions to be inconsistent in the optimal solution space. Besides, the high-dimensional non-convex optimization solution may be easily affected by the multi-saddle points, which leads to suboptimal solutions [46]. Therefore, the proposed solution can more easily achieve the optimal solution for the economic dispatch problem. The numerical results also verify the above conclusions.

V. CONCLUSION

This paper proposes an IL-based decision-making solution to realize microgrid economic dispatch in a real-time fashion. The proposed solution is capable of effectively addressing the economic dispatch problem with high operation uncertainties caused by the intermittency of renewable energy generation and the stochasticity in market prices and loads. By learning the optimal dispatch of the historical operation patterns in a data-driven way, the proposed solution with good generalization performance can make intelligent decisions close to the optimal dispatch. The proposed solution is easy to deploy in practice and suitable for the cloud-edge collaborative communication and computing architecture of the future microgrid.

The proposed solution is evaluated through simulation tests subject to various uncertainties. Compared with the benchmark solutions of day-ahead stochastic optimization, MPC-based rolling optimization, and basic clone IL-based dispatch, the numerical results demonstrate that the total operation cost of the proposed solution is reduced by 10.5%, 6.3%, and 2.8%, respectively, for all the test months.

REFERENCES

[1] S. Eslami, Y. Noorollahi, M. Marzband *et al.*, "District heating planning with focus on solar energy and heat pump using GIS and the supervised learning method: case study of Gaziantep, Turkey," *Energy Conversion and Management*, vol. 269, p. 116131, Oct. 2022.

[2] Z. Wu, J. Wang, H. Zhong *et al.*, "Sharing economy in local energy markets," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 3, pp. 714-726, May 2023.

[3] S. Islam, A. Iqbal, M. Marzband *et al.*, "State-of-the-art vehicle-to-everything mode of operation of electric vehicles and its future perspectives," *Renewable and Sustainable Energy Reviews*, vol. 166, p. 112574, Sept. 2022.

[4] D. Sadeghi, N. Amiri, M. Marzband *et al.*, "Optimal sizing of hybrid renewable energy systems by considering power sharing and electric vehicles," *International Journal of Energy Research*, vol. 46, no. 6, pp. 8288-8312, May 2022.

[5] A. Bharatee, P. K. Ray, and A. Ghosh, "A Power management scheme for grid-connected PV integrated with hybrid energy storage system," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 4, pp. 954-963, Jul. 2022.

[6] H. Shuai, J. Fang, X. Ai *et al.*, "Stochastic optimization of economic dispatch for microgrid based on approximate dynamic programming," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2440-2452, May 2019.

[7] D. Prudhviraj, P. B. S. Kiran, and N. M. Pindoriya, "Stochastic energy management of microgrid with nodal pricing," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 1, pp. 102-110, Jan. 2020.

[8] F. Conte, S. Massucco, M. Saviozzi *et al.*, "A stochastic optimization method for planning and real-time control of integrated PV-storage systems: design and experimental validation," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1188-1197, Jul. 2018.

[9] S. E. Ahmadi, M. Marzband, A. Ikpehai *et al.*, "Optimal stochastic scheduling of plug-in electric vehicles as mobile energy storage systems for resilience enhancement of multi-agent multi-energy networked microgrids," *Journal of Energy Storage*, vol. 55, p. 105566, Nov. 2022.

[10] D. Sadeghi, S. E. Ahmadi, N. Amiri *et al.*, "Designing, optimizing and comparing distributed generation technologies as a substitute system for reducing life cycle costs, CO_2 emissions, and power losses in residential buildings," *Energy*, vol. 253, p. 123947, Aug. 2022.

[11] M. Moafi, R. R. Ardeshiri, M. W. Mudiyanselage *et al.*, "Optimal coalition formation and maximum profit allocation for distributed energy resources in smart grids based on cooperative game theory," *International Journal of Electrical Power & Energy Systems*, vol. 144, p. 108492, Jan. 2023.

[12] W. Dong, Q. Yang, X. Fang *et al.*, "Adaptive optimal fuzzy logic based energy management in multi-energy microgrid considering operational uncertainties," *Applied Soft Computing*, vol. 98, p. 106882, Jan. 2021.

[13] J. Zhang, M. Cui, Y. He *et al.*, "Multi-period two-stage robust optimization of radial distribution system with cables considering time-of-use price," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 1, pp. 312-323, Jan. 2023.

[14] S. Sharma, A. Verma, Y. Xu *et al.*, "Robustly coordinated bi-level energy management of a multi-energy building under multiple uncertain-

ties," *IEEE Transaction on Sustainable Energy*, vol. 12, no. 1, pp. 3-13, Jan. 2021.

[15] L. Tian, L. Cheng, J. Guo *et al.*, "System modeling and optimal dispatching of multi-energy microgrid with energy storage," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 5, pp. 809-819, Sept. 2020.

[16] N. Nasiri, S. Zeynali, S. N. Ravadanegh *et al.*, "A tactical scheduling framework for wind farm-integrated multi-energy systems to take part in natural gas and wholesale electricity markets as a price setter," *IET Generation, Transmission & Distribution*, vol. 16, no. 9, pp. 1849-1864, May 2022.

[17] W. Dong and Q. Yang, "Data-driven solution for optimal pumping units scheduling of smart water conservancy," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 1919-1926, Mar. 2020.

[18] M. Daneshvar, B. Mohammadi-Ivatloo, K. Zare *et al.*, "Two-stage robust stochastic model scheduling for transactive energy based renewable microgrids," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 11, pp. 6857-6867, Nov. 2020.

[19] W. Hu, P. Wang, and H. B. Gooi, "Toward optimal energy management of microgrids via robust two-stage optimization," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 1161-1174, Mar. 2018.

[20] M. A. Velasquez, J. Barreiro-Gomez, N. Quijano *et al.* "Intra-hour microgrid economic dispatch based on model predictive control," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 1968-1979, May 2020.

[21] A. Parisio, E. Rikos, and L. Glielmo, "Stochastic model predictive control for economic/environmental operation management of microgrids: an experimental case study," *Journal of Process Control*, vol. 43, pp. 24-37, Jul. 2016.

[22] J. Sachs and O. Sawodny, "A two-stage model predictive control strategy for economic diesel-PV-battery island microgrid operation in rural areas," *IEEE Transaction on Sustainable Energy*, vol. 7, no. 3, pp. 903-913, Jul. 2016.

[23] Y. Yoldas, S. Goren, and A. Onen, "Optimal control of microgrids with multi-stage mixed-integer nonlinear programming guided *Q*-learning algorithm," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1151-1159, Nov. 2020.

[24] C. Keerthisinghe, A. C. Chapman, and G. Verbič, "Energy management of PV-storage systems: policy approximations using machine learning," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 257-265, Jan. 2019.

[25] E. Foruzan, L. Soh, and S. Asgarpoor, "Reinforcement learning approach for optimal distributed energy management in a microgrid," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5749-5758, Sept. 2018.

[26] J. Duan, Z. Yi, D. Shi *et al.*, "Reinforcement-learning-based optimal control of hybrid energy storage systems in hybrid AC-DC microgrids," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 9, pp. 5355-5364, Sept. 2019.

[27] W. Liu, P. Zhuang, H. Liang *et al.*, "Distributed economic dispatch in microgrids based on cooperative reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2192-2203, Jun. 2018.

[28] Y. Shang, W. Wu, J. Guo *et al.*, "Stochastic dispatch of energy storage in microgrids: an augmented reinforcement learning approach," *Applied Energy*, vol. 261, p. 114423, Mar. 2020.

[29] S. Dey, T. Marzullo, and G. Henze, "Inverse reinforcement learning control for building energy management," *Energy and Buildings*, vol. 286, p. 112941, May 2023.

[30] Q. Tang, H. Guo, and Q. Chen, "Multi-market bidding behavior analysis of energy storage system based on inverse reinforcement learning," *IEEE Transactions on Power Systems*, vol. 37, no. 6, pp. 4819-4831, Nov. 2022.

[31] S. Gao, C. Xiang, M. Yu *et al.*, "Online optimal power scheduling of a microgrid via imitation learning," *IEEE Transactions on Smart Grid*, vol. 13, no. 2, pp. 861-876, Mar. 2022.

[32] Y. Zhang, Q. Yang, D. Li *et al.*, "A reinforcement and imitation learning method for pricing strategy of electricity retailer with customers' flexibility," *Applied Energy*, vol. 323, p. 119543, Oct. 2022.

[33] W. Dong, Q. Yang, W. Li *et al.*, "Machine learning-based real-time economic dispatch in islanding microgrids in a cloud-edge computing environment," *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13703-13711, Sept. 2021.

[34] S. Kulkarni, Q. Gu, E. Myers *et al.*, "Enabling a decentralized smart grid using autonomous edge control devices," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7406-7419, Oct. 2019.

[35] Z. Gong, P. Zhong, and W. Hu, "Diversity in machine learning," *IEEE Access*, vol. 7, pp. 64323-64350, May 2019.

[36] J. G. Vlachogiannis and K. Y. Lee, "A comparative study on particle swarm optimization for optimal steady-state performance of power systems," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1718-1728, Nov. 2006.

[37] W. Dong and M. Zhou, "A supervised learning and control method to improve particle swarm optimization algorithms," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 7, pp. 1135-1148, Jul. 2017.

[38] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of International Conference on Neural Networks*, Perth WA, Australia, Nov. 1995, pp. 1942-1948.

[39] W. Dong, X. Chen, and Q. Yang, "Data-driven scenario generation of renewable energy production based on controllable generative adversarial networks with interpretability," *Applied Energy*, vol. 308, p. 118387, Feb. 2022.

[40] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798-1828, Aug. 2013.

[41] P. Vincent, H. Larochelle, I. Lajoie *et al.*, "Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion," *Journal of Machine Learning Research*, vol. 11, pp. 3371-3408, Dec. 2010.

[42] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, pp. 1735-1780, Nov. 1997.

[43] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM," *Neural Computation*, Vol. 12, pp. 2451-2471, Oct. 2000.

[44] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *Proceedings of International Conference on Learning Representations*, San Diego, USA, May 2015, pp. 1-13.

[45] Y. Wang, W. Dong, and Q. Yang, "Multi-stage optimal energy management of multi-energy microgrid in deregulated electricity markets," *Applied Energy*, vol. 310, p. 118528, Mar. 2022.

[46] Y. Dauphin, R. Pascanu, C. Gulcehre *et al.*, "Identifying and attacking the saddle point problem in high-dimensional non-convex optimization," in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, Montreal, Canada, Dec. 2014, pp. 2933-2941.

Wei Dong received the B.S. degree from Huazhong University of Science and Technology, Wuhan, China, in 2017, and the Ph.D. degree from the College of Electrical Engineering, Zhejiang University, Hangzhou, China, in 2022. He is currently an Associate Professor with the School of Automation, Hangzhou Dianzi University, Hangzhou, China. His research interests include energy internet, complex system modeling, learning-based optimization, and data-driven control.

Fan Zhang received the B.S. and Ph.D. degrees from the College of Control Science and Engineering, Zhejiang University, Hangzhou, China, in 1997 and 2002, respectively. He is currently a Full Professor with the School of Automation, Hangzhou Dianzi University, Hangzhou, China. His research interests include smart grids, integrated energy systems, electric Internet of Things, and complex system modeling and optimization.

Meng Li received the B.S. degree in School of Automation, Hangzhou Dianzi University, Hangzhou, China, in 2020. She is currently pursuing the M.Sc. degree with Hangzhou Dianzi University. Her research interests include learning-based optimization in smart grid.

Xiaolun Fang received the B.S. degree in electrical engineering and its automation in Zhejiang University of Technology, Hangzhou, China, in 2018, and the Ph. D. degree in the electrical engineering in Zhejiang University, Hangzhou, China, in 2023. She is currently an Associate Professor with the School of Automation, Hangzhou Dianzi University, Hangzhou, China. Her research interests include hydrogen-electricity integration, renewable energy analysis, energy efficiency, and optimization.

Qiang Yang received the Ph. D. degree in electronic engineering and computer science from the Queen Mary University of London, London, U.K., in 2007. He has worked with the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K., from 2007 to 2010. He is currently a Full Professor with the College of Electrical Engineering, Zhejiang University, Hangzhou, China. His research interests include smart energy systems, large-scale complex network modeling, control and optimization, and learning-based optimization and control.