Article:

Zhang, H, Yue, D, Dou, C et al. (3 more authors) (Accepted: 2021) Resilient optimal defensive strategy of TSK fuzzy-model-based micro-grids system via a novel reinforcement learning approach. IEEE Transactions on Neural Networks and Learning Systems. ISSN 2162-237X (In Press)

Resilient optimal defensive strategy of TSK fuzzy-model-based micro-grids system via a novel reinforcement learning approach

Huifeng Zhang, *Member, IEEE,* Dong Yue, *Fellow, IEEE,* and Chunxia Dou, *Member, IEEE,*
Xiangpeng Xie, *Member, IEEE,* Kang Li, *Senior member, IEEE,* Gerhardus P. Hancke, *Life Fellow, IEEE*

***Abstract*—With consideration of false data injection (FDI) on demand side, it brings great challenge for optimal defensive strategy with security issue, voltage stability, power flow and economic cost indexes. This paper proposes a Takagi-Sugeuo-Kang (TSK) fuzzy system based reinforcement learning approach for resilient optimal defensive strategy of inter-connected micro-grids. Due to FDI uncertainty of system load, TSK based deep deterministic policy gradient (DDPG) is proposed to learn actor network and critic network, where multiple indexes assessment occurs in critic network and security switching control strategy is made in actor network. Alternating direction method of multipliers (ADMM) method is improved for policy gradient with on-line coordination between actor network and critic network learning, and its convergence and optimality are proved properly. On the basis of security switching control strategy, penalty-based boundary intersection (PBI) based multi-objective optimization method is utilized to solve economic cost and emission issue simultaneously with considering voltage stability, rate-of-change of frequency (RoCoF) limits. According to simulation results, it reveals that the proposed resilient optimal defensive strategy can be a viable and promising alternative for tackling uncertain attack problem on inter-connected micro-grids.**

***Index Terms*—Takagi-Sugeuo-Kang fuzzy system, resilient optimal defensive, reinforcement learning, micro-grids.**

I. Introduction

THE resilience in power system is generally defined as the ability of power system to withstand severe disturbances without experiencing any large disruption, and further enabling a quick recovery to the normal operation state [1], [2], [3]. In spite of increasing situational awareness and automatic control of power grids, it also brings additional vulnerabilities as system scale increases and potential risks rise, which requires resilience enhancement measurements for power grids. Hence, resilient operation strategy can be a good choice for avoiding those risks. Actually, some researches on resilient operation have been taken in many existing literatures, where resilient operation strategies were proposed with considering natural disaster, misbehaving of power generators, FDI [4], [5], [6], [7], [8], [9], [10], [11]. In literature [4], a resilient distribution network planning problem was formulated as a two-stage robust optimization model with spatial and temporal dynamics of uncertain Hurricane disaster. Literature [6] proposed a practical framework for identifying network investments to offer highest hedge against risk of earthquake. Literature [9] proposed a three-stage resilient unit commitment model with considering typhoon paths and line outages, and optimal solution was deduced with worst-case scenario of possible typhoon paths. In literature [11], two-layer optimal augmented controller with observer to mitigate cyber disruptions, and it also models an intelligent type of FDI effect on cyber-physical interconnected micro-grids. However, those strategies lacks of assessment on the security risk or economic cost, while those assessment results can further guide measurements for decreasing those risks to minimization level [12].

Recently, many scholars proposed several kinds of assessment approaches [13], [14], [15], [16], [17], [18]. Literature [13] proposed a novel assessment approach of required reserve capability for meeting forecasting uncertainty dynamics in micro-grid. Literature [14] proposed an efficient and accurate model of inverter-based micro-grids with reduced-order, the improved model accounts for the effects of network dynamics as well as similar to quasi-stationary model. In literature [16], a comprehensive method was proposed to assess the uncertainty characteristics of overall photo-voltaic resources with gray-box model, which consists of physical and data-driven sub-models with currents, voltage information of photo-voltaic resources. Literature [18] integrated impact assessment model and optimal restoration model of active distribution network together with non-sequential Monte Carlo Simulation framework. Here, reinforcement learning (RL) method is utilized to assess security risk and economic indexes of interconnected micro-grids, as it is more powerful than other alternatives due to its model-free

H. Zhang is with the institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Jiangsu Province, China, 210023, e-mail: zhanghuifeng_520@163.com

D. Yue is with the institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Jiangsu Province, China, 210023, e-mail: medongy@vip.163.com (Corresponding author)

C. Dou is with the institute of Advanced Technology, Nanjing University of Posts and Telecommunications,Jiangsu Province, China, 210023, e-mail: cxdou@ysu.edu.cn (Corresponding author)

X. Xie is with the institute of Advanced Technology, Nanjing University of Posts and Telecommunications,Jiangsu Province, China, 210023, e-mail: xiexiangpeng1953@163.com

K. Li is with the school of Electronic and Electrical Engineering, University of Leeds, Leeds, LS2 9JT, United Kingdom, e-mail: K.Li1@leeds.ac.uk

Gerhardus P. Hancke is with college of automation, Nanjing University of Posts and Telecommunications, Jiangsu Province, China, and also University of Pretoria, South Africa, e-mail: g.hancke@ieee.org

strategy and knowledge-cumulative mechanisms [19], [20]. Literature [21] developed a short-term voltage stability assessment strategy with on-line systematic imbalance learning machine. In literature [22], a novel deep learning based feature extraction framework was developed for creating security rules of electricity system operation. Literature [23] proposed a multi-stage game between attacker and the defender with reinforcement learning to identify optimal attack sequences given certain objectives. Literature [24] presented deep reinforcement learning models for vulnerability assessment of power system topology optimization with considering data perturbation and cyber-attack. Those existing learning based assessment method can be efficient for dealing with deterministic issue, while it can not tackle with uncertain input problem.

Since smart meter on demand side can be affected by different factors, such as electricity burglary action, device fault and cyber attack, these factors can lead the false injection of consumers' load information. This FDI attack can mislead energy management, it can make wrong dispatch scheme, which will increase economic cost/emission rate, and even cause security problem. The resilience of interconnected micro-grids mainly lies on two parts: robustness for uncertain input and recovery from FDI. In this paper, a TSK fuzzy system based learning strategy is improved to solve the uncertain input problem, and multi-objective optimal defensive strategy is also proposed to recover power system into normal operation state. The main contribution of this paper can be summarized as follows:

1) This paper proposes an optimal recovery model with frequency-awareness under FDI on demand side. Since FDI of system load can affect operation scheme of interconnected micro-grids and even lead to security risks, frequency-awareness limits with RoCoF have been taken into consideration.

2) This paper improves a TSK fuzzy system based deep reinforcement learning based strategy for assessment on multiple indexes of interconnected micro-grids. With consideration of uncertainty nature of FDI on demand side, TSK fuzzy system is modeled for uncertain input of load parameters, and accumulation knowledge with actor network and critic network learning is utilized to guide multiple indexes assessment with ADMM algorithm, and security control strategy is also made to reduce security risk with actor network learning.

3) This paper utilizes a decomposition based multi-objective differential evolution algorithm for optimal recovery of interconnected micro-grids. On the basis of assessment and security control of interconnected micro-grids, the interconnected micro-grids require to recover to normal operation state with low security risk as well as economic cost, a decomposition based multi-objective differential evolution is utilized to obtain optimal scheme for recovery of interconnected micro-grids.

The arrangement of paper structure is presented as: Problem formulation is presented in Section II, the proposed assessment method is described in Section III, and active defensive strategy is introduced in Section IV, and the simulation results and conclusion are shown in Section V and Section VI.

II. Problem formulation of assessment analysis and optimal defensive model

With consideration of FDI attack, the measurement for enhancing inter-connected micro-grids consists of two procedures: assessment on multiple indexes and optimal recovery after FDI. The assessment procedure mainly detects the FDI effect on multiple indexes of interconnected micro-grids, and then optimal recovery procedure controls micro-grids to normal state.

A. *Multiple indexes of interconnected micro-grids*

Security issue, economic cost and environment protection can be the most important metrics for interconnected micro-grids, while power supply voltage and frequency stability can be the main metric for system security, emission rate can be an important factor for environment protection, and power generation cost can be the main metric for economic cost. The assessment on those indexes can be necessary for defensive of FDI attack, the indexes are presented as follows: (1) Power supply security: The power supply index mainly presents the security level of deviation between power supply and load demand with considering adjustment ability from power electronic devices.

$$F_1 = Security_{con} = Prob(\sum_{n \in N_G} \alpha_n |P_{G,n,t} + P_{S,n,t} - P_{load,n,t} - P_{loss,n,t}| < \epsilon_t) \tag{1}$$

where N_G represents the number of interconnected micro-grids, α_n is state parameter of nth micro-grid, $\epsilon_t > 0$ denotes the permitted deviation, $P_{tot,n,t}$, $P_{load,n,t}$ and $P_{loss,n,t}$ denote total power output, load demand and transmission loss respectively. The power transmission $P_{loss,n,t}$ can lead power loss among interconnected micro-grids, it can be described as:

$$P_{loss,n,t} = \sum_{n_1 \in N_{node}} d_{n_1,n} R_{n_1,n} \frac{P^2_{n_1,n} + Q^2_{n_1,n}}{U^2_{n_1}} \tag{2}$$

where $d_{n_1,n}$ and $R_{n_1,n}$ represent the distance and resistance between n_1th micro-grid and nth micro-grid, U_{n_1} is the rating line voltage, $P_{n_1,n}$ and $Q_{n_1,n}$ denote the power flow from n_1th micro-grid to nth micro-grid.
(2) Voltage stability: The voltage index $F_2 = 1/N_G \sum_{n=1}^{N_G} U_n(t)$ of each micro-grid must be controlled in security domain as follows:

$$U_{n,min} \leq U_n(t) \leq U_{n,max} \tag{3}$$

where $U_{n,min}$ and $U_{n,max}$ represent the minimum and maximum voltage at nth micro-grid.

(3) Frequency stability: Since interconnected micro-grids have low inertial characteristic, the frequency index $F_3 = f(t)$ can change sharply during an islanding transition [25]. The RoCoF can be taken into consideration as follows:

$$\begin{cases} F_4 = RoCoF_t = (\sum_{n\in N_G} \triangle P_{G,n,t} + \sum_{n\in N_G} \triangle P_{S,n,t} \\ -D \triangle f(t) - P_{M,t})/2H_t \\ -RoCoF^{max} \leq RoCoF_t \leq RoCoF^{max} \end{cases} \tag{4}$$

where $\triangle P_{G,n,t}$ represents power deviation of power generator, $\triangle P_{S,n,t}$ denotes power deviation of energy storage, D is load damping factor, $\triangle f(t)$ denotes frequency deviation, and $P_{M,t}$ represents power imbalance, it can also be described as $\sum_{n\in N_G} P_{G,n,t} + \sum_{n\in N_G} P_{S,n,t} - \sum_{n\in N_G} P_{load,n,t} - \sum_{n\in N_G} P_{loss,n,t}$, and H_t denotes the inertial of interconnected micro-grids, and it can be calculated as $(\sum_{n\in N_G} h_{G,n} P_{G,n}^{max} + \sum_{n\in N_G} h_{S,n} P_{S,n}^{max})/f^0$, where $h_{G,n}$ and $h_{S,n}$ represent inertial constant of power generator and energy storage, $RoCoF^{max}$ represents maximum fluctuation limit of RoCoF, $P_{G,n}^{max}$ and $P_{S,n}^{max}$ denote maximum output of power generator and energy storage, and f^0 is initial frequency, which can be set as 50HZ.

(4) Economic cost: Since renewable energy can be considered with no power generation cost, economic cost can be made merely by CHP generators. The economic cost can be described as [26]:

$$\begin{aligned} F_5 =& Eco_t = \sum_{n=1}^{N_G} H_{n,t}[\alpha_{n,1} + \alpha_{n,2} P_{CHP,n,t} + \alpha_{n,3} P_{CHP,n,t}^2 \\ &+ |\alpha_{n,4} sin(\alpha_{n,5}(P_{CHP,n,min} - P_{CHP,n,t}))] \end{aligned} \tag{5}$$

where $H_{n,t}$ represents the on/off state of nth micro-grid with binary variable 0 or 1, $\alpha_{n,1}$, $\alpha_{n,2}$, $\alpha_{n,3}$, $\alpha_{n,4}$ and $\alpha_{n,5}$ denote the coefficients of cost, $P_{CHP,n,min}$ is the minimum limit of CHP output.

(5) Emission rate: Since power generation of CHP generator can produce emission pollutant, which can affect social lives. Hence, emission rate is taken as another index as [26]:

$$\begin{aligned} F_6 =& Emi_t = \sum_{n=1}^{N_G} H_{n,t}[\beta_{n,1} + \beta_{n,2} P_{CHP,n,t} \\ &+ \beta_{n,3} P_{CHP,n,t}^2 + \beta_{n,4} exp(\beta_{n,5} P_{CHP,n,t})] \end{aligned} \tag{6}$$

where $\beta_{n,1}$,$\beta_{n,2}$,$\beta_{n,3}$,$\beta_{n,4}$ and $\beta_{n,5}$ represent the coefficients of emission rate.

B. The constraint limits of optimal active defensive model

(1) The power flow constraints: The power flow among different micro-grids can be described as follows:

$$\begin{cases} \sum_{n_2\in \Xi_{G,n_1,t}} [P_{n_2,n_1}(t) - R_{n_2,n_1} \frac{(P_{n_2,n_1}(t))^2 + (Q_{n_2,n_1}(t))^2}{(U_{n_2}(t))^2}] \\ = P_{n_1}(t)' + \sum_{n_2'\in \Xi_{G,n_1,t}'} P_{q,n_2'}(t) \\ \sum_{n_2\in \Xi_{G,n_1,t}} [Q_{n_2,n_1}(t) - Z_{n_2,n_1} \frac{(P_{n_2,n_1}(t))^2 + (Q_{n_2,n_1}(t))^2}{(U_{n_2}(t))^2}] \\ = Q_{n_1}(t)' + \sum_{n_2'\in \Xi_{G,n_1,t}'} Q_{n_1,n_2'}(t) \end{cases} \tag{7}$$

The above formulation mainly represents reactive power flow balance and active power flow balance with considering resistance and reactance of transmission line, and the voltage of micro-grids must satisfy following conditions:

$$\begin{aligned} &(U_{n_1}(t))^2 - (U_{n_2}(t))^2 + 2(R_{n_2,n_1} P_{n_2,n_1}(t) \\ &+ Z_{n_2,n_1} Q_{n_2,n_1}(t)) - [(R_{n_2,n_1})^2 \\ &+ (Z_{n_2,n_1})^2] \frac{(P_{n_2,n_1}(t))^2 + (Q_{n_2,n_1}(t))^2}{(U_{n_2}(t))^2} = 0 \end{aligned} \tag{8}$$

where $\Xi_{n_1,t}$ represents the interconnected micro-grids of power flow to n_1th micro-grid, and power flow from n_1th micro-grid to micro-grid set $\Xi_{n_1,t}'$. R_{n_2,n_1} and Z_{n_2,n_1} denote resistance and reactance between micro-grid n_2 and micro-grid n_1, $U_{n_1}(t)$ is the voltage of micro-grid n_1, and those following constraint limits should also be satisfied:

$$\begin{cases} U_{n_1}^{min} \leq U_{n_1}(t) \leq U_{n_1}^{max} \\ Q_{n_1}(t) = P_{n_1}(t) tan\varphi_{n_1} \\ Q_{n_1}^{min} \leq Q_{n_1}(t) \leq Q_{n_1}^{max} \end{cases} \tag{9}$$

where $U_{n_1}^{min}$ and $U_{n_1}^{max}$ represent the minimum and maximum voltage limits, and $Q_{n_1}^{min}$ and $Q_{n_1}^{max}$ denote the minimum and maximum reactive power limits, φ_{n_1} is the power factor angle.

(2) The output constraint limit: The output at each micro-grid can be describe as follows:

$$P_n(t) = P_{G,n,t} + P_{S,n,t} \tag{10}$$

where $P_{G,n,t}$ can also be described as $P_{CHP,n,t} + P_{w,n,t} + P_{v,n,t}$, $P_{CHP,n,t}$ represents the output of combined heat and power generator, $P_{w,n,t}$ and $P_{v,n,t}$ denote the output of wind power and PV generator. The above output must satisfy following limits:

$$\begin{cases} P_{CHP,n,min} \leq P_{CHP,n,t} \leq P_{CHP,n,max} \\ Ramp_{down,n} \leq P_{CHP,n,t} - P_{CHP,n,t-1} \leq Ramp_{up,n} \end{cases} \tag{11}$$

where $P_{CHP,n,min}$ and $P_{CHP,n,max}$ represent the minimum and maximum output of CHP generator, $Ramp_{down,n}$ and $Ramp_{up,n}$ denote the ramp down and ramp up of CHP generator.

(3) The limits of energy storage: The energy storage can be a reliable energy resource for ensuring interconnected micro-grids system stability, $P_{S,n,t}$ can be considered as $\sum_{r=1}^{N_{n,ESS}} P_{S,n,r,t}$, which represents the summation of provided power of $N_{n,ESS}$ energy storages in r micro-grid. With consideration of charging or discharging state, $P_{S,n,r,t}$ can be taken as $P_{S,n,r,t}^{dis}$ if energy storage is discharging, else it can be taken as $P_{S,n,r,t}^{cha}$, and the charging/discharging process must satisfy following limits:

$$\begin{cases} E_{n,r}^{ESS}(t+1) = E_{n,r}^{ESS}(t) + \eta_{n,r}^{cha} P_{S,n,r,t}^{cha} - P_{S,n,r,t}^{dis}/\eta_{n,r}^{dis} \\ E_{n,r}^{ESS,min} \leq E_{n,r}^{ESS}(t) \leq E_{n,r}^{ESS,max} \\ 0 \leq P_{S,n,r,t}^{dis} \leq P_{S,n,r,t}^{dis,max} \\ 0 \leq P_{S,n,r,t}^{cha} \leq P_{S,n,r,t}^{cha,max} \\ E_{n,r}^{ESS}(0) = E_{n,r}^{ESS,initial} \end{cases} \tag{12}$$

where $E_{n,r}^{ESS}(t)$ is the rth energy storage of nth micro-grid, $E_{n,r}^{ESS,min}$ and $E_{n,r}^{ESS,max}$ are the minimum and maximum bounds of rth energy storage in nth micro-grid, $P_{S,n,r,t}^{dis}$ and $P_{S,n,r,t}^{cha}$ are the output of rth energy storage in discharging and charging state, $P_{S,n,r,t}^{dis,max}$ and $P_{S,n,r,t}^{cha,max}$ are the maximum discharging and charging output of rth energy storage, $\eta_{n,r}^{dis}$ and $\eta_{n,r}^{cha}$ are the efficiency factor of rth energy storage in discharging and charging state, $E_{n,r}^{ESS,initial}$ denotes the initial storage of energy storage.

III. THE LEARNING WITH TSK FUZZY APPROACH FOR SECURITY ASSESSMENT AND SWITCHING CONTROL OF INTERCONNECTED MICRO-GRIDS

With consideration of FDI on demand side, it can cause large deviation between actual system load and observed data, which may cause security issue when observed data is directly taken for optimal operation of interconnected micro-grids. For system load $P_{load,n,t}$ at node n, it can be described as:

$$\begin{cases} P_{load,n,t} = \bar{P}_{load,n,t} + \xi_{n,t}\widetilde{P_{load,n,t}} \\ \widetilde{P_{load,n,t}} \in [\widetilde{P_{load,n,t}^{min}}, \widetilde{P_{load,n,t}^{max}}] \end{cases} \tag{13}$$

where $\bar{P}_{load,n,t}$ denotes actual system load, $\xi_{n,t} > 0$ and $\widetilde{P_{load,n,t}}$ represent uncertain parameter and uncertain deviation of FDI on system load, $\widetilde{P_{load,n,t}^{min}}$ and $\widetilde{P_{load,n,t}^{max}}$ are the minimum and maximum bounds of uncertainty deviation. Suppose uncertain budget $\Delta_t > 0$ can be described as [27]:

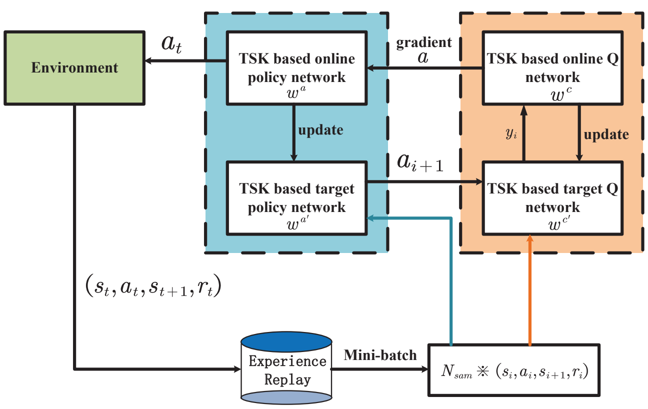


Fig. 1. The TSK model based learning strategy with DDPG approach

$$\sum_{n=1}^{N_G} \xi_{n,t} \leq \Delta_t \tag{14}$$

Hence, security assessment is required to evaluate the system safety, the TSK fuzzy model based DDPG learning strategy is developed to tackle with this problem, its framework is presented in Fig.1. Some definitions are defined as follows:
State set: The state set can be defined as $s_t = [\xi_{1,t}, \xi_{2,t}, \cdots, \xi_{N_G,t}]$, those variables has main effect on those security and economic indexes.
Action set: The action is mainly implemented to prevent decreasing of maximization index and increasing of minimization index, action set can be defined as the set of on/off state of power generators, which can be described as $a_t = [H_{1,t}, H_{2,t}, \cdots, H_{N_G,t}]$.
Reward: During the learning process, state and action must be rewarded or punished according to the effect on those indexes, and reward function can be defined as $r(s_t, a_t)$. If evaluation index is to be maximized, $r(s_t, a_t)$ can be specifically expressed as F_i, otherwise, it can be described as $-F_i$. Specially, if evaluation index is bounded with upper bound F^{max} and lower bound F^{min}, it can be described as minimization index with $|F_i - (F^{max} + F^{min})/2|$.

A. The TSK fuzzy model based network learning with FDI uncertain input

Here, a deep reinforcement learning approach is developed to learn the relationship between uncertain parameters $[x_1, x_2, \cdots, x_{N_G}] = [\xi_{1,t}, \xi_{2,t}, \cdots, \xi_{N_G,t}]$ and each index F_i. With consideration of input uncertainty, the TSK fuzzy model can be created to learn the actor network as follows:

$$\begin{aligned} &Rule\ i : IF(x_1 \in A_{i1}) AND ... AND(x_{N_G} \in A_{iN_G}) \\ &THEN \quad q_i = g_i(x_1, ..., x_{N_G}) \quad i = 1, 2..., N_R \end{aligned} \tag{15}$$

where $x_j (j = 1, 2..., N_G)$ represents the jth input variable, $A_{i1}, A_{i2}, ...A_{iN_G}$ denote the fuzzy sets, q_i is the output of ith rule consequent, $g(\cdot)$ represents the output function, and N_R is the number of rules. For each fuzzy set A_{ij}, its membership function can be described as:

$$M_{ij}(x_j) = exp\{-\frac{(x_{ij} - m_{ij})^2}{b_{ij}^2}\} \tag{16}$$

where m_{ij} represents the central value of fuzzy set A_{ij}, b_{ij} denotes the width of fuzzy set. The fuzzy AND operation is calculated with algebraic product, and the firing strength ϕ_i of ith rule can be deduced as follows:

$$\phi_i = \prod_{j=1}^{N_G} M_{ij}(x_j) = exp\{-\sum_{j=1}^{N_G} \frac{(x_{ij} - m_{ij})^2}{b_{ij}^2}\} \tag{17}$$

The output of the fuzzy system can be calculated with weighted de-fuzzification method as:

$$q = \frac{\sum_{i=1}^{N_G} \phi_i q_i}{\sum_{i=1}^{N_G} \phi_i} \tag{18}$$

The output function $g(\cdot)$ can be approximated with RBF neural network as follows:

$$g_i(X) = \sum_{j=1}^{N_G} w_{ij} / \sqrt{\| X - c_{1j} \|_2^2 + \kappa_j^2} \tag{19}$$

where X represents the input vector, w_{ij} denotes the weights of network, c_{1j} and κ_j represent the control vector and parameter of RBF function. With consideration of over-fitting and generalization issue of each y_i, the following L_2 regularized loss function is created:

$$L_\lambda = \frac{1}{2} \sum_{n=1}^{N_{sam}} (q_i^{(n)} - g_i(X_n))^2 + \frac{\lambda}{2} \sum_{j=1}^{N_G} {w_{ij}}^2 \tag{20}$$

where $X_n = [x_1^n, x_2^n, ..., x_N^n]$, N_{sam} represents the number of samples, $y_i^{(n)}$ denotes the actual value of nth sample, n denotes the sample index, $\lambda \geq 0$ is the regularization parameter.

B. The actor network and critic network learning with policy gradient action

With consideration of uncertain income, output function of above TSK fuzzy model can convert uncertain vector into deterministic vector, the Pseudo code of network learning is shown in **Algorithm**. Since critic network mainly evaluates current state and action, it can be considered as the function of state and action vector. As it is known in reinforcement learning, the state-action value function can be deduced with Bellman theory as follows:

$$Q^k(s,a) = \max_{a\in\Omega}[\tau Q^{k-1}(s,a) + r^k(s,a)] \tag{21}$$

where $Q^k(s,a)$ represents the state-action value function with state vector s and action vector a at kth step,τ is discount factor, and $r^k(s,a)$ denotes the reward function at kth step. The optimal action of $Q^k(s,a)$ can be trained with actor network $a(s) = g_i(s)$, which represents TSK based network. The loss function can be described with L_{λ_1} as:

$$L_{\lambda_1} = \frac{1}{2}\sum_{n=1}^{N_{sam}}(a_i^{(n)} - a_i(s))^2 + \frac{\lambda}{2}\sum_{j=1}^{N_G} {w_{ij}^a}^2 \tag{22}$$

where w_{ij}^a denotes the weight of actor network. Here, the state vector and action vector can be both treated as input variables, the state-action value function can be approximated as:

$$Q^k(s,a) = \sum_{j=1}^{N_G} w_j^{ck}\psi_j^k(s) + \sum_{l=1}^{N_a} w_{N+l}^{ck}\psi_{N+l}^k(a) \tag{23}$$

where w_j^{ck} represents the weights of network, and $\psi_j(\cdot)$ denotes TSK based RBF neural network function, which can be described as follows:

$$\psi_j^k(X) = exp[-\frac{1}{2\delta_k^2} \parallel X - c_{2j}^k \parallel^2] \tag{24}$$

where δ_k and c_{2j}^k represent the scaling and altitude control parameters. Generally, the objective function can be treated as reward function. Here, each objective function f_i can be evaluated with critic network according to this approach. For ensuring the learning efficiency, the Lagrangian function can be created as follows:

$$L_{\lambda_2} = \frac{1}{2}\sum_{n=1}^{N_{sam}}\sum_{j=1}^{N_G}(y_n - Q(s^{(n)}, a^{(n)}))^2 + \frac{\lambda_2}{2}\sum_{j=1}^{N_G+N_a} {w_j^{ck}}^2 \tag{25}$$

where R_n represents the reward value of nth sample, $s^{(n)}$ and $a^{(n)}$ are the state and action value of nth sample, λ_2 denotes the regularization parameter. Combined with actor network weight w^{c*}, it can be taken as the function of state variable s. Since s and a are trained together with critic network weight w^c, the remaining task is to deduce optimal w^c for minimizing L_{λ_2}.

C. On-line coordinated policy gradient between actor network and critic network with ADMM approach

Since action vector in critic network is deduced with actor network learning, coordination between actor network and critic network can promote optimization efficiency. Here, the ADMM approach is developed to coordinate above two models. Suppose $\chi_1(w^a) = \mu_1 L_{\lambda_1}(w^a)$ and $\chi_2(w^c, z) = \mu_2 L_{\lambda_2}(w^c, z)$, where $w^a = [w_{i1}^a, w_{i2}^a, \cdots, w_{iN_G}^a]^T$, $w^c = [w_{i1}^a, w_{i2}^a, \cdots, w_{iN_G}^a]^T$ and $z = [z_{i1}^a, z_{i2}^a, \cdots, z_{iN_G}^a]^T$,$\mu_1$ and μ_2 are two weight parameters. The augmented Lagrangian problem can be equalized as follows:

$$\min L_\rho = \chi_1(w^a) + \chi_2(w^c, z) + y^T(w^a - z) + \frac{\rho}{2} \parallel w^a - z \parallel^2 \tag{26}$$

where ρ and y^T represent the control parameter and control vector. Suppose $u = y/\rho$, then the iteration algorithm in ADMM scaled form can be presented as follows:

$$\begin{cases} w^{a\,k+1} = \arg\min_{w^a}[\chi_1(w^a) + \frac{\rho}{2} \parallel w^a - z^k + u^k \parallel_2^2] \\ z^{k+1} = \arg\min_{z}[\chi_2(w^{ck}, z) + \frac{\rho}{2} \parallel w^{a\,k+1} - z + u^k \parallel_2^2] \\ u^{k+1} = u^k + (w^{a\,k+1} - z^{k+1}) \end{cases} \tag{27}$$

Due to mere existence of w^c in $\chi_2(w^c, z)$, weight w^c of critic network can be deduced with local iteration as follows:

$$w^{ck+1} = w^{ck} + \eta_{w^c}\frac{\partial\chi_2(w^{ck}, z^{k+1})}{\partial w^{ck}} \tag{28}$$

Combined with above iteration framework, it can be further rewritten as:

$$\begin{cases} w^{a\,k+1} = w^{a\,k} + \eta_{w^a}[\nabla\chi_1(w^{a\,k}) + \rho(w^{a\,k} - z^k + u^k)] \\ z^{k+1} = z^k + \eta_z[\frac{\partial\chi_2(w^{ck}, z^k)}{\partial z^k} - \rho(w^{a\,k+1} - z^k + u^k)] \\ u^{k+1} = u^k + (w^{a\,k+1} - z^{k+1}) \end{cases} \tag{29}$$

where ∇ represents the gradient operator, η_{w^a} and η_z denote iteration step length parameter. In addition, suppose $r^k = w^{a\,k} - z^k$, the stopping criteria can be presented as follows:

$$\parallel r^k \parallel_2 \leq \epsilon^{pri} \tag{30}$$

where ϵ^{pri} denotes the primal residual, which can be described as:

$$\epsilon^{pri} = \sqrt{N_G}\epsilon^{abs} + \epsilon^{rel}\max\{\parallel w^a \parallel_2, \parallel z \parallel_2\} \tag{31}$$

where $\epsilon^{abs} > 0$ denotes the absolute tolerance, and $\epsilon^{rel} > 0$ is the relative tolerance, which is generally taken as 10^{-3} or 10^{-4}. With above algorithm, optimal action can be deduced with current state by training actor network weight w^a and critic network weight w^a. In addition, the weights of target actor network and target critic network can be updated as follows:

$$\begin{cases} w^{a'} = \mu w^a + (1-\mu)w^{a'} \\ w^{c'} = \mu w^c + (1-\mu)w^{c'} \end{cases} \tag{32}$$

where $\mu \in [0,1]$ denotes update parameter, $w^{a'}$ and $w^{c'}$ represent weight of target actor network and target critic network.

Algorithm Assessment and security control with FDI

```
procedure TSK fuzzy system based DDPG algorithm
Initialization: Initialize critic network w^c and actor network w^a, and their copied target networks w^{c'} and w^{a'}, ReplayBuffer = φ, episode = 1;
    Collect input parameter [ξ_{1,t}, ξ_{2,t}, ···, ξ_{N_G,t}];
    while episode < maxcount1 do
        Store transition information (s, a, s', r) in ReplayBuffer;
        while count < maxcount2 do
            Update weights w^{ak} in actor network with fuzzy system input;
            Update weights w^{ck} in critic network with fuzzy system input;
            Update parameter u^k and z^k;
            k = k + 1;
        end while
        Update weights w^{c'} and w^{a'} of target networks;
        episode = episode + 1;
    end while
end procedure
```

D. The analysis on convergence and optimality of proposed algorithm

Before the convergence and optimality analysis, some remarks are required as:

Remark 1: The functions $\chi_1(w^a)$ and $\chi_2(w^c, z)$ are both closed, proper and convex, which can be described as $\chi_1(w^a) = \{(w^a, \Omega) \in R^n \times R | \chi_1(w^a) \le \Omega\}$ and $\chi_2(w^c, z) = \{(z, \Omega) \in R^n \times R | \chi_2(w^c, z) \le \Omega\}$

Remark 2: The Lagrangian function L_0 exists at least one saddle point, suppose $L_0(w^a, w^c, z, y) = \chi_1(w^a) + \chi_2(w^c, z) + y^T(w^a - z)$ and $\Theta = [w^a, w^c, z]$, it can be described as:

$$L_0(\Theta^*, y) \le L_0(\Theta^*, y^*) \le L_0(\Theta, y^*) \tag{33}$$

As it is illustrated in literature [28], the following Lyapunov function $V^k \to 0$:

$$V^k = (1/\rho) \parallel y^k - y^* \parallel_2^2 + \rho \parallel z^k - z^* \parallel_2^2 \tag{34}$$

It also means that $r^k \to 0$, $w^{ak} \to z^k$ and $z^k \to z^*$ when $k \to +\infty$, the convergence and optimality of w^a, z and u hold. However, the proposed coordinated algorithm has extra variable w^c, here the convergence and optimality are proved in the following section.

(1) The optimality proof

Proof: Since Remark 1 holds, $\chi_2(w^c, z)$ is closed, proper and convex, so it is L_ρ. For simplicity, suppose $p^k = \chi_1(w^a) + \chi_2(w^{ck}, z)$ and $p^* = \chi_1(w^{a*}) + \chi_2(w^{c*}, z^*)$. As w^{ck} minimizes L_ρ, it satisfies the necessary and sufficient condition:

$$0 \in \partial L_\rho(w^a, w^{ck}, z) = \partial \chi_2(w^{ck}, z) \tag{35}$$

It can also mean that:

$$p^k = \chi_1(w^a) + \chi_2(w^{ck}, z) \le \chi_1(w^a) + \chi_2(w^{c*}, z) \tag{36}$$

Specially, if $w^a = w^{a*}$ and $z = z^*$, it has $p^k \le p^*$. Simultaneously, as Remark 2 holds, $w^{a*} = z^*$ and $w^a \to z$, it obtains:

$$L_0(w^{a*}, w^{c*}, z^*, y^*) \le L_0(w^a, w^{ck}, z, y^*) \tag{37}$$

It can be described as:

$$p^* = \chi_1(w^{a*}) + \chi_2(w^{c*}, z^*) \le \chi_1(w^a) + \chi_2(w^{ck}, z) = p^k \tag{38}$$

Combined with formula(36) and formula(38), it has $p^k \to p^*$, the optimality holds. ■

(2) The convergence proof

Proof: The convergence can also hold with satisfying following condition:

$$\eta_c \le \frac{1}{\sqrt{M}} \cdot \frac{1}{K} \tag{39}$$

where $K > 0$ is a large enough number, constant $M > 0$ can be described as $\max\{\Psi^T\Psi\}$, $\Psi = [\psi_1(s), \cdots, \psi_{N_G}(s), \psi_{N_G+1}(a), \cdots, \psi_{N_G+N_a}(a)]^T$.

$$\begin{aligned} \parallel \chi_2(w^{ck}, z) - \chi_2(w^{c*}, z) \parallel_2^2 &\le \eta^{c2}\Psi^T\Psi \parallel w^{ck} - w^{c*} \parallel_2^2 \\ &\le \delta_c^2/(K^2) \le \epsilon_c \end{aligned} \tag{40}$$

where $\epsilon_c > 0$ denotes the accuracy parameter, $\delta_c = \max\{\parallel w^{ck} - w^{c*} \parallel_2\}$, and K can take $\lceil \frac{\delta_c}{\sqrt{\epsilon_c}} \rceil$, then the convergence holds. ■

IV. The optimal active defensive strategy of interconnected micro-grids

A. The active defensive model with multiple objectives

With above assessment on multiple indexes, the security on/off switching strategy can also be made to avoid potential risk of interconnected micro-grids. However, the switching strategy can merely tackle with one security or economic issue, further defensive strategy must be made for ensuring system security as well as minimizing economic cost. The supply security index can be considered as equality constraint, and voltage stability and frequency stability indexes can be treated as inequality constraint limits. Hence, the active defensive model can be created as follows:

$$\begin{cases} \min & \{\sum_{t=1}^{T} Eco_t, \sum_{t=1}^{T} Emi_t\} \\ s.t. & \sum_{n \in N_G} P_{G,n,t} + \sum_{n \in N_G} P_{S,n,t} - \sum_{n \in N_G} \\ & P_{loss,n,t} = \sum_{n \in N_G} P_{load,n,t} \\ & V_{n,min} \le V_{n,t} \le V_{n,max} \\ & -RoCoF^{max} \le RoCoF_t \le RoCoF^{max} \\ & Power\ flow\ constraints; \\ & Output\ constraints; \\ & The\ limits\ of\ energy\ storage. \end{cases} \tag{41}$$

Since economic cost and emission rate can be contradict with each other. Hence, the optimal active defensive model can be treated as a multi-objective optimization problem, and an efficient multi-objective optimization algorithm is required to optimize this model.

B. Decomposition based multi-objective optimization approach for optimal recovery of micro-grid

Here, the penalty-based boundary intersection (PBI) approach with gradient decent based differential evolution is developed to solve this problem combined with several constraint handling techniques. Without loss of generality, the active defensive model can be described as follows:

$$\begin{cases} \min F(x) = (f_1(x), f_2(x), ..., f_m(x))^T \\ s.t. \quad h_1^{j_1}(x) < 0, j_1 = 1, 2, ..., J_1; \\ \qquad h_2^{j_2}(x) = 0, j_2 = 1, 2, ..., J_2; \\ \qquad x \in \mathbb{R}^n \end{cases} \tag{42}$$

where $h_1(\cdot)$ and $h_2(\cdot)$ represent the inequality and equality constraint function, m is the number of objectives, J_1 and J_2 denote the number of inequality and equality constraint limits. The main idea of PBI is to search optimal solutions with guidance of utopian point $z^* = (z_1^*, z_2^*, ..., z_m^*)^T$, the above problem can be equalized as follows:

$$\begin{cases} \min \; g^{pbi}(x|\lambda^i, z^*) = d_1^i + \beta d_2^i \\ s.t. \; d_1^i = \parallel (F(x) - z^*)^T \lambda^i \parallel / \parallel \lambda^i \parallel \\ \qquad d_2^i = \parallel F(x) - z^* - d_1^i \lambda^i \parallel \\ \qquad x \in \Omega \end{cases} \tag{43}$$

where d_1^i represent the distance between z^* and projection of $F(x)$ on the ith subproblem, d_2^i denotes the distance between $F(x)$ and direction line of ith subproblem, β is the preset penalty parameter, Ω denotes the feasible domain, which is determined by equality and inequality constraint limits. λ^i represents the direction vector, its component λ_j^i satisfies $\sum_{j=1}^{m} \lambda_j^i = 1$ ($\lambda_j^i \geq 0$).
Obviously, the optimal model can be converted into above version, which optimizes several single-objective subproblems with alternating weights. To properly solve each subproblem, gradient decent based differential evolution is utilized to enhance the search ability. Here, differential evolution procedure is taken with mutation operator of improved DE/rand/1/bin strategy, which can be described as follows:

$$\begin{cases} X_{G+1}^j = X_{r,G}^j + \beta_1^j (X_{r1,G} - X_{r2,G}) + \beta_2^j (X_{r3,G} - X_{r4,G}) \\ r \neq r1 \neq r2 \neq r3 \neq r4 \end{cases} \tag{44}$$

where $X_{r,G}$, $X_{r1,G}$, $X_{r2,G}$, $X_{r3,G}$ and $X_{r4,G}$ are randomly selected individuals from non-dominated solutions, X_{G+1}^j is the generated individual for $G+1$ generation, β_1^j and β_2^j represent the control parameters. With consideration of convergence ability, β_1^j and β_2^j can be updated as follows:

$$\begin{cases} \beta_1^j = -\dfrac{\Upsilon_G \kappa_1 sgn(f_1(X_{r1,G}) - f_1(X_{r2,G}))}{(X_{r1,G} - X_{r2,G})^2 \sqrt{\sum_{j \in n} \frac{1}{(X_{r1,G} - X_{r2,G})^2}}} \\ \beta_2^j = -\dfrac{\Upsilon_G \kappa_2 sgn(f_2(X_{r3,G}) - f_2(X_{r4,G}))}{(X_{r3,G} - X_{r4,G})^2 \sqrt{\sum_{j \in n} \frac{1}{(X_{r3,G} - X_{r4,G})^2}}} \\ \Upsilon_G = \Upsilon_0 [(G_{max} - G + 1)/G_{max}]^p \end{cases} \tag{45}$$

where Υ_G, p, κ_1 and κ_2 are control parameters, G_{max} denotes the maximum generation number, $sgn(\cdot)$ represents the sign function.

V. CASE STUDY

In the case study, a six micro-grid system is utilized to verify the efficiency of proposed approach, each micro-grid consists of four CHP generators, three energy storages, one wind farm and system load. Those micro-grids are interconnected, and related data can be found in literature [32], [33]. The analysis can be classified into two parts: multiple indexes assessment and optimal recovery, assessment results mainly show the evaluation of each index, and also consists of security control scheme, optimal recovery results mainly present optimal Pareto fronts and optimal scheme of each micro-grid.

A. Multiple indexes assessment and security control of interconnected micro-grids

Since FDI from system load can destroy some indexes of micro-grid systems, the assessment on those indexes must be evaluated before defensive action, assessment results of economic cost, frequency, RoCoF, security level, voltage stability and emission rate are shown in Fig.2, where voltage is obtained with the average value of interconnected micro-grids, and the transmission loss of each micro-grid at tth time period does not exceed 8% of total output. In Fig.2, each index can converge well under different FDI effects with alternating uncertainty budget Δ, it can also be seen that large uncertainty of FDI has more damage on index value. After assessment, action strategy can also be made to reduce the security risk with on/off state of four CHP generators in each micro-grid, which are presented in Table.I, it can be seen that the on /off state of each CHP generator is presented (1 means on state and 0 means off). The security level after action strategy is shown in Fig.3, where security valve is set as 0.8. It can be seen that security index at some periods is lower than 0.8 after FDI, and security level is significantly improved after implementation of action strategy with proposed reinforcement learning method in comparison to RO.

B. Multi-objective Optimal recovery strategy

On the basis of above security control scheme, the remaining task is to recover other indexes such as economic cost, emission rate, frequency, RoCoF and voltage stability. The time step of base case scheduling is set as one hour, frequency dynamics and RoCoF is mainly taken into consideration within 1min at sharp time of hourly scheduling. Here, PBI based multi-objective algorithm is implemented to optimize these indexes simultaneously, and the obtained Pareto-optimal fronts are shown in Fig.4, where the results of proposed method and Chebyshev method [34] are presented with different uncertainty budgets. It can be seen that PBI based method can have better optimal results as well as have better diversity distribution. Here, scheme (10) of Pareto solutions with $\Delta = 1.5$ is taken for further analysis on optimal scheme, the uncertain parameters under $\Delta = 1.5$ is evenly taken for treating each micro-grid equally. In this

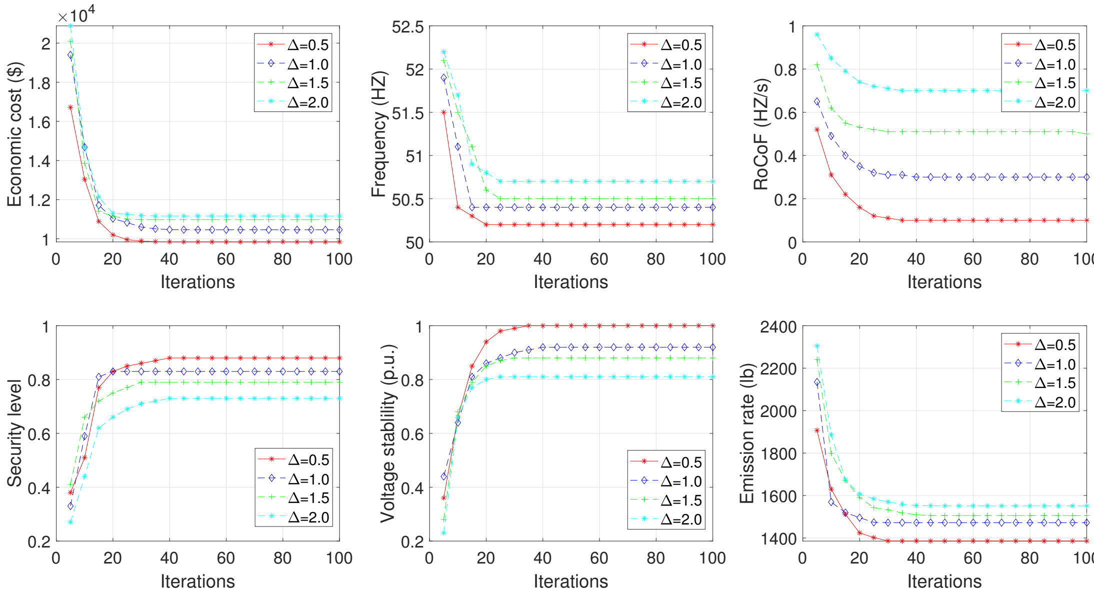


Fig. 2. The index evaluation with improved deep reinforcement learning

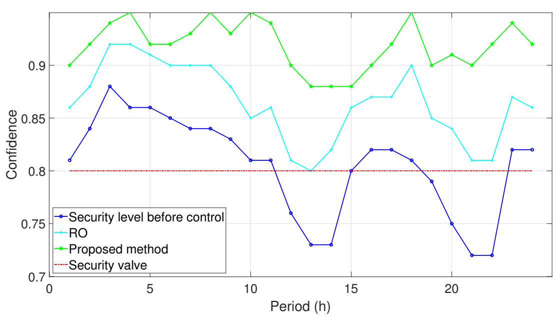


Fig. 3. The caused security risk and control by FDI

optimal scheme, the voltage distribution of each micro-grid is presented in Fig.5, where it can be seen that the voltage at each period is controlled in $[0.95, 1.05]$, which means that voltage security can be properly satisfied. The output process of CHPs in each micro-grid is shown in Fig.6, where it can be seen that 00:00-05:00 and 22:00-24:00 can be two output valleys, 10:00-13:00 and 19:00-20:00 are two peaks. Besides, the charging/discharging process of energy storages is also shown in Fig.7, where those periods with 0 charging/discharging state mean that this energy storage is out of work. It can be seen that charging state mainly occurs at 00:00-05:00, and discharging state occurs when load peak comes. The comparison of obtained results are presented in Table.II, where voltage stability is calculated with average value of voltage at each micro-grid. It can be seen that the proposed method can have better economic cost and emission rate with satisfying security level, voltage stability, frequency and RoCoF requirements in comparison to other existing alternatives.

VI. Conclusion

With consideration of resilience of interconnected micro-grids after FDI on system load, this paper proposes a TSK fuzzy system based deep reinforcement learning approach for assessment and security control on the interconnected micro-grids, and active defensive strategy is also made to recover interconnected micro-grids to normal operation state with improved decomposition based multi-objective differential evolution algorithm. According to those obtained simulation results, it can be verified that the proposed TSK based learning assessment strategy can be valid for uncertain input, it can also make correct security control scheme for improving security levels, and the optimal defensive strategy can recover interconnected micro-grids to normal state with low security risk and economic cost.

Acknowledgment

This work is supported in part by the National Natural Science Fund Under Grant 61973171,Basic research project of leading technology of Jiangsu Province under Grant (BK20202011), National Key R&D Program of China under Grant (2018YFA0702200), and National Natural Science Key Fund under Grant 61833008.

References

[1] Y. Wang, C. Chen, J. Wang, and R. Baldick, “Research on resilience of power systems under natural disasters—a review,” *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1604–1613, 2016.

[2] H. Zhang, D. Yue, X. Xie, S. Hu, and S. Weng, “Multi-elite guide hybrid differential evolution with simulated annealing technique for dynamic economic emission dispatch,” *Applied Soft Computing*, vol. 34, pp. 312 – 323, 2015.

TABLE I
THE ON/OFF STATE OF CHPS IN EACH MICRO-GRID

Period	Microgrid #1	Microgrid #2	Microgrid #3	Microgrid #4	Microgrid #5	Microgrid #6
1	{1,0,1,0}	{1,0,1,1}	{1,0,0,1}	{1,0,1,0}	{1,0,0,1}	{1,0,0,1}
2	{1,0,1,0}	{1,0,1,0}	{1,0,0,1}	{1,0,1,0}	{1,0,0,1}	{1,0,0,1}
3	{1,0,0,0}	{1,0,1,0}	{1,0,0,1}	{1,0,0,0}	{1,0,0,1}	{1,0,0,1}
4	{1,0,0,0}	{1,0,1,1}	{1,0,0,1}	{1,0,0,0}	{1,0,0,1}	{1,0,0,1}
5	{1,0,0,0}	{1,0,1,1}	{1,0,0,1}	{1,0,0,0}	{1,0,0,1}	{1,0,0,1}
6	{1,0,1,0}	{1,0,1,1}	{1,0,1,1}	{1,0,1,0}	{1,0,0,1}	{1,0,1,1}
7	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
8	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
9	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
10	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
11	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
12	{1,1,1,1}	{1,1,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
13	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
14	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
15	{1,0,1,0}	{1,0,1,1}	{1,0,1,1}	{1,0,1,0}	{1,0,1,1}	{1,0,1,1}
16	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,0}	{1,0,1,1}	{1,0,1,1}
17	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,0}	{1,0,1,1}	{1,0,1,1}
18	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
19	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
20	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
21	{1,0,1,1}	{1,1,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,1,1}
22	{1,0,1,1}	{1,1,1,1}	{1,0,1,1}	{1,0,1,1}	{1,0,0,1}	{1,0,1,1}
23	{1,0,0,0}	{1,0,1,1}	{1,0,0,1}	{1,0,1,0}	{1,0,0,1}	{1,0,0,1}
24	{1,0,0,0}	{1,0,1,1}	{1,0,0,1}	{1,0,1,0}	{1,0,0,1}	{1,0,0,1}

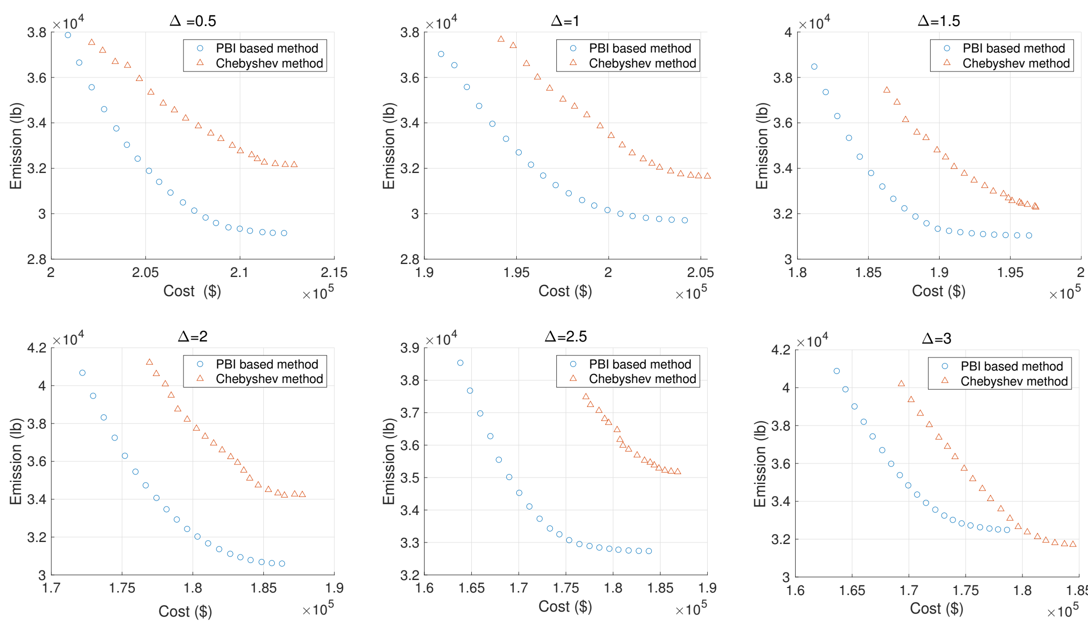


Fig. 4. The Pareto fronts with different FDI uncertainty budgets

TABLE II
THE COMPARISON OF OBTAINED RESULTS WITH OTHER ALTERNATIVES

Index	Ref. [29]	Ref. [30]	Ref. [31]	This method
Security level	0.87	0.87	0.88	0.92
Economic cost ($)	197743	198865	197685	188316
Voltage (p.u.)	0.96	0.98	0.97	0.99
Emission rate (lb)	33357	33226	32993	31878
Frequency (HZ)	50.14	50.15	50.10	50.02
RoCoF (HZ/s)	0.22	0.19	0.18	0.15

[3] M. Vahedipour-Dahraie, H. Rashidizadeh-Kermani, and A. Anvari-Moghaddam, "Risk-based stochastic scheduling of resilient microgrids considering demand response programs," *IEEE Systems Journal*, pp. 1–10, 2020.

[4] W. Yuan, J. Wang, F. Qiu, C. Chen, C. Kang, and B. Zeng, "Robust optimization-based resilient distribution network planning against natural disasters," *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2817–2826, 2016.

[5] T. Ding, Y. Lin, G. Li, and Z. Bie, "A new model for resilient distribution systems by microgrids formation," *IEEE Transactions on Power Systems*, vol. 32, no. 5, pp. 4145–4147, 2017.

[6] T. Lagos, R. Moreno, A. N. Espinosa, M. Panteli, R. Sacaan, F. Or-

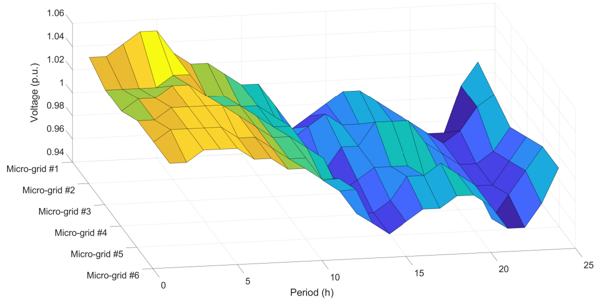


Fig. 5. The voltage of micro-grids system

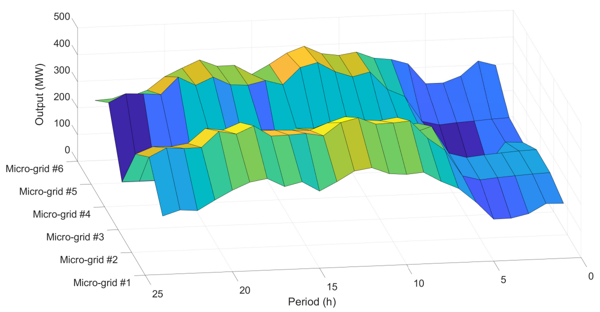


Fig. 6. The output of CHPs in each micro-grid

donez, H. Rudnick, and P. Mancarella, "Identifying optimal portfolios of resilient network investments against natural hazards, with applications to earthquakes," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1411–1421, 2020.

[7] J. Duan, W. Zeng, and M. Chow, "Resilient distributed dc optimal power flow against data integrity attack," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3543–3552, 2018.

[8] F. Yang, X. Feng, and Z. Li, "Advanced microgrid energy management system for future sustainable and resilient power grid," *IEEE Transactions on Industry Applications*, vol. 55, no. 6, pp. 7251–7260, 2019.

[9] T. Ding, M. Qu, Z. Wang, B. Chen, C. Chen, and M. Shahidehpour, "Power system resilience enhancement in typhoons using a three-stage day-ahead unit commitment," *IEEE Transactions on Smart Grid*, pp. 1–1, 2020.

[10] W. Zeng, Y. Zhang, and M. Chow, "Resilient distributed energy management subject to unexpected misbehaving generation units," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 208–216, 2017.

[11] M. R. Khalghani, J. Solanki, S. K. Solanki, M. H. Khooban, and A. Sargolzaei, "Resilient frequency control design for microgrids under false data injection," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 3, pp. 2151–2162, 2021.

[12] S. A. Arefifar, M. Ordonez, and Y. A. I. Mohamed, "Energy management in multi-microgrid systems—development and assessment," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 910–922, 2017.

[13] S. Khan, W. Gawlik, and P. Palensky, "Reserve capability assessment considering correlated uncertainty in microgrid," *IEEE Transactions on Sustainable Energy*, vol. 7, no. 2, pp. 637–646, 2016.

[14] P. Vorobev, P. Huang, M. Al Hosani, J. L. Kirtley, and K. Turitsyn, "High-fidelity model order reduction for microgrids stability assessment," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 874–887, 2018.

[15] J. Guo, W. Liu, F. R. Syed, and J. Zhang, "Reliability assessment of a cyber physical microgrid system in island mode," *CSEE Journal of Power and Energy Systems*, vol. 5, no. 1, pp. 46–55, 2019.

[16] E. Scolari, L. Reyes-Chamorro, F. Sossan, and M. Paolone, "A comprehensive assessment of the short-term uncertainty of grid-connected pv systems," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 3, pp. 1458–1467, 2018.

[17] V. Venkataramanan, A. Hahn, and A. Srivastava, "Cp-sam: Cyber-physical security assessment metric for monitoring microgrid resiliency," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1055–1065, 2020.

[18] P. Gautam, P. Piya, and R. Karki, "Resilience assessment of distribution systems integrated with distributed energy resources," *IEEE Transactions on Sustainable Energy*, pp. 1–1, 2020.

[19] C. Liu and Y. L. Murphey, "Optimal power management based on q-learning and neuro-dynamic programming for plug-in hybrid electric vehicles," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 6, pp. 1942–1954, 2020.

[20] W. Liu, P. Zhuang, H. Liang, J. Peng, and Z. Huang, "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, 2018.

[21] L. Zhu, C. Lu, Z. Y. Dong, and C. Hong, "Imbalance learning machine-based power system short-term voltage stability assessment," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 5, pp. 2533–2543, 2017.

[22] M. Sun, I. Konstantelos, and G. Strbac, "A deep learning-based feature extraction framework for system security assessment," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5007–5020, 2019.

[23] Z. Ni and S. Paul, "A multistage game in smart grid security: A reinforcement learning solution," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 9, pp. 2684–2695, 2019.

[24] Y. Zheng, Z. Yan, K. Chen, J. Sun, Y. Xu, and Y. Liu, "Vulnerability assessment of deep reinforcement learning models for power system topology optimization," *IEEE Transactions on Smart Grid*, pp. 1–1, 2021.

[25] Y. Wen, C. Y. Chung, X. Liu, and L. Che, "Microgrid dispatch with frequency-aware islanding constraints," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 2465–2468, 2019.

[26] H. Zhang, D. Yue, W. Yue, K. Li, and M. Yin, "Moea/d-based probabilistic pbi approach for risk-based optimal operation of hybrid energy system with intermittent power uncertainty," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–11, 2019.

[27] H. Zhang, D. Yue, C. Dou, X. Xie, and G. P. Hancke, "Two-stage optimal operation strategy of isolated microgrid with tsk fuzzy identification of supply security," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 3731–3743, 2020.

[28] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*, 2011.

[29] Y. Xiang, J. Liu, and Y. Liu, "Robust energy management of microgrid with uncertain renewable generation and load," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 1034–1043, March 2016.

[30] W. Zhang, Y. Xu, Z. Dong, and K. P. Wong, "Robust security constrained-optimal power flow using multiple microgrids for corrective control of power systems under uncertainty," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 1704–1713, Aug 2017.

[31] H. Zhang, D. Yue, C. Dou, K. Li, and X. Xie, "Event-triggered multiagent optimization for two-layered model of hybrid energy system with price bidding-based demand response," *IEEE Transactions on Cybernetics*, pp. 1–12, 2019.

[32] J. Aghaei, T. Niknam, R. Azizipanah-Abarghooee, and J. M. Arroyo, "Scenario-based dynamic economic emission dispatch considering load and wind power uncertainties," *International Journal of Electrical Power & Energy Systems*, vol. 47, pp. 351 – 367, 2013.

[33] H. Zhang, D. Yue, X. Xie, C. Dou, and F. Sun, "Gradient decent based multi-objective cultural differential evolution for short-term hydrothermal optimal scheduling of economic emission with integrating wind power and photovoltaic power," *Energy*, vol. 122, pp. 748 – 766, 2017.

[34] X. Ma, Q. Zhang, G. Tian, J. Yang, and Z. Zhu, "On tchebycheff decomposition approaches for multiobjective evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 2, pp. 226–244, 2018.

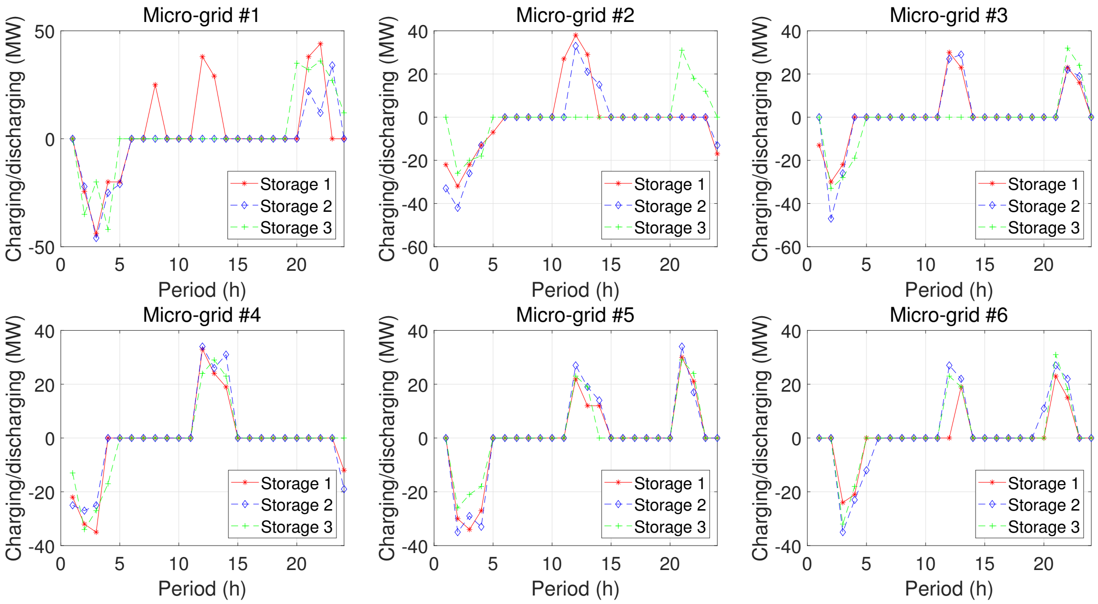
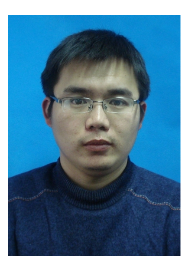


Fig. 7. The charging/discharging process of energy storages in each micro-grid

Huifeng Zhang (M'16) received Ph.D. degree from Huazhong University of Science and Technology, Wuhan, China, in 2013. From 2014 to 2016, he was a Post-Doctoral Fellow with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China. From 2017 to 2018, He was granted as visiting research fellow by China Scholarship Council to study in Queen's University Belfast and University of Leeds, UK. He is currently an Associate Professor at the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China. His current research interest includes electrical power management, optimal operation of power system, distributed optimization, and multi-objective optimization.

Chunxia Dou (M'18) received the B.S. and M.S. degrees in automation from the Northeast Heavy Machinery Institute, Qiqihaer, China, in 1989 and 1994, respectively, and the Ph.D. degree in Institute of Electrical Engineering from Yanshan University, Qinhuangdao, China, in 2005. In 2010, she joined the Department of Engineering, Peking University, Beijing, China, where she was a Post-doctoral Fellow for two years. Since 2005, she has been a Professor in Institute of Electrical Engineering, Yanshan University. Her current research interests include multi-agent based control, event-triggered hybrid control, distributed coordinated control, and multi-mode switching control and their applications in power systems, Microgrids and smart grids.

Dong Yue (SM'08-F'20) received the Ph.D. degree from the South China University of Technology, Guangzhou, China, in 1995. He is currently a Professor and the Dean with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China, and also a Changjiang Professor with the Department of Control Science and Engineering, Huazhong University of Science and Technology, Wuhan, China. His current research interests include analysis and synthesis of networked control systems, multiagent systems, optimal control of power systems, and internet of things. He has published over 100 papers in international journals, domestic journals, and international conferences. Prof. Yue is currently an Associate Editor of the IEEE Control Systems Society Conference Editorial Board and the International Journal of Systems Science.



Xiangpeng Xie (M'18) received the B.S. and Ph.D. degrees from Northeastern University, Shenyang, China, in 2004 and 2010, respectively, both in engineering. From 2012 to 2014, he was a Post-Doctoral Fellow with the Department of Control Science and Engineering, Huazhong University of Science and Technology, Wuhan, China. He is currently a Professor with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China. His current research interests include fuzzy modeling and control synthesis, state estimations, optimization in process industries, and intelligent optimization algorithms.

Kang Li (M'05–SM'11) received the B.Sc. degree in Industrial Automation from Xiangtan University, Hunan, China, in 1989, the M.Sc. degree in Control Theory and Applications from Harbin Institute of Technology, Harbin, China, in 1992, and the Ph.D. degree in Control Theory and Applications from Shanghai Jiaotong University, Shanghai, China, in 1995. He also received D.Sc. degree in Engineering from Queen's University Belfast, UK, in 2015. He currently holds the Chair of Smart Energy Systems at the University of Leeds, UK. His research interests cover nonlinear system modelling, identification, and control, and artificial intelligence, with substantial applications to energy and power systems, smart grid, electric vehicles, railway systems, and energy management in energy intensive manufacturing processes.

Gerhard P. Hancke (M'88–SM'00–F'16) received the B.Sc. and B.Eng. degrees, in1970, and the M.Eng. degree in Electronic Engineering, in 1973, from the University of Stellenbosch, South Africa, and the Ph.D. degree from the University of Pretoria, South Africa, in 1983. He is a Professor with the University of Pretoria, South Africa and recognized internationally as a pioneer and leading scholar in industrial wireless sensor networks research. He initiated and co-edited the first Special Section on Industrial Wireless Sensor Networks in the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS in 2009 and the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS in 2013. He co-edited a textbook, Industrial Wireless Sensor Networks: Applications, Protocols and Standards (2013), the first on the topic. Prof. Hancke has been serving as an Associate Editor and Guest Editor for the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE ACCESS, and previously the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS. Currently, he is a Co-Editor-in-Chief for the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.