Resilient Optimal Defensive Strategy of TSK Fuzzy-Model-Based Microgrids' System via a Novel Reinforcement Learning Approach

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

***Abstract*— With consideration of false data injection (FDI) on the demand side, it brings a great challenge for the optimal defensive strategy with the security issue, voltage stability, power flow, and economic cost indexes. This article proposes a Takagi–Sugeuo–Kang (TSK) fuzzy system-based reinforcement learning approach for the resilient optimal defensive strategy of interconnected microgrids. Due to FDI uncertainty of the system load, TSK-based deep deterministic policy gradient (DDPG) is proposed to learn the actor network and the critic network, where multiple indexes' assessment occurs in the critic network, and the security switching control strategy is made in the actor network. Alternating direction method of multipliers (ADMM) method is improved for policy gradient with online coordination between the actor network and the critic network learning, and its convergence and optimality are proved properly. On the basis of security switching control strategy, the penalty-based boundary intersection (PBI)-based multiobjective optimization method is utilized to solve economic cost and emission issues simultaneously with considering voltage stability and 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 problems on interconnected microgrids.**

***Index Terms*— Microgrids, reinforcement learning (RL), resilient optimal defensive, Takagi–Sugeuo–Kang (TSK) fuzzy system.**

I. Introduction

THE resilience in the power system is generally defined as the ability of the power system to withstand severe disturbances without experiencing any large disruption and further enabling a quick recovery to the normal operation state [1]–[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, a resilient operation strategy can be a good choice for avoiding those risks. In fact, some studies on resilient operation have been taken in many existing studies, where resilient operation strategies were proposed with considering natural disasters, misbehaving of power generators, and false data injection (FDI) [4]–[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 the highest hedge against the risk of earthquake. Literature [9] proposed a three-stage resilient unit commitment model considering typhoon paths and line outages, and the optimal solution was deduced with worst case scenario of possible typhoon paths. In literature [11], the two-layer optimal augmented controller with observer mitigates cyberdisruptions, and it also models an intelligent type of FDI effect on cyber–physical interconnected microgrids. However, those strategies lack 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]–[18]. Literature [13] proposed a novel assessment approach of required reserve capability for meeting forecasting uncertainty dynamics in the microgrid. Literature [14] proposed an efficient and accurate model of inverter-based microgrids with reduced order; the improved model accounts for the effects of network dynamics and is similar to the quasi-stationary model. In literature [16], a comprehensive method was proposed to assess the uncertainty characteristics of overall photovoltaic resources with the gray-box model, which consists of physical and data-driven submodels with currents, and voltage information of photovoltaic resources. Literature [18] integrated impact assessment model and optimal restoration model of active distribution network together with nonsequential Monte Carlo simulation framework. Here, the reinforcement learning (RL) method is utilized to assess

Manuscript received March 23, 2021; revised June 20, 2021; accepted August 14, 2021. This work was supported in part by the National Natural Science Fund Under Grant 61973171, in part by the Basic Research Project of Leading Technology of Jiangsu Province under Grant BK20202011, in part by the National Key Research and Development Program of China under Grant 2018YFA0702200, and in part by the National Natural Science Key Fund under Grant 61833008. *(Corresponding authors: Dong Yue; Chunxia Dou.)*

Huifeng Zhang, Dong Yue, Chunxia Dou, and Xiangpeng Xie are with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210023, China (e-mail: zhanghuifeng_520@163.com; medongy@vip.163.com; cxdou@ysu.edu.cn; xiexiangpeng1953@163.com).

Kang Li is with the School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, U.K. (e-mail: k.li1@leeds.ac.uk).

Gerhardus P. Hancke is with the College of Automation, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210023, China, and also with the Department of Computer Engineering, University of Pretoria, Pretoria 0028, South Africa (e-mail: g.hancke@ieee.org).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TNNLS.2021.3105668.

Digital Object Identifier 10.1109/TNNLS.2021.3105668

security risk and economic indexes of interconnected microgrids, as it is more powerful than other alternatives due to its model-free strategy and knowledge-cumulative mechanisms [19], [20]. Literature [21] developed a short-term voltage stability assessment strategy with an online systematic imbalance learning machine. In literature [22], a novel deep learning-based feature extraction framework was developed for creating security rules of the electricity system operation. Literature [23] proposed a multistage game between attacker and the defender with RL to identify optimal attack sequences given certain objectives. Literature [24] presented deep RL models for vulnerability assessment of power system topology optimization with considering data perturbation and cyberattack. Those existing learning-based assessment methods can be efficient for dealing with the deterministic issue, while it cannot tackle uncertain input problems.

Since the smart meter on the demand side can be affected by different factors, such as electricity burglary action, device fault, and cyberattack, these factors can lead to 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 problems. The resilience of interconnected microgrids mainly lies in two parts: robustness for uncertain input and recovery from FDI. In this article, a Takagi–Sugeuo–Kang (TSK) fuzzy system-based learning strategy is improved to solve the uncertain input problem, and the multiobjective optimal defensive strategy is also proposed to recover the power system into a normal operation state. The main contribution of this article can be summarized as follows.

1) This article proposes an optimal recovery model with frequency-awareness under FDI on the demand side. Since FDI of the system load can affect the operation scheme of interconnected microgrids and even lead to security risks, frequency-awareness limits with rate-of-change of frequency (RoCoF) have been taken into consideration.
2) This article improves a TSK fuzzy system-based deep RL-based strategy for assessment on multiple indexes of interconnected microgrids. With consideration of the uncertainty nature of FDI on the demand side, the TSK fuzzy system is modeled for uncertain input of load parameters, accumulation knowledge with the actor network and the critic network learning is utilized to guide multiple indexes assessment with the alternating direction method of multipliers (ADMM) algorithm, and the security control strategy is also made to reduce security risk with actor network learning.
3) This article utilizes a decomposition-based multiobjective differential evolution algorithm for optimal recovery of interconnected microgrids. On the basis of assessment and security control of interconnected microgrids, the interconnected microgrids require recovering normal operation state with low security risk and economic cost; a decomposition-based multiobjective differential evolution is utilized to obtain optimal scheme for the recovery of interconnected microgrids.

The arrangement of article structure is presented as follows. Problem formulation is presented in Section II. The proposed assessment method is described in Section III. An active defensive strategy is introduced in Section IV. The simulation results and conclusion are shown in Sections V and VI.

II. Problem Formulation of Assessment Analysis and Optimal Defensive Model

With consideration of the FDI attack, the measurement for enhancing interconnected microgrids 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 microgrids, and then, the optimal recovery procedure controls microgrids to the normal state.

A. Multiple Indexes of Interconnected Microgrids

The security issue, economic cost, and environmental protection can be the most important metrics for interconnected microgrids, 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

$$F_1 = \text{Security}_{\text{con}} = \text{Prob}\left(\sum_{n \in N_G} \alpha_n | P_{G,n,t} + P_{S,n,t} - P_{\text{load},n,t} - P_{\text{loss},n,t} | < \epsilon_t \right) \quad (1)$$

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

$$P_{\text{loss},n,t} = \sum_{n_1 \in N_{\text{node}}} d_{n_1,n} R_{n_1,n} \frac{P_{n_1,n}^2 + Q_{n_1,n}^2}{U_{n_1}^2} \quad (2)$$

where $d_{n_1,n}$ and $R_{n_1,n}$ represent the distance and resistance between the n_1th microgrid and the nth microgrid, U_{n_1} is the rating line voltage, and $P_{n_1,n}$ and $Q_{n_1,n}$ denote the power flow from the n_1th microgrid to the nth microgrid.

2) Voltage Stability: The voltage index $F_2 = 1/N_G \sum_{n=1}^{N_G} U_n(t)$ of each microgrid must be controlled in security domain as follows:

$$U_{n,\min} \le U_n(t) \le U_{n,\max} \quad (3)$$

where $U_{n,\min}$ and $U_{n,\max}$ represent the minimum and maximum voltages at the nth microgrid.

3) Frequency Stability: Since interconnected microgrids 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:

$$\begin{cases} F_4 = \mathrm{RoCoF}_t = \left(\sum_{n\in N_G} \triangle P_{G,n,t} + \sum_{n\in N_G} \triangle P_{S,n,t} \right. \\ \qquad\qquad \left. -D \triangle f(t) - P_{M,t} \right) / 2H_t \\ -\mathrm{RoCoF}^{\max} \le \mathrm{RoCoF}_t \le \mathrm{RoCoF}^{\max} \end{cases} \tag{4}$$

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

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

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

where $H_{n,t}$ represents the ON/OFF state of the nth microgrid 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, and $P_{\mathrm{CHP},n,\min}$ is the minimum limit of CHP output.

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

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

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

B. Constraint Limits of Optimal Active Defensive Model

1) Power Flow Constraints:: The power flow among different microgrids can be described as follows:

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

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

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

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

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

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

2) Output Constraint Limit: The output at each microgrid 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 $P_{\mathrm{CHP},n,t} + P_{w,n,t} + P_{v,n,t}$, $P_{\mathrm{CHP},n,t}$ represents the output of combined heat and power generator, and $P_{w,n,t}$ and $P_{v,n,t}$ denote the output of wind power and PV generator. The above output must satisfy following limits:

$$\begin{cases} P_{\mathrm{CHP},n,\min} \le P_{\mathrm{CHP},n,t} \le P_{\mathrm{CHP},n,\max} \\ \mathrm{Ramp}_{\mathrm{down},n} \le P_{\mathrm{CHP},n,t} - P_{\mathrm{CHP},n,t-1} \le \mathrm{Ramp}_{\mathrm{up},n} \end{cases} \tag{11}$$

where $P_{\mathrm{CHP},n,\min}$ and $P_{\mathrm{CHP},n,\max}$ represent the minimum and maximum outputs of the CHP generator, and $\mathrm{Ramp}_{\mathrm{down},n}$ and $\mathrm{Ramp}_{\mathrm{up},n}$ denote the ramp down and ramp up of the CHP generator.

3) Limits of Energy Storage: The energy storage can be a reliable energy resource for ensuring interconnected microgrids' system stability, and $P_{S,n,t}$ can be considered as $\sum_{r=1}^{N_{n,\mathrm{ESS}}} P_{S,n,r,t}$, which represents the summation of provided power of $N_{n,\mathrm{ESS}}$ energy storages in r microgrid. With consideration of charging or discharging state, $P_{S,n,r,t}$ can be taken as $P_{S,n,r,t}^{\mathrm{dis}}$ if energy storage is discharging; else, it can be taken as $P_{S,n,r,t}^{\mathrm{cha}}$, and the charging/discharging process must satisfy following limits:

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

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

III. Learning With TSK Fuzzy Approach for Security Assessment and Switching Control of Interconnected Microgrids

With consideration of FDI on the demand side, it can cause a large deviation between actual system load and observed data, which may cause security issues when observed data are directly taken for optimal operation of interconnected microgrids. For system load $P_{\mathrm{load},n,t}$ at node n, it can be described as

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

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

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

Hence, security assessment is required to evaluate the system safety, the TSK fuzzy-model-based deep deterministic policy gradient (DDPG) learning strategy is developed to tackle this problem, and its framework is presented in Fig. 1. Some definitions are defined as follows.

State Set: The state set can be defined as $s_t = [\xi_{1,t}, \xi_{2,t}, \ldots, \xi_{N_G,t}]$; those variables has the main effect on those security and economic indexes.

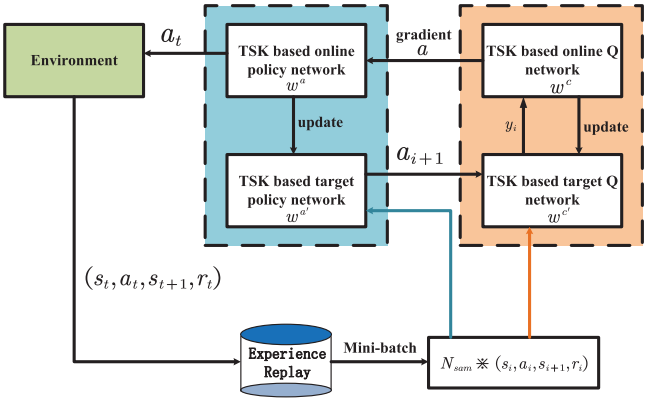


Fig. 1. TSK model-based learning strategy with the DDPG approach.

Action Set: The action is mainly implemented to prevent decreasing of maximization index and increasing of minimization index, and the action set can be defined as the set of ON/OFF state of power generators, which can be described as $a_t = [H_{1,t}, H_{2,t}, \ldots, H_{N_G,t}]$.

Reward: During the learning process, state and action must be rewarded or punished according to the effect on those indexes, and the 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 $-F_i$. Specifically, if the 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. TSK Fuzzy Model-Based Network Learning With FDI Uncertain Input

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

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

where $x_j (j = 1, 2 \ldots, N_G)$ represents the jth input variable, $A_{i1}, A_{i2}, \ldots, A_{iN_G}$ denote the fuzzy sets, q_i is the output of the 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

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

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

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

The output of the fuzzy system can be calculated with weighted defuzzification 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 the 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, and c_{1j} and κ_j represent the control vector and parameter of the RBF function. With consideration of overfitting and generalization issue of each y_i, the following L_2 regularized loss function is created:

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

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

B. Actor Network and Critic Network Learning With Policy Gradient Action

With consideration of uncertain income, the output function of the above TSK fuzzy model can convert the uncertain vector into the deterministic vector, and the pseudocode of network learning is shown in Algorithm 1. Since the critic network mainly evaluates the current state and action, it can be considered as the function of state and action vector. As it is known in RL, the state-action value function can be deduced with the Bellman theory as follows:

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

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 the kth step. The optimal action of $Q^k(s,a)$ can be trained with actor network $a(s) = g_i(s)$, which represents the TSK-based network. The loss function can be described with L_{λ_1} as

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

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

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

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

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

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 the critic network according to this approach. For ensuring the learning efficiency, the Lagrangian function can be created as follows:

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

where R_n represents the reward value of the nth sample, $s^{(n)}$ and $a^{(n)}$ are the state and action values of the nth sample, and 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. Online Coordinated Policy Gradient Between Actor Network and Critic Network With ADMM Approach

Since the action vector in the critic network is deduced with actor network learning, coordination between the actor network and the critic network can promote optimization efficiency. Here, the ADMM approach is developed to coordinate above two models. Suppose that $\chi_1(w^a) = \mu_1 L_{\lambda_1}(w^a)$ and $\chi_2(w^c, z) = \mu_2 L_{\lambda_2}(w^c, z)$, where $w^a = [w_{i1}^a, w_{i2}^a, \ldots, w_{iN_G}^a]^T$, $w^c = [w_{i1}^a, w_{i2}^a, \ldots, w_{iN_G}^a]^T$, $z = [z_{i1}^a, z_{i2}^a, \ldots, z_{iN_G}^a]^T$, and μ_1 and μ_2 are two weight parameters. The augmented Lagrangian problem can be equalized as follows:

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

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

$$\begin{cases} w^{ak+1} = \arg\min_{w^a} \left[\chi_1(w^a) + \frac{\rho}{2} \left\| w^a - z^k + u^k \right\|_2^2 \right] \\ z^{k+1} = \arg\min_z \left[\chi_2(w^{ck}, z) + \frac{\rho}{2} \left\| w^{ak+1} - z + u^k \right\|_2^2 \right] \\ u^{k+1} = u^k + \left(w^{ak+1} - z^{k+1} \right). \end{cases} \tag{27}$$

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

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

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

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

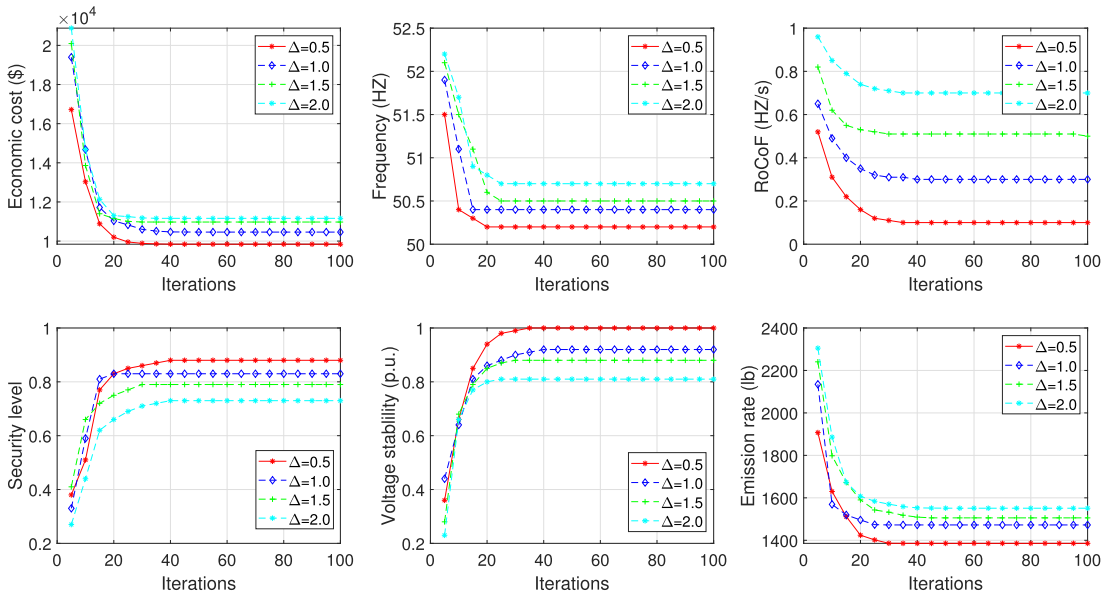


Fig. 2. Index evaluation with improved deep RL.

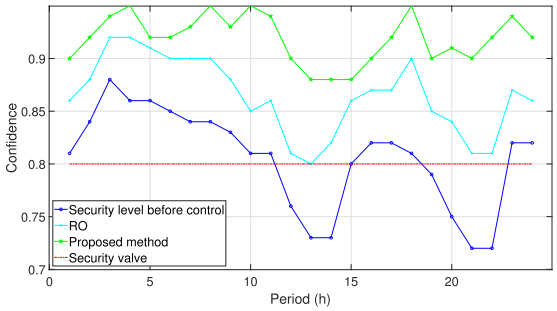


Fig. 3. Caused security risk and control by FDI.

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

$$\| r^k \|_2 \leq \epsilon^{\text{pri}} \tag{30}$$

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

$$\epsilon^{\text{pri}} = \sqrt{N_G}\epsilon^{abs} + \epsilon^{\text{rel}} \max\{\| w^a \|_2, \| z \|_2\} \tag{31}$$

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

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

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

Algorithm 1 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 $w^{c\prime}$ and $w^{a\prime}$, $ReplayBuffer = \phi$, $episode = 1$;
  Collect input parameter $[\xi_{1,t}, \xi_{2,t}, \ldots, \xi_{N_G,t}]$;
  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 $w^{c\prime}$ and $w^{a\prime}$ of target networks;
    $episode = episode + 1$;
  end while
end procedure

D. Analysis on Convergence and Optimality of Proposed Algorithm

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

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 $\chi_1(w^a) = \{(w^a, \Omega) \in R^n \times R | \chi_1(w^a) \leq \Omega\}$ and $\chi_2(w^c, z) = \{(z, \Omega) \in R^n \times R | \chi_2(w^c, z) \leq \Omega\}$.

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

TABLE I
ON/OFF STATE OF CHPS IN EACH MICROGRID

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}

described as

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

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

$$V^k = (1/\rho)\big\|y^k - y^*\big\|_2^2 + \rho\big\|z^k - z^*\big\|_2^2. \tag{34}$$

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 as follows:

1) 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 that $p^k = \chi_1(w^a) + \chi_2(w^{ck}, z)$ and $p^* = \chi_1(w^{a*}) + \chi_2(w^{c*}, z^*)$. As w^{ck} minimizes L_ρ, it satisfies the necessary and sufficient condition

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

It can also mean that

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

Specifically, 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

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

It can be described as

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

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

2) 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 $\max\{\Psi^T\Psi\}$, and $\Psi = [\psi_1(s), \ldots, \psi_{N_G}(s), \psi_{N_G+1}(a), \ldots, \psi_{N_G+N_a}(a)]^T$

$$\begin{aligned}\Big\|\chi_2\Big(w^{ck}, z\Big) - \chi_2\big(w^{c*}, z\big)\Big\|_2^2 &\le \eta^{c2}\Psi^T\Psi\big\|w^{ck} - w^{c*}\big\|_2^2 \\ &\le \delta_c^2/\big(K^2\big) \le \epsilon_c\end{aligned} \tag{40}$$

where $\epsilon_c > 0$ denotes the accuracy parameter, $\delta_c = \max\{\| w^{ck} - w^{c*} \|_2\}$, and K can take $\lceil(\delta_c/\sqrt{\epsilon_c})\rceil$; then, the convergence holds. ■

IV. OPTIMAL ACTIVE DEFENSIVE STRATEGY OF INTERCONNECTED MICROGRIDS

A. 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 microgrids. However, the switching strategy can merely tackle with one security or economic issues; further defensive strategy must be made for ensuring system security and minimizing economic cost. The supply security index can be considered as equality constraint, and the voltage stability and frequency stability indexes can be treated as inequality constraint limits. Hence, the active defensive

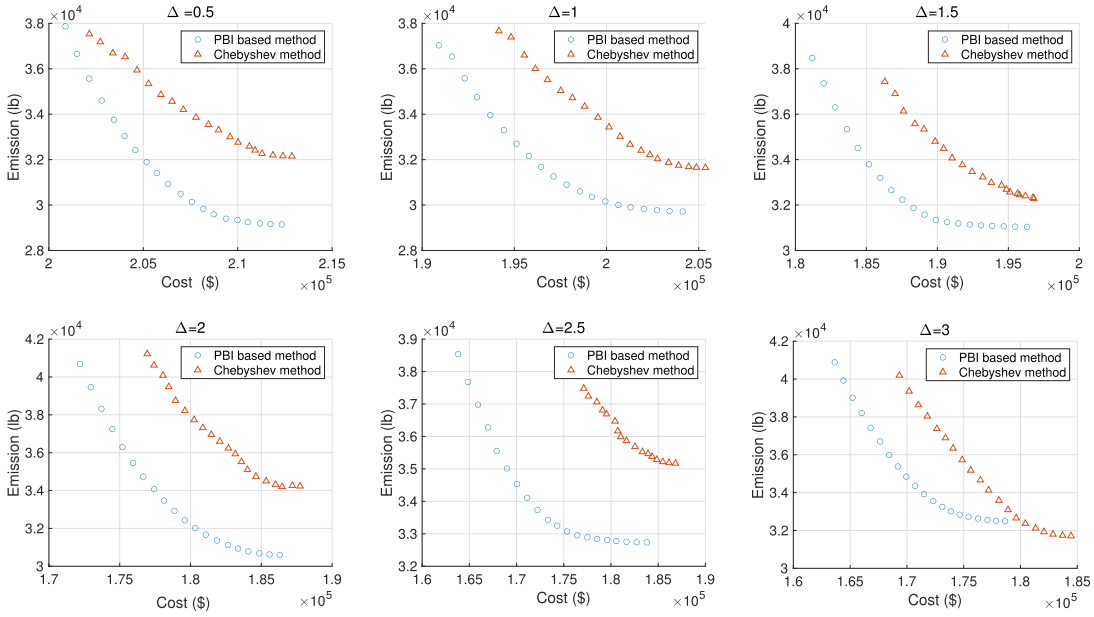


Fig. 4. Pareto fronts with different FDI uncertainty budgets.

model can be created as follows:

$$\begin{cases} \min \left\{ \sum_{t=1}^{T} \mathrm{Eco}_t, \sum_{t=1}^{T} \mathrm{Emi}_t \right\} \\ \text{s.t.} \quad \sum_{n\in N_G} P_{G,n,t} + \sum_{n\in N_G} P_{S,n,t} - \sum_{n\in N_G} \\ \quad P_{\mathrm{loss},n,t} = \sum_{n\in N_G} P_{\mathrm{load},n,t} \\ \quad V_{n,\min} \le V_{n,t} \le V_{n,\max} \\ \quad -\mathrm{RoCoF}^{\max} \le \mathrm{RoCoF}_t \le \mathrm{RoCoF}^{\max} \\ \quad \text{Power flow constraints;} \\ \quad \text{Output constraints;} \\ \quad \text{The limits of energy storage.} \end{cases} \tag{41}$$

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

B. Decomposition-Based Multiobjective Optimization Approach for Optimal Recovery of Microgrid

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:

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

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

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

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, and Ω denotes the feasible domain, which is determined by equality and inequality constraint limits. λ^i represents the direction vector, and its component λ_j^i satisfies $\sum_{j=1}^{m} \lambda_j^i = 1$ ($\lambda_j^i \ge 0$).

It is obvious that the optimal model can be converted into the above version, which optimizes several single-objective subproblems with alternating weights. To properly solve each subproblem, gradient descent-based differential evolution is utilized to enhance the searchability. Here, differential evolution procedure is taken with mutation operator of improved DE/rand/1/bin strategy, which can be described as follows:

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

where $X_{r,G}$, $X_{r1,G}$, $X_{r2,G}$, $X_{r3,G}$, and $X_{r4,G}$ are randomly selected individuals from nondominated solutions, X_{G+1}^j is the generated individual for $G+1$ generation, and β_1^j and β_2^j represent the control parameters. With consideration of

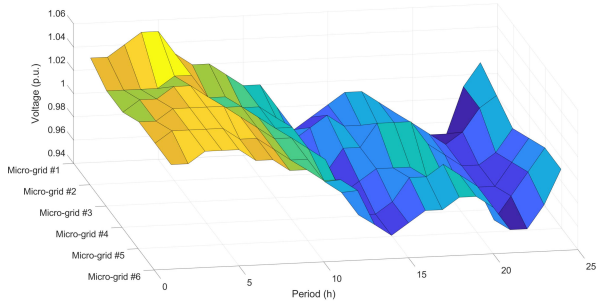


Fig. 5. Voltage of the microgrids' system.

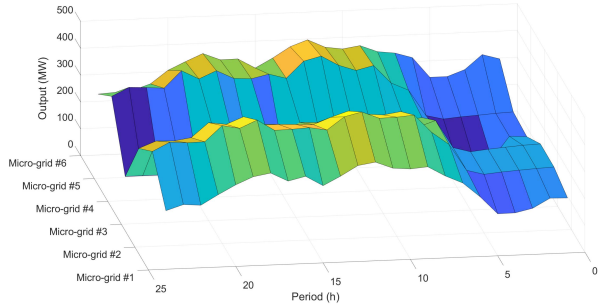


Fig. 6. Output of CHPs in each microgrid.

convergence ability, β_1^j and β_2^j can be updated as follows:

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

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

V. Case Study

In the case study, a six microgrid system is utilized to verify the efficiency of the proposed approach, each microgrid consists of four CHP generators, and three energy storages, one wind farm, and system load. Those microgrids are interconnected, and related data can be found in the 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 microgrid.

A. Multiple Indexes Assessment and Security Control of Interconnected Microgrids

Since FDI from the system load can destroy some indexes of microgrid systems, the assessment on those indexes must be evaluated before defensive action, and 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 microgrids, and the transmission loss of each microgrid at the tth time period does not exceed 8% of total output. In Fig. 2,

TABLE II
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

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 the index value. After assessment, action strategy can also be made to reduce the security risk with the ON/OFF state of four CHP generators in each microgrid, 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 the security valve is set as 0.8. It can be seen that the security index at some periods is lower than 0.8 after FDI, and the security level is significantly improved after the implementation of the action strategy with the proposed RL method in comparison to RO.

B. Multiobjective Optimal Recovery Strategy

On the basis of the 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 1 h, and frequency dynamics and RoCoF are mainly taken into consideration within 1 min at the sharp time of hourly scheduling. Here, the PBI-based multiobjective algorithm is implemented to optimize these indexes simultaneously, and the obtained Pareto-optimal fronts are shown in Fig. 4, where the results of the proposed method and the Chebyshev method [34] are presented with different uncertainty budgets. It can be seen that the PBI-based method can have better optimal results and have better diversity distribution. Here, scheme (10) of Pareto solutions with $\Delta = 1.5$ is taken for further analysis on the optimal scheme, and the uncertain parameters under $\Delta = 1.5$ are evenly taken for treating each microgrid equally. In this optimal scheme, the voltage distribution of each microgrid 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 microgrid 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 and 10:00–13:00 and 19:00–20:00 are two peaks. Besides, the charging/discharging process of energy storage 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 the charging state mainly occurs at 00:00–05:00, and the discharging state occurs when load peak comes. The comparison of obtained results is presented in Table II, where the voltage stability is calculated with an average value of the voltage at each microgrid. It can be seen that the proposed method can have better economic cost and emission rate with a satisfying security level, voltage stability, frequency, and RoCoF requirements in comparison to other existing alternatives.

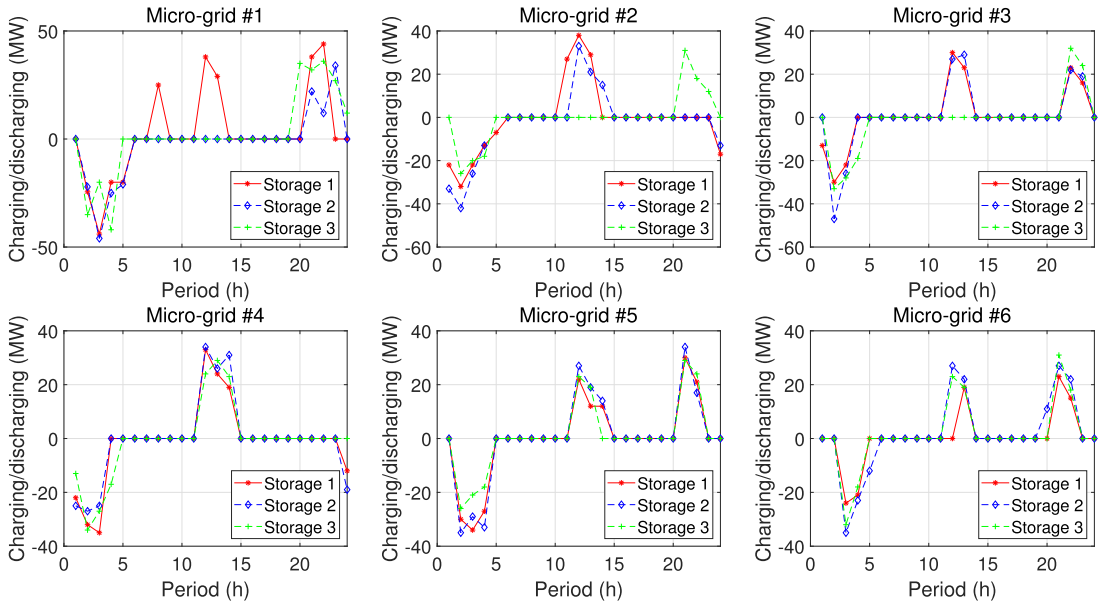


Fig. 7. Charging/discharging process of energy storages in each microgrid.

VI. Conclusion

With consideration of resilience of interconnected microgrids after FDI on system load, this article proposes a TSK fuzzy system-based deep RL approach for assessment and security control on the interconnected microgrids, and the active defensive strategy is also made to recover interconnected microgrids to normal operation state with the improved decomposition-based multiobjective 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 microgrids to the normal state with low security risk and economic cost.

References

[1] Y. Wang, C. Chen, J. Wang, and R. Baldick, "Research on resilience of power systems under natural disasters—A review," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1604–1613, Mar. 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," *Appl. Soft Comput.*, vol. 34, pp. 312–323, Sep. 2015.

[3] M. Vahedipour-Dahraie, H. Rashidizadeh-Kermani, and A. Anvari-Moghaddam, "Risk-based stochastic scheduling of resilient microgrids considering demand response programs," *IEEE Syst. J.*, vol. 15, no. 1, pp. 971–980, Mar. 2021.

[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 Trans. Smart Grid*, vol. 7, no. 6, pp. 2817–2826, Nov. 2016.

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

[6] T. Lagos *et al.*, "Identifying optimal portfolios of resilient network investments against natural hazards, with applications to earthquakes," *IEEE Trans. Power Syst.*, vol. 35, no. 2, pp. 1411–1421, Mar. 2020.

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

[8] F. Yang, X. Feng, and Z. Li, "Advanced microgrid energy management system for future sustainable and resilient power grid," *IEEE Trans. Ind. Appl.*, vol. 55, no. 6, pp. 7251–7260, Nov. 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 Trans. Smart Grid*, vol. 12, no. 3, pp. 2153–2164, May 2020.

[10] W. Zeng, Y. Zhang, and M.-Y. Chow, "Resilient distributed energy management subject to unexpected misbehaving generation units," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 208–216, Feb. 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 Trans. Ind. Electron.*, vol. 68, no. 3, pp. 2151–2162, Mar. 2021.

[12] S. Ali Arefifar, M. Ordonez, and Y. A.-R. I. Mohamed, "Energy management in multi-microgrid systems—Development and assessment," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 910–922, Mar. 2017.

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

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

[15] J. Guo *et al.*, "Reliability assessment of a cyber physical microgrid system in island mode," *CSEE J. Power Energy Syst.*, 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 Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1458–1467, Jul. 2018.

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

[18] P. Gautam, P. Piya, and R. Karki, "Resilience assessment of distribution systems integrated with distributed energy resources," *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 338–348, Jan. 2021.

[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 Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 6, pp. 1942–1954, Jun. 2020.

[20] W. Liu, P. Zhuang, H. Liang, J. Peng, and Z. Huang, "Distributed economic dispatch in microgrids based on cooperative reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2192–2203, Jun. 2018.

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

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

[23] Z. Ni and S. Paul, "A multistage game in smart grid security: A reinforcement learning solution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2684–2695, Sep. 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 Trans. Smart Grid*, vol. 12, no. 4, pp. 3613–3623, Jul. 2021.
[25] Y. Wen, C. Y. Chung, X. Liu, and L. Che, "Microgrid dispatch with frequency-aware islanding constraints," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 2465–2468, May 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 Trans. Syst., Man, Cybern. Syst.*, vol. 51, no. 4, pp. 2080–2090, Apr. 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 Trans. Ind. Informat.*, vol. 16, no. 6, pp. 3731–3743, Jun. 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," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
[29] Y. Xiang, J. Liu, and Y. Liu, "Robust energy management of microgrid with uncertain renewable generation and load," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 1034–1043, Mar. 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 Trans. Ind. Informat.*, 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 Trans. Cybern.*, vol. 51, no. 4, pp. 2068–2079, Apr. 2021.
[32] J. Aghaei, T. Niknam, R. Azizipanah-Abarghooee, and J. M. Arroyo, "Scenario-based dynamic economic emission dispatch considering load and wind power uncertainties," *Int. J. Electr. Power Energy Syst.*, vol. 47, pp. 351–367, May 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, Mar. 2017.
[34] X. Ma, Q. Zhang, G. Tian, J. Yang, and Z. Zhu, "On Tchebycheff decomposition approaches for multiobjective evolutionary optimization," *IEEE Trans. Evol. Comput.*, vol. 22, no. 2, pp. 226–244, Apr. 2018.

Huifeng Zhang (Member, IEEE) received the Ph.D. degree from the 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 a Visiting Research Fellow by the China Scholarship Council to study at Queen's University Belfast, Belfast, U.K., and the University of Leeds, Leeds, U.K. He is currently an Associate Professor with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications. His current research interests include electrical power management, optimal operation of power systems, distributed optimization, and multiobjective optimization.

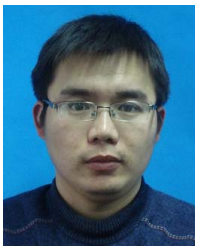
Dong Yue (Fellow, IEEE) received the Ph.D. degree from the South China University of Technology, Guangzhou, China, in 1995.

He is currently a Professor and the Dean of the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China, and a Changjiang Professor with the Department of Control Science and Engineering, Huazhong University of Science and Technology, Wuhan, China. He has published over 100 articles in international journals, domestic journals, and international conferences. His current research interests include the analysis and synthesis of networked control systems, multiagent systems, optimal control of power systems, and the Internet of Things.

Prof. Yue is also an Associate Editor of the IEEE Control Systems Society Conference Editorial Board and the *International Journal of Systems Science*.

Chunxia Dou (Member, IEEE) received the B.S. and M.S. degrees in automation from the Northeast Heavy Machinery Institute, Qiqihar, China, in 1989 and 1994, respectively, and the Ph.D. degree from the Institute of Electrical Engineering, 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. From 2005 to 2016, she was a Professor with the Institute of Electrical Engineering, Yanshan University. Since 2016, she has been a Professor with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications. Her current research interests include multiagent-based control, event-triggered hybrid control, distributed coordinated control, and multimode switching control and their applications in power systems, microgrids, and smart grids.

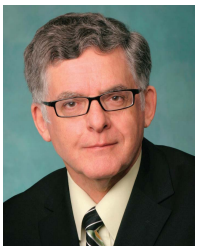


Xiangpeng Xie (Member, IEEE) received the B.S. and Ph.D. degrees in engineering from Northeastern University, Shenyang, China, in 2004 and 2010, respectively.

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 (Senior Member, IEEE) 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, the Ph.D. degree in control theory and applications from Shanghai Jiao Tong University, Shanghai, China, in 1995, and the D.Sc. degree in engineering from Queen's University Belfast, Belfast, U.K., in 2015.

He is currently the Chair of Smart Energy Systems with the University of Leeds, Leeds, U.K. His research interests cover nonlinear system modeling, identification, and control, and artificial intelligence, with substantial applications to energy and power systems, smart grids, electric vehicles, railway systems, and energy management in energy-intensive manufacturing processes.



Gerhardus P. Hancke (Life Fellow, IEEE) received the B.Sc. and B.Eng. degrees and the M.Eng. degree in electronic engineering from Stellenbosch University, Stellenbosch, South Africa, in 1970 and 1973, respectively, and the Ph.D. degree from the University of Pretoria, Pretoria, South Africa, in 1983.

He is currently a Professor with Nanjing University of Posts and Telecommunications and the University of Pretoria and is recognized internationally as a pioneer and leading scholar in industrial wireless sensor networks research. He initiated and coedited 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 coedited a textbook *Industrial Wireless Sensor Networks: Applications, Protocols and Standards* (2013), the first on the topic.

Prof. Hancke has served as an Associate Editor and a Guest Editor for the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS. He has been serving as an Associate Editor and a Guest Editor for the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS and IEEE ACCESS. He is also the Co-Editor-in-Chief of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.