Deep Reinforcement Learning Enabled Bi-level Robust Parameter Optimization of Hydropower-dominated Systems for Damping Ultra-low Frequency Oscillation

Guozhou Zhang, Junbo Zhao, *Senior Member, IEEE*, Weihao Hu, Di Cao, *Member, IEEE*, Nan Duan, *Senior Member, IEEE*, Zhe Chen, *Fellow, IEEE*, and Frede Blaabjerg, *Fellow, IEEE*

Abstract—This paper proposes a robust and computationally efficient control method for damping ultra-low frequency oscillations (ULFOs) in hydropower-dominated systems. Unlike the existing robust optimization based control formulation that can only deal with a limited number of operating conditions, the proposed method reformulates the control problem into a bi-level robust parameter optimization model. This allows us to consider a wide range of system operating conditions. To speed up the bi-level optimization process, the deep deterministic policy gradient (DDPG) based deep reinforcement learning algorithm is developed to train an intelligent agent. This agent can provide very fast lower-level decision variables for the upper-level model, significantly enhancing its computational efficiency. Simulation results demonstrate that the proposed method can achieve much better damping control performance than other alternatives with slightly degraded dynamic response performance of the governor under various types of operating conditions.

Index Terms—Bi-level robust parameter optimization, deep reinforcement learning, deep deterministic policy gradient, ultra-low frequency oscillation, damping control stability.

I. Introduction

IN recent years, the negative effects of the fossil fuel based power system attract more and more attentions. In this context, increasing the penetration of renewable energy with low carbon and sustainable characteristics in the power

G. Zhang, W. Hu (corresponding author), and D. Cao are with School of Mechanical and Electrical Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: zgz@std. uestc. edu. cn; whu@uestc.edu.cn; caodi@std.uestc.edu.cn).

J. Zhao is with Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT, 06269, USA (e-mail: junbo@uconn.edu).

N. Duan is with Lawrence Livermore National Laboratory, Livermore, USA (e-mail: duan4@llnl.gov).

Z. Chen and F. Blaabjerg are with the Department of Energy Technology, Aalborg University, 9220 Aalborg, Denmark (e-mail: zch@et.aau.dk; fbl@et.aau.dk).

DOI: 10.35833/MPCE.2022.000529

system is a significant pathway to address this issue [1]. Hydropower is one of the cheapest and reliable sources of renewable energy and it accounts for nearly 50% of renewable power generation [2]. However, recent studies indicate that the hydropower-dominanted systems are easily subject to the ultra-low frequency oscillation (ULFO) [3], which have been observed in China [4] as well as Nordic and Colombia [5], [6]. If ULFO is not properly controlled, there is a risk of system instability.

In [7], the root cause of ULFO is investigated based on the vector margin method, where it is shown that the hydraulic turbine-governor creates a negative damping torque in the ultra-low frequency band. In [8], via the damping torque analysis, it is found that the unreasonable parameter settings of the hydraulic governor proportional-integral (PI) controller lead to ULFO. Furthermore, if the ratio of the proportional parameter to the integral parameter is too small for the hydropower governor, the ULFO can occur [9]. To damp out ULFO, various strategies have been proposed. In [10], a high-voltage direct current frequency limiting controller (HVDC-FLC) is developed. Note that for partial hydropower-dominated systems, only the AC transmission is used for the transmission of electric energy and thus it is difficult to adopt this method. In [11], the power system stabilizer (PSS) is configured on the governor to mitigate the impacts of the negative damping torque. By contrast, in [12], the PSS and proportion resonant (PR) are integrated to form the PR-PSS for ULFO control. A multi-band PSS, named PSS4B, is also proposed [13] with low-band, intermediate-band, and high-band controllers. The low-band controller provides the damping of the ULFO modes. However, these configurations are not always available in practice, limiting their applications in power systems.

Optimizing the proportional-integral-derivative (PID) parameters of the governor is another alternative for ULFO control. It is also shown that the re-tuning of the governor settings has high practical value [14]. In [15], [16], a robust optimization method for tuning governor is proposed. A robust fixed order control design method is also developed in [2]. However, these methods are limited to several typical operating conditions. To this end, a bi-level optimization

model (min-max or max-min) is proposed [17]. In particular, the performance of optimization variables under extreme operating conditions is taken as the objective to achieve generalizability under all operating conditions. Although the results are conservative, it is appropriate considering the importance of system stability unless other effective methods are available. To solve the bi-level model, it is advocated to replace the lower-level model using Karush-Kuhn-Tucker formulation [18], duality principle [19], and penalty function [20], yielding a single-level optimization model. However, for the decision variables provided by the upper-level model at each iteration, the lower-level model needs to repeat optimizations to find the corresponding extreme operating condition, which is usually time-consuming.

To this end, a novel bi-level PID parameter optimization model is proposed for ULFO control. It has the following contributions:

1) Based on the Routh-Hurwitz criterion, the mechanism of the ULFO is studied and the feasibility of formulating ULFO control as optimizing PID parameters is demonstrated. In particular, the problem of ULFO control is formulated as a bi-level robust PID optimization model. This is in contrast with the formulation in [14], [15] and allows us to deal with a wider range of extreme system operating conditions. In addition, the optimized PID parameters retain the dynamic performance of the governor.

2) To improve the efficiency of solving bi-level model, this paper forms the lower-level model into the Markov decision process (MDP) solved by a deep deterministic policy gradient (DDPG) based algorithm. After that, the decision variables transferred by the upper-level model can be quickly addressed via the well-trained DDPG agent without the repeated optimization. This is a novel method to solve min-max optimization model and is different from previous min-max model [18]-[20], yielding significantly improved computational efficiency.

The rest of this paper is organized as follows. Section II introduces the system model. In Section III, the formulation of governor parameter optimization is presented. The proposed bi-level robust parameter optimization model is presented in Section IV. The case study is provided in Section V. Section VI presents the actual hydropower-dominated system and conclusions are given in Section VII.

II. SYSTEM MODEL

A. Two-machine System

Figure 1 shows a two-machine system, which is introduced to act as the studied system to investigate ULFO, including hydropower unit (G1), thermal power unit (G2), and load. The key parameter settings of this system are provided in Appendix A. The hydropower mainly consists of a generator, a governor, and a turbine, as shown in Fig. A1 of Appendix A. Among them, the linearized dynamic equation of generator can be described as:

$$T_J \frac{d\Delta\omega}{dt} = \Delta P_m - \Delta P_e - D_s\Delta\omega \qquad (1)$$

where T_J is the inertia time constant; $\Delta\omega$ is the rotor speed

deviation; ΔP_m is the mechanical power deviation; ΔP_e is the electromagnetic power deviation; and D_s is the damping coefficient.
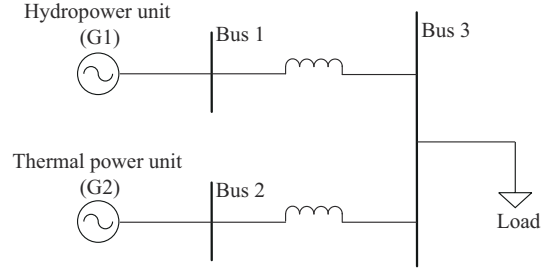

Fig. 1.   Schematic of two-machine system.

Assuming frequency-dependent load as $\Delta P_e = K_L\Delta\omega$, where K_L is the load frequency sensitivity, and substituting $\Delta P_e = K_L\Delta\omega$ into (1), the transfer function of the generator $G_{gen}(s)$ can be described as [15]:

$$G_{gen}(s) = \frac{\Delta\omega(s)}{\Delta P_m(s)} = \frac{1}{T_J s + D_s + K_L} = \frac{1}{T_J s + D} \qquad (2)$$

where $D = D_s + K_L$. The transfer function of prime mover (governor and turbine) $G_m(s)$ can be defined as [4]:

$$\begin{cases} G_m(s) = \dfrac{\Delta P_m(s)}{-\Delta\omega(s)} = G_{gov}(s)G_t(s) \\[2mm] G_{gov}(s) = \dfrac{K_D s^2 + K_P s + K_I}{b_p K_I + s} \dfrac{1}{1 + T_G s} \\[2mm] G_t(s) = \dfrac{1 - T_w s}{1 + 0.5 T_w s} \end{cases} \qquad (3)$$

where $G_{gov}(s)$ and $G_t(s)$ are the transfer functions of governor and turbine, respectively; K_P, K_I, and K_D are the proportional, integral, and differential coefficients of the governor, respectively; b_p is the adjustment coefficient of the governor; T_G is the response time of the governor; and T_w is the water hammer effect and it depends on the operating conditions of hydroturbine [2]:

$$T_w = \frac{\sum\limits_{i=1}^{n} \dfrac{L_i}{A_i} Q_r}{g H_r} \qquad (4)$$

where L_i and A_i are the length and sectional area of the ith diversion pipeline, respectively; Q_r is the rated water flow; g is the gravitational acceleration; H_r is the rated water head; and n is the number of diversion pipelines.

For the thermal power unit, this paper ignores the boiler dynamic process and mainly considers the steam turbine and governor [8]:

$$\Delta P_{mh}(s) = -G_h(s)\Delta\omega(s) = \frac{-K_a\Delta\omega(s)}{(1 + sT_g)(1 + sT_{ch})} \qquad (5)$$

where $\Delta P_{mh}(s)$ is the damping torque of thermal generator; $G_h(s)$ is the transfer function of thermal power unit; K_a is the gain; T_g and T_{ch} are the time-steps of hydraulic system and high-pressure cylinder, respectively.

B. ULFO Mechanism Analysis

The dynamic responses of G1 and G2 under fault 1, i.e., a

double-phase short-circuit fault at bus 3 from 2.0 s to 2.2 s, are shown in Fig. 2. It can be observed that, the angle speed of two generators varies with the same phase and amplitude. No oscillations between two units are observed. It is different from traditional low-frequecy oscillation (LFO). The Prony method [21] is utilized to identify the oscillation model and the identification results show that a ULFO mode $-0.0001+\mathrm{j}0.403$ exists in the system.
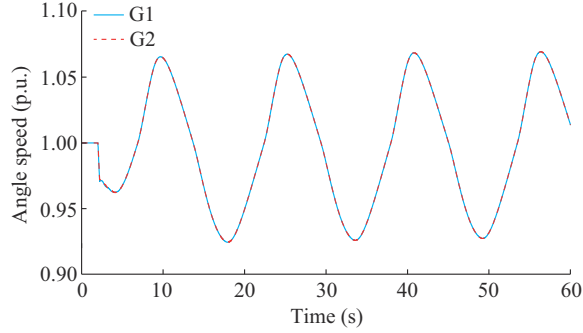

Fig. 2. Dynamic responses of G1 and G2 under fault 1.

To investigate the cause of ULFO in the studied system, we calculate the damping torque coefficient of two units' primary frequency regulation (PFR). Among them, the detailed calculation process of damping torque coefficient of hydropower unit is provided in [12] and can be defined as:

$$D_H(s) = \frac{1-0.5T_G T_w \omega_d^2 - \omega_d^2 T_\omega(T_G+0.5T_w)}{1-0.5T_G T_w \omega_d^2 + \omega_d^2(T_G+0.5T_w)} \quad (6)$$

Submitting $s=\mathrm{j}\omega_d$ into (3), the damping torque coefficient of stream power unit can be obtained as [8]:

$$\Delta P_m = -D_T \Delta \omega - \mathrm{j}D_{IM}\Delta \omega \quad (7)$$

$$D_T(s) = \frac{K_a\left(1-T_g T_{ch}\omega_d^2\right)}{\left(1-T_g T_{ch}\omega_d^2\right)^2 + \left(T_g+T_{ch}\right)^2 \omega_d^2} \quad (8)$$

where D_T denotes the damping torque coefficient; and D_{IM} denotes the synchronous torque coefficient.

Submitting $\omega_d=2\pi f$ into (6) and (8), the trajectory of damping torque coefficients D_H and D_T changing with frequency f can be obtained, as shown in Fig. 3.

It can be observed that the hydropower unit will produce a negative damping under the ultra-low frequency band (below 0.1 Hz). In contrast, the thermal power unit would produce a positive damping in this band. Therefore, the ULFO is strongly related with hydropower unit.

To further investigate the relationship of ULFO and hydropower unit, the characteristic equation of closed-loop transfer function of the hydropower unit is calculated based on (2) and (3) and can be written as:

$$\left(T_J s+D\right)\left(b_p K_I+s\right)\left(1+T_G s\right)\left(1+0.5T_w s\right) + \left(1-T_w s\right)\left(K_D s^2+K_P s+K_I\right)=0 \quad (9)$$

After simplification, (9) can be rewritten as:

$$s^4 + a_1 s^3 + a_2 s^2 + a_3 s + a_4 = 0 \quad (10)$$

where a_1, a_2, a_3, and a_4 are the Routh-Hurwitz criterion coef-

ficients, which are related to T_w. Specifically, the Routh-Hurwitz criterion coefficients and oscillation frequency changing with T_w are shown in Fig. 4. Note that, the detailed expression of these coefficients and the calculation process of oscillation frequency are provided in Appendix A.
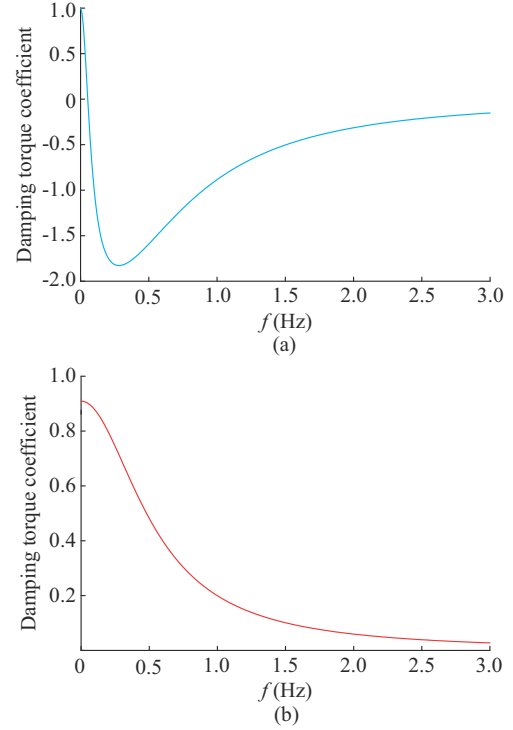

Fig. 3. Trajectory of damping torque coefficients of hydropower and stream power units. (a) Hydropower unit. (b) Stream power unit.
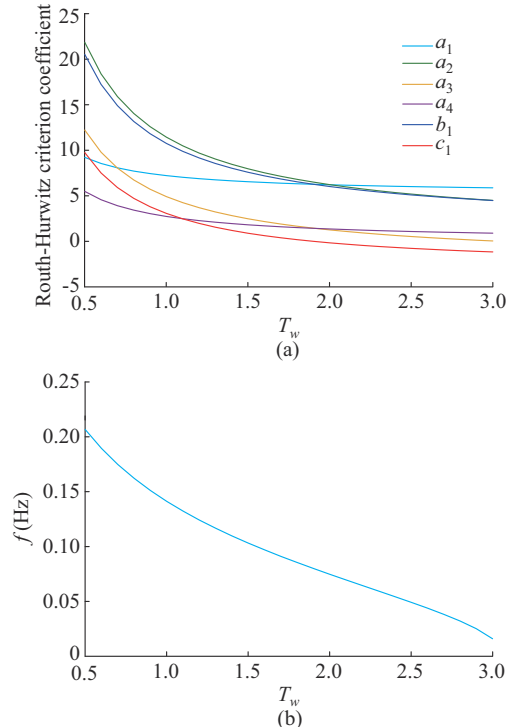

Fig. 4. Routh-Hurwitz criterion coefficients and oscillation frequency changing with T_w. (a) Routh-Hurwitz criterion coefficients. (b) Oscillation frequency.

It can be observed from Fig. 4(a) that with the increase of T_w, all Routh-Hurwitz criterion coefficients decrease and in particular, the coefficient c_1 becomes negative. Base on the Routh-Hurwitz criterion [3], it can be concluded that the system is unstable, and there is a negative damping oscillation mode in the system. In addition, Fig. 4(b) shows that, with the increase of T_w, the oscillation frequency decreases to below 0.1 Hz. It means that this dominant negative damping oscillation mode is ULFO mode.

In summary, the ULFO is strongly related to the hydropower units, and it is caused by the PFR of the hydropower units. More specifically, due to the water hammer effect, the hydraulic governor easily produces negative damping torque, as shown in Fig. 3, resulting in the negative ULFO mode.

Besides, Routh-Hurwitz criterion coefficients are related to the hydrogovernor PID parameters. In fact, we can adjust PID parameter settings to make these coefficients keep positive. Based on this consideration, we test the trajectory of c_1 with different PID parameter settings and the results are shown in Fig. 5. It can be observed that the tuning of PID parameters can make the c_1 positive. In this way, the system becomes stable, which means that optimizing PID parameters contributes to preventing ULFO.
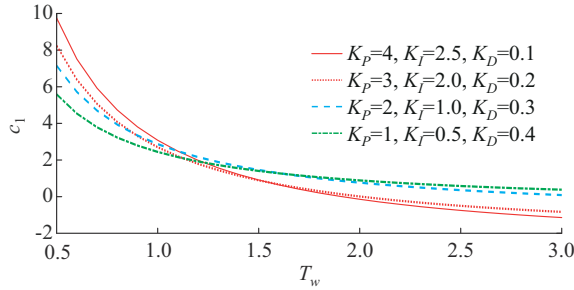

Fig. 5. c_1 changing with different PID parameter settings.

III. FORMULATION OF GOVERNOR PARAMETER OPTIMIZATION

Based on the system linearization technology, the state matrix A can be obtained [22]. Then, (11) can be utilized to diagonalize matrix A to obtain eigenvalues:

$$\begin{cases} A = V\Lambda \\ UV^{\mathrm{T}} = I \end{cases} \quad (11)$$

$$\begin{cases} \Lambda = \mathrm{diag}\left(\lambda_1, \lambda_2, ..., \lambda_{\mathrm{ULFO}}, ..., \lambda_m\right) \\ \lambda_{\mathrm{ULFO}} = \sigma_{\mathrm{ULFO}} + \mathrm{j}\beta_{\mathrm{ULFO}} \\ \xi_{\mathrm{ULFO}} = -\sigma_{\mathrm{ULFO}} \Big/ \sqrt{\sigma_{\mathrm{ULFO}}^2 + \beta_{\mathrm{ULFO}}^2} \end{cases} \quad (12)$$

where U and V are the left and right feature vectors, respectively; Λ is the diagonal matrix; λ_m denotes the mth eigenvalue; λ_{ULFO} is the ULFO mode; σ_{ULFO} and β_{ULFO} are the real part and imaginary part of λ_{ULFO}, respectively; and ξ_{ULFO} is the damping ratio.

As mentioned above, optimizing PID parameter settings contributes to stable ULFO mode λ_{ULFO}. However, there are some requirements: ① less effect on other oscillation modes: we should avoid weakening the damping of other modes when improving the damping of ULFO mode; ②

good dynamic performance for the PFR: previous studies indicate that if without proper design, optimizing PID parameters may deteriorate PFR dynamic performance, and thus weakening the frequency adjustment ability of governor [14], [15]; ③ robustness to the change of operating conditions: in actual operation, the operating condition of the turbine is time-varying. The optimized PID parameters that show good performance under one operating condition may not work well under other conditions. Therefore, during the optimization process, extreme operating conditions should be considered. Based on the above considerations, the suppression of ULFO is modeled as a bi-level min-max optimization model with constraints, which is described as:

$$\begin{cases} \min J^{upper}\left(K_{P,i}, K_{I,i}, K_{D,i}, T_{w,i}^*\right) \\ K_{P,i}^{\min} \le K_{P,i} \le K_{P,i}^{\max} \\ K_{I,i}^{\min} \le K_{I,i} \le K_{I,i}^{\max} \\ K_{D,i}^{\min} \le K_{D,i} \le K_{D,i}^{\max} \qquad i = 1, 2, ..., N \quad (13) \\ T_{w,i}^* = \arg\max y_z^{lower}\left(K_{P,i}, K_{I,i}, K_{D,i}\right) \\ T_{w,i}^{\min} \le T_{w,i} \le T_{w,i}^{\max} \\ \xi_j'\left(K_{P,i}, K_{I,i}, K_{D,i}, T_{w,i}\right) > \xi_0 \end{cases}$$

where $J^{upper}(\cdot)$ is the objective function of the optimization problem; $T_{w,i}^*$ is the time constant of water hammer effect of the ith governor representing the operating conditions of the governor and limited between 0.5 and 3 [15]; y_z^{lower} is the objective function of the lower model; $\xi_j'(\cdot)$ is the damping of the jth oscillation mode except for the ULFO mode; ξ_0 is the desired damping ratio of the oscillation mode that is typically set to be 0.05 [12]; and N is the number of the governors. $J^{upper}(\cdot)$ can be reformulated as [15]:

$$J^{upper}\left(K_{P,i}, K_{I,i}, K_{D,i}, T_{w,i}\right) = J_{\mathrm{STB}} + J_{\mathrm{ITAE}}$$

$$\begin{cases} J_{\mathrm{STB}} = \sum_{\xi_{\mathrm{ULFO}} < \xi_{\mathrm{set}}} \left|\xi_{\mathrm{ULFO}} - \xi_{\mathrm{set}}\right| + \sum_{\sigma_{\mathrm{ULFO}} > \sigma_{\mathrm{set}}} \left|\sigma_{\mathrm{ULFO}} - \sigma_{\mathrm{set}}\right| \\ J_{\mathrm{ITAE}} = \sum_{i=1}^{N}\left(\int_0^T \left|t\left(\Delta P_m(t) - \Delta P_m^\infty\right)\right|\mathrm{d}t\right) \end{cases} \quad (14)$$

where J_{STB} is the stability of the ULFO; J_{ITAE} is the primary frequency control performance of the prime mover; σ_{set} is the desired real part of the eigenvalues; ξ_{set} is the damping ratio of the eigenvalues; $\Delta P_m(t)$ is the dynamic response of governor under disturbance; and ΔP_m^∞ is the steady-state value of the prime mover under disturbance. In fact, J_{STB} can be further written as:

$$J_{\mathrm{STB}} = \begin{cases} \xi_{\mathrm{set}} - \xi_{\mathrm{ULFO}} + \sigma_{\mathrm{ULFO}} - \sigma_{\mathrm{set}} & \xi_{\mathrm{ULFO}} < \xi_{\mathrm{set}}, \sigma_{\mathrm{ULFO}} > \sigma_{\mathrm{set}} \\ 0 & \xi_{\mathrm{ULFO}} \ge \xi_{\mathrm{set}}, \sigma_{\mathrm{ULFO}} \le \sigma_{\mathrm{set}} \end{cases} \quad (15)$$

It can be observed from (15) that, when $\xi_{\mathrm{ULFO}} < \xi_{\mathrm{set}}, \sigma_{\mathrm{ULFO}} > \sigma_{\mathrm{set}}$, J_{STB} is defined as $\xi_{\mathrm{set}} - \xi_{\mathrm{ULFO}} + \sigma_{\mathrm{ULFO}} - \sigma_{\mathrm{set}}$, by minimizing this target, ξ_{ULFO} and σ_{ULFO} of ULFO mode would be pushed close to the predetermined damping and real part. Note that the damping of ULFO mode is set to be a bigger value. By optimizing PID parameter settings to make the damping of ULFO mode close to the predetermined damping, the damp-

ing of ULFO mode is improved. Once the damping and real parts of ULFO mode reach to the predetermined value, J_{STB} becomes 0.

By minimizing J_{STB}, the output deviation of the prime mover can be optimized to close to steady-state point. In this way, both the oscillation amplitude and oscillation time of the prime mover under disturbance can be optimized, which means that the dynamic performance of the primary frequency control can be improved.
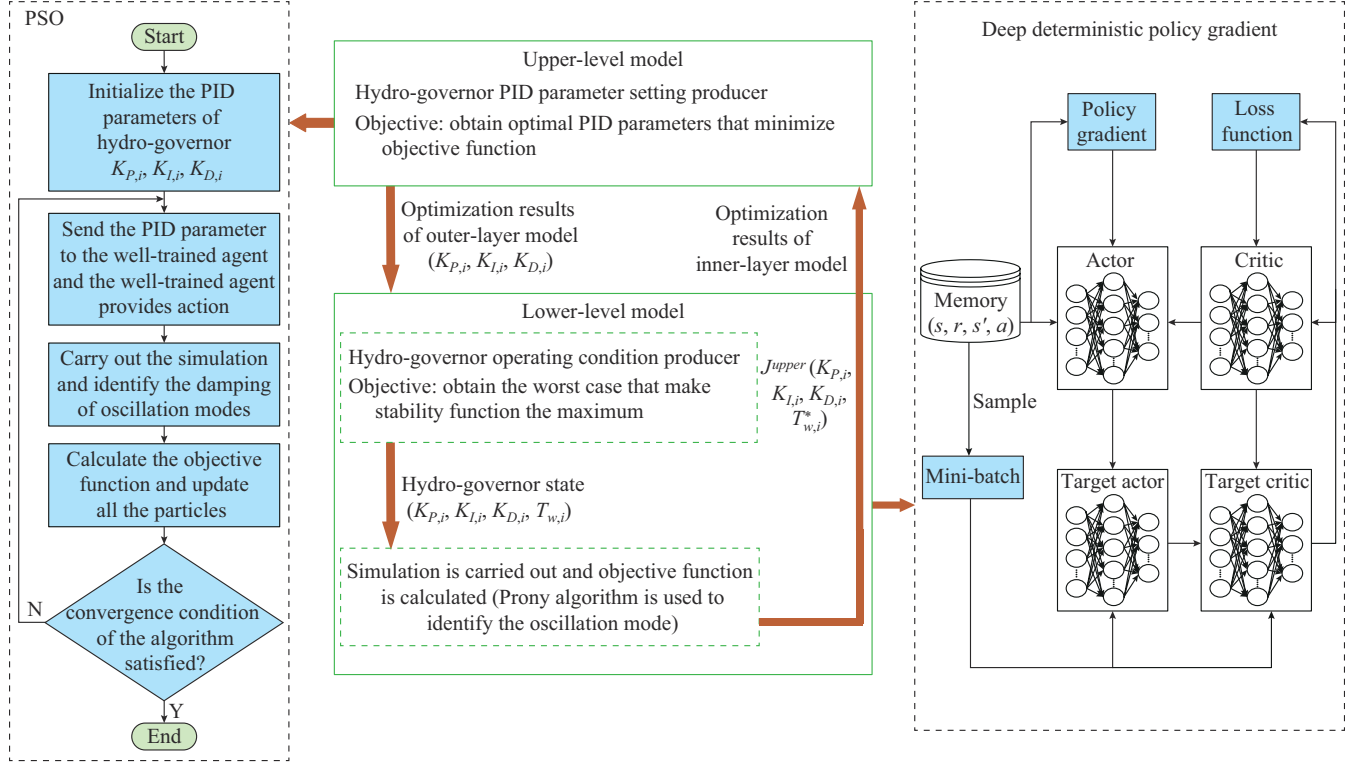
IV. PROPOSED BI-LEVEL ROBUST PARAMETER OPTIMIZATION MODEL

For the bi-level robust parameter optimization model, the lower-level model is reformulated as the MDP and solved by the DDPG algorithm. Subsequently, the trained agent is employed to assist particle swarm optimization (PSO) to solve the upper-level model to obtain the optimal solution. The overall scheme of the optimization process is shown in Fig. 6.

Fig. 6.   Overall scheme of optimization process.

A. Solving Lower-level Model via DDPG Algorithm

The core functions of lower-level model shown in Fig. 6 can be described as:

$$\begin{cases} T_{w,i}^* = \arg\max y_z^{lower}\left(K_{P,i}, K_{I,i}, K_{D,i}\right) \\ T_{w,i}^{\min} \le T_{w,i} \le T_{w,i}^{\max} \qquad i = 1, 2, \ldots, N \\ \xi_j'\left(K_{P,i}, K_{I,i}, K_{D,i}, T_{w,i}\right) > \xi_0 \end{cases} \quad (16)$$

In each episode, the upper-level model would provide a set of PID parameter setting $\left(K_{P,i}, K_{I,i}, K_{D,i}\right)$ to the lower-level model. The lower-level model searches the extreme operating condition $T_{w,i}^*$ with the worst performance under this PID setting condition. In this way, both the upper-level and lower-level variables are determined and the objective function (14) can be calculated and sent to the upper-level model.

In fact, the function of the lower-level model is to find the corresponding extreme operating condition under every PID setting condition. Therefore, the key to solving this lower-level model is to find the policy function $T_{w,i}^* = \pi\left(K_{P,i}, K_{I,i}, K_{D,i}\right)$ that maximizes the objective function. It is

a decision-making problem in an uncertain environment. This paper reformulates it as an MDP, and the key elements related to MDP are defined as a tuple, $\langle S, A, P, R \rangle$, where S denotes the state and is composed of the PID parameter settings of each hydrogovernor; A denotes the action and is represented as the operation conditions of governor; P denotes the transition probability; and R denotes the reward and is used to evaluate the action taken by the agent at each time step. In this paper, the reward $r\left(s_k, a_k\right)$ is defined as:

$$r\left(s_k, a_k\right) = J\left(K_{P,i}^k, K_{I,i}^k, K_{D,i}^k, T_{w,i}^k\right) + \sum_{j=1}^{m} \sum_{\xi_0 < \xi_j'} \left(\xi_0 - \xi_j'\left(K_{P,i}^k, K_{I,i}^k, K_{D,i}^k, T_{w,i}^k\right)\right) \quad (17)$$

where the superscript k denotes the kth turbine governor.

The reinforcement learning algorithm is a common method to solve this MDP [22]. The background of reinforcement learning is described as follows:

1) Q-learning: it is a widely used reinforcement learning method [23], where an agent is trained to learn the optimal control policy to maximize the cumulative reward. In Q-

learning, each state-action pair (s_k, a_k) is assigned with a Q-value, stored in the Q-table, and denoted as $Q(s_k, a_k)$. For a given state, the higher Q-value for an action denotes a higher potential cumulative reward and it can be utilized to update Q-value via Bellman equation [24]:

$$Q(s_k, a_k) \leftarrow Q(s_k, a_k) + \\ \alpha \left[r(s_k, a_k) + \gamma \max Q(s_{k+1}, a_{k+1}) - Q(s_k, a_k) \right] \quad (18)$$

where γ is the discount rate; and α is the learning rate. The ε-greedy policy is adopted in Q-learning, where the agent will choose the action with the highest Q-value during the training process. Note that Q-learning is only suitable for the case where state-action space is small. The increase of state-action space size would make the Q-table become too big, resulting in each Q-value being rarely updated. In this context, deep Q-network (DQN) is proposed to address it.

2) DQN: DQN is proposed for solving high dimension state-action problem via combining deep neural network (DNN) with Q-learning method. Specifically, a DNN is utilized to approximate Q-table, named Q-network, represented as $Q(s_k, a_k, \theta)$, where θ denotes the parameters of DNN. The Q-network takes the state as input and outputs the Q-value for each state-action pair. It can be trained via minimizing the loss function:

$$L(\theta) = \mathbb{E} \left[y - Q(s_k, a_k, \theta)^2 \right] \quad (19)$$

where $\mathbb{E}(\cdot)$ is the expected function; $y = r(s_k, a_k) + \gamma \max Q(s_{k+1}, a_{k+1}, \theta')$ denotes the target Q-value, and θ' denotes the parameter of target Q-network and is updated via soft-update method [25]. To stabilize training process, it can be calculated via the target Q-network $Q(s_{k+1}, a_{k+1}, \theta')$.

Moreover, a replay buffer is employed in DQN to break the identically distributed state samples and reduce the correlation between them, leading to improved data efficiency. Specifically, during the training process, all information is saved as an experience $e_k = (s_t, a_t, r_t, s_{t+1})$ and stored in the memory $D = \{e_1, e_2, ..., e_M\}$. After the buffer is full, the oldest experience will be replaced by the newly obtained one. Subsequently, at each iteration, the agent will sample a mini-batch of experience from the replay buffer.

3) DDPG: DDPG is the standard DQN method only working effectively for solving control problems with continuous states and low-dimension discrete action sets. It is not suitable to solve the optimization problem of PID parameter setting. To this end, DDPG is introduced and it can achieve better performance in solving control problems with continuous action space than DQN. Figure 6 shows the procedure of the DDPG and it consists of two eponymous ingredients: an actor network is utilized to fit the state s_k to action a_k, denoted as policy function $\mu(s_k|\theta^\mu)$ via adjusting the network parameter $\theta^\mu = \{W_1^a, b_1^a, W_2^a, b_2^a, ..., W_n^a, b_n^a\}$; and a critic network is utilized to fit the action-value function $Q(s_k, a_k|\theta^Q)$ via adjusting network parameter $\theta^Q = \{W_1^c, b_1^c, W_2^c, b_2^c, ..., W_m^c, b_m^c\}$.

During the training process, these two networks are trained against each other. Among them, the critic network can be updated by the loss function [26]:

$$\begin{cases} L(\theta^Q) = \dfrac{1}{N_{mm}} \sum_{i=1}^{n} \left(Q(s_k, a_k|\theta^Q) - y_k \right)^2 \\ y_k = r(s_k, a_k) + \gamma Q\left(s_{k+1}, \mu(s_k|\theta^\mu) \big| \theta^Q \right) \\ \theta_{k+1}^Q = \theta_k^Q + \alpha^Q \nabla_{\theta^\mu} L(\theta_k^Q) \end{cases} \quad (20)$$

where N_{mm} is the number of mini-batch; and α^Q is the learning rate of the critic networks, Then, the parameters of the actor network can be updated by the gradient descent [27]:

$$\begin{cases} \nabla_{\theta^\mu} J(\theta^\mu) = \\ \quad \dfrac{1}{N_{mm}} \sum_{i=1}^{N} \nabla_a Q(s, a|\theta^Q) \Big|_{s=s_k, a=\mu(s_k)+\mathcal{N}} \nabla_{\theta^\mu} \mu(s|\theta^\mu) \Big|_{s=s_k} \\ \theta_{k+1}^\mu = \theta_k^\mu + \alpha^\mu \nabla_{\theta^\mu} J(\theta_k^\mu) \end{cases}$$
$$(21)$$

where α^μ is the learning rate of the actor network; and $\mathcal{N}$ is the Gaussian noise.

Moreover, to stabilize the training process, both target actor-critic pair networks are added in DDPG, which can be parameterized by $\theta^{\mu'}$ and $\theta^{Q'}$. At each iteration, the soft updating (22) is utilized to synchronize its parameters to the target actor-critic networks [26].

$$\begin{cases} \theta^{Q'} \leftarrow \tau \theta^Q + (1-\tau) \theta^{Q'} \\ \theta^{\mu'} \leftarrow \tau \theta^\mu + (1-\tau) \theta^{\mu'} \end{cases} \quad (22)$$

where τ is the soft-update rate.

B. Combining PSO and Well-trained DDPG-enabled Agent to Solve Upper-level Model

After off-line training, the well-trained DDPG-enabled agent can learn the optimal control policy and provide extreme operating conditions for each PID parameter setting of the system.

$$T_{w,i}^* = \mu\left(K_{P,i}, K_{I,i}, K_{D,i} \big| \theta^\mu \right) \quad i = 1, 2, ..., N \quad (23)$$

In this way, the proposed bi-level min-max optimization model (13) can be transformed into a single-level mathematical programming problem with constraints, i.e.,

$$\begin{cases} \min J\left(K_{P,i}, K_{I,i}, K_{D,i}, T_{w,i}^* \right) \\ K_{P,i}^{\min} \leq K_{P,i} \leq K_{P,i}^{\max} \\ K_{I,i}^{\min} \leq K_{I,i} \leq K_{I,i}^{\max} \qquad i = 1, 2, ..., N \\ K_{D,i}^{\min} \leq K_{D,i} \leq K_{D,i}^{\max} \\ T_{w,i}^* = \mu\left(K_{P,i}, K_{I,i}, K_{D,i} \big| \theta^\mu \right) \end{cases} \quad (24)$$

It can be observed from (23) that the bi-level model is converted into a nonlinear optimization problem. A heuristic algorithm is a good choice to solve that and this paper uses the PSO algorithm [28]. The detailed solution processes of the optimization model are as follows.

Step 1: define the solution space and fitness function. The PSO is used to find the optimal PID parameter settings for governor. In this context, the particle position is designed as the PID parameter settings $(K_{P,i}, K_{I,i}, K_{D,i})$. The fitness function is applied to evaluate the training error and the good-

ness of a given solution, which is defined as the objective function and shown in (14).

Step 2: initialize random swarm location and velocities. Before beginning to search the optimal position, each particle is initialized with the random PID parameter setting within the allowable ranges. Moreover, the direction and length of movement of the particle at each episode are named velocity, which is also initialized.

Step 3: calculate the fitness of each particle. The PID parameter setting carried by each particle is transmitted to the well-trained DDPG agent. Then, the actor network in DDPG can provide the extreme operating condition T_{w-i}^* with the worst performance for each particle under the corresponding PID parameter setting. Next, both PID parameter setting and operating condition data of each particle can be updated to the studied system to calculate the fitness via (14).

Step 4: update the particle position and velocity. The velocity and position of each particle can be updated via:

$$\begin{cases} v_{i+1}^d = \omega v_i^d + c_1 r_1 \left(p_i^d - x_i^d \right) + c_2 r_2 \left(p_g - x_i^d \right) \\ x_{i+1}^d = x_i^d + v_i^d \end{cases} \quad (25)$$

where v_i^d is the velocity of the dth particle at the ith iteration; x_i^d is the position of the dth particle at the ith iteration; p_g is the best position among all particles in the population up to the nth iteration; p_i^d is the optimal position of the dth particle up to the nth iteration; ω is the inertia factor; c_1 and c_2 are the acceleration coefficients; and r_1 and r_2 are the random numbers in [0,1]. In this paper, $c_1 = c_2 = 2$; $\omega = 0.6$ [15].

Step 5: iterate to find the optimal solution. Repeat *Steps 2* and *3* until the minimum error is met, or the maximum number of iterations is reached. Output the final result as the solution to the above optimization problem.

V. CASE STUDY

In this section, the performance of the proposed method is tested on the IEEE 10-machine 39-bus system. All generators adopt a fifth-order model and a simplified excitation system [14]. The eigenvalue analysis results show that a ULFO mode of $0.00045 + j0.55$ exists in the system. The generators strongly related with this mode include G1, G2, G5, and G8. To damp out that oscillation, we formulate the ULFO suppression problem as a bi-level PID parameter setting optimization model. Then, the DDPG and PSO algorithms are combined to solve this model. The simulation is carried out via Python and power system analysis software package (PSASP) [29]. Among them, the studied system is constructed in PSASP. PSO and DDPG algorithms are modelled in Python.

A. Performance Evaluation of Trained Agents

The numbers of layers and neurons for the networks in the DDPG algorithm are set as follows: both actor and critic networks adopt the same structure, which contains three hidden layers; and the numbers of neurons for the hidden layers are 128, 64, and 64, respectively. The hyperparameters of the DDPG algorithm can be found in Table I.

To obtain the corresponding extreme operating conditions

with different PID parameter settings, which maximize the objective function, numerous scenarios are constructed for training. Specifically, we generate scenarios by randomly sampling from the upper- and lower-limits of PID parameter settings of each hydroturbine ($0 \le K_{P,i} \le 25$, $0 \le K_{I,i} \le 0.5$, $0 \le K_{D,i} \le 3$). Note that 95% of these PID parameter settings are used for training while the remaining 5% are used for testing. The cumulative reward changing with the episode during the training process is shown in Fig. 7, where the blue solid lines and the shaded areas represent the average accumulated reward and the standard deviation of the reward, respectively. At the beginning of the training process, as the parameters of DDPG are randomly initialized, the agent does not learn the optimal policy that satisfies all constraints, leading to low accumulated reward. As the training process goes further, the agent interacts with the system to obtain experiences to optimize the parameters of DDPG, and the cumulative reward increases gradually. After about 5000 episodes, the algorithm converges.

TABLE I
HYPERPARAMETERS OF DDPG ALGORITHM

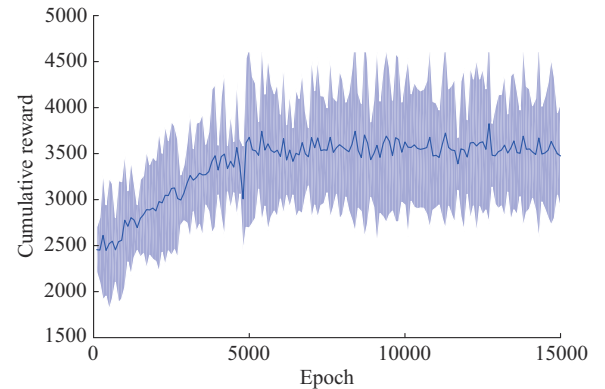| Parameter	Value
Learning rate for actor network	0.001
Learning rate for critic network	0.002
Experience replay memory capacity	8000
Step size of each episode	12
Mini-batch size	40

Fig. 7.   Cumulative reward changes with episode during training process.

After training, the well-trained agent can provide each PID parameter setting to the corresponding extreme operating condition. After that, this agent can be applied to assist the PSO algorithm to solve the upper-level model. To show the computational efficiency of this agent, we compare the solution methods of the lower-level model in [19] and [20] with our trained agent. The computational efficiency comparison of different methods is shown in Table II. It can be observed that our proposed method achieves the highest computational efficiency.

Moreover, we further investigate the performance of the PSO and differential evolution (DE) to test the performance of these two methods. To balance the training time and optimization effect, by careful trial and error, the population size

of these two methods is set to be 30. Except the population size, the hyperparameters of the method have a great impact on the optimization performance. To deal with it, each hyperparameter in the method will first be discretized into several points in its feasible region. Then, these discrete points form a test table by the orthogonal method.

TABLE II
COMPUTATIONAL EFFICIENCY COMPARISON OF DIFFERENT METHODS

Method	Average solution time of lower-level model (s)
Optimization method in [19]	13.40
Optimization method in [20]	3.70
Proposed method	0.05

In this way, we can form nine sets of hyperparameter settings for DE and PSO algorithms, respectively. Then, we can train the PSO and DE algorithms with each set of hyperparameter settings, as shown in Fig. 8, where the solid lines and the shaded areas represent the average and the standard deviations of the objective function under different trials, respectively. Moreover, some quantitative indicators are introduced to evaluate the optimization performance of PSO and DE algorithms, as shown in Table III. It can be observed that, the performance of both PSO and DE algorithms is almost similar in terms of the best fitness value obtained in the 18 runs. Specifically, DE algorithm is slightly better than PSO algorithm. However, the PSO algorithm performs better compared with DE algorithm in terms of the average, the worst, and the standard deviations of the objective function.
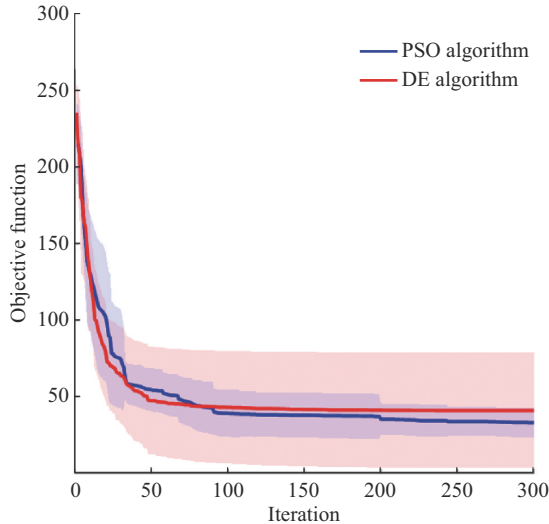

Fig. 8.   Convergence process of DE and PSO algorithms with different hyperparameter settings.

TABLE III
QUANTITATIVE INDICATORS OF OPTIMIZATION PERFORMANCE OF PSO AND DE ALGORITHMS WITH DIFFERENT HYPERPARAMETER SETTINGS

Algorithm	Value of quantitative indicator			
	Best	Average	Worst	Standard deviation
PSO	26.42	33.18	53.30	9.71
DE	25.01	41.07	140.27	37.68

B. Control Performance Comparisons

To test the effectiveness of the proposed method, five cases are tested by setting different time constants of the hydro-turbine T_w, as shown in Table IV. Other PID parameter setting tuning methods are compared as well, including the optimization method shown in [14] (called robust method I), and the one in [15] (called robust method II). The optimized parameters are used in the time-domain simulations to validate the damping control performances. Taking G1 and G2 as examples, the frequency deviations of units under fault 1, i.e., a double-phase short-circuit fault at bus 3 from 2.0 s to 2.2 s, are shown in Fig. 9. The damping of ULFO mode by different optimization methods is presented in Table V. It can be observed from Table V and Fig. 9 that ULFO exists in the system before PID parameter tuning. After PID parameter optimization, the damping of ULFO mode is enhanced and the oscillation is suppressed. However, the PID parameter settings optimized by different methods show different performances. Compared with the other two robust optimization methods, the proposed method achieves the best damping of ULFO mode and allows the generators to reach a steady state in the shortest time.

TABLE IV
T_w OF HYDRO-TURBINES IN DIFFERENT CASES

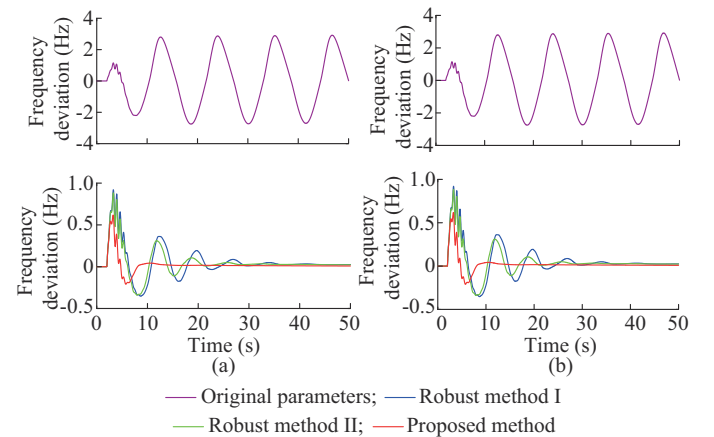| Generator	T_w of hydro-turbine				
	Case 1	Case 2	Case 3	Case 4	Case 5
G1	0.5	2.5	1.5	1.5	2.5
G2	1.5	1.5	2.5	0.5	0.5
G3	0.5	1.5	1.5	1.5	0.5
G4	2.5	2.5	2.5	1.5	1.5
G5	1.5	2.5	0.5	1.5	0.5
G6	1.5	0.5	0.5	1.5	2.5
G7	0.5	1.5	0.5	0.5	2.5
G8	0.5	0.5	1.5	1.5	1.5
G9	1.5	2.5	1.5	2.5	1.5
G10	2.5	0.5	1.5	0.5	2.5

Fig. 9.   Frequency deviations of units under fault 1. (a) G1. (b) G2.

To further elaborate on the results, the dynamic characteristic of G1 and step responses of PFR are shown in Figs. 10 and 11. It can be observed that the original parameters make

the units achieve the best dynamic response but the worst damping performance of ULFO. The parameter settings re-tuned by the three optimization methods can improve the damping characteristics of the governor with the cost of re-duced dynamic characteristics. Among them, the proposed method not only improves the damping characteristics of the governors but also retains a better dynamic response of PFR than the other two methods.

TABLE V
DAMPING OF ULFO MODE BY DIFFERENT OPTIMIZATION METHODS

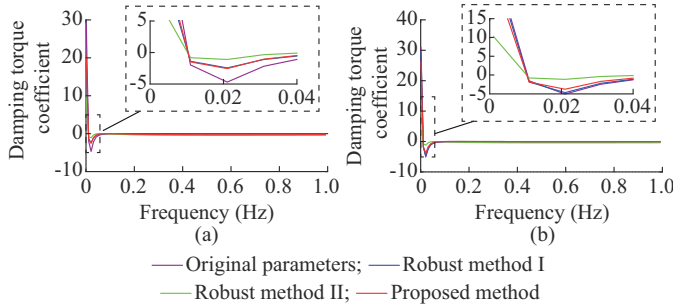| Method	Real part	Imaginary part	Damping (%)
Original parameters	0.0005	0.54	−0.093
Robust method I	−0.0410	0.59	6.930
Robust method II	−0.0540	0.55	9.770
Proposed method	−0.0920	0.52	17.420

Fig. 10.   Dynamic characteristic of G1. (a) Damping curve. (b) Step response.
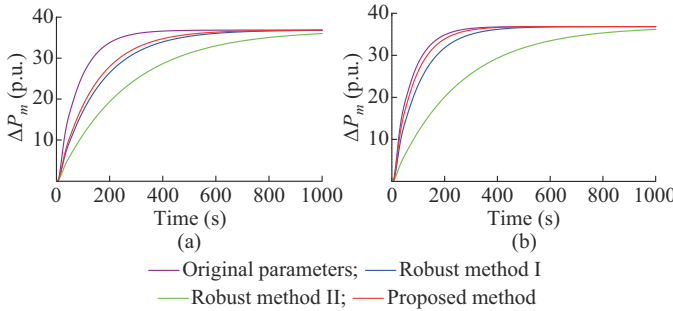

Fig. 11.   Step responses of PFR. (a) G1. (b) G2.

C. Robustness to Various Operating Conditions

To test the robustness of all three methods under different disturbances, four different two-phase short-circuit faults are applied with 100 ms fault duration and the faulty buses are 8, 16, 23, and 25, which are named as fault 2, fault 3, fault 4, and fault 5, respectively. The simulation results are shown in Fig. 12. It can be observed that all methods can damp out the oscillations. Among them, the proposed method takes the least time to damp out the oscillations as compared with the other two robust methods. The proposed method also has the smallest oscillation magnitude, demonstrating the best performance of damping out UFLO under different faults.

Furthermore, different T_w settings for Cases 2-5 in Table IV are taken as comparative cases to test the robustness of the proposed method under different cases.
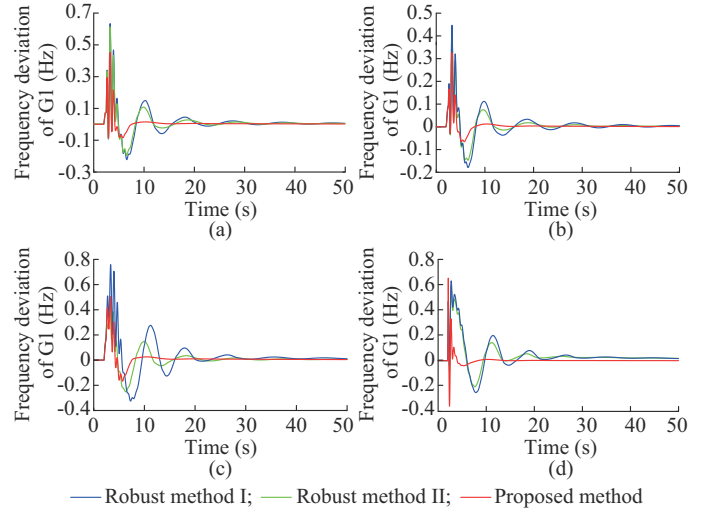

Fig. 12.   Dynamic response of G1 under different disturbances. (a) Fault 2. (b) Fault 3. (c) Fault 4. (d) Fault 5.

For each case, time-domain simulations are carried out with different PID parameter settings. The results are shown in Fig. 13.
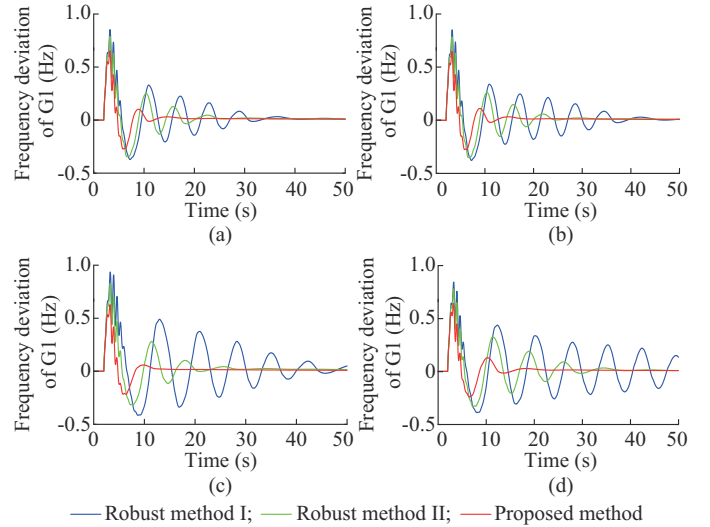

Fig. 13.   Dynamic response of G1 under different cases. (a) Case 2. (b) Case 3. (c) Case 4. (d) Case 5.

It can be concluded that the PID parameters optimized by the proposed method can achieve better performance in comparison with two other optimization methods. To further investigate the performance of the proposed method, we adopt the Monte Carlo method to sample time constant T_w of turbines between 0.5 s and 3 s to generate 300 cases. For each case, the ULFO mode is calculated and the probability density function (PDF) of ULFO nodes is also calculated, as shown in Fig. 14. Note that the red curve denotes the PDF distribution and is formed by fitting the eigenvalues produced via different PID parameter tuning method. It can be observed that before PID parameter optimization, the damping of the ULFO mode is less than 0 for some cases and not located in the secure region ($\xi > 5\%$). This means that there is a risk of ULFO induced instability during the operation of

the system. Figure 14(b), (c), and (d) shows the distributions of ULFO modes under different PID parameter settings. The statistical hypothesis for damping ratio of ULFO mode under 300 operating conditions is listed in Table VI. Clearly, compared with two other robust methods, the proposed method can make the ULFO mode achieve better mean value and smaller standard deviation of the damping ratio, yielding a larger stability margin.
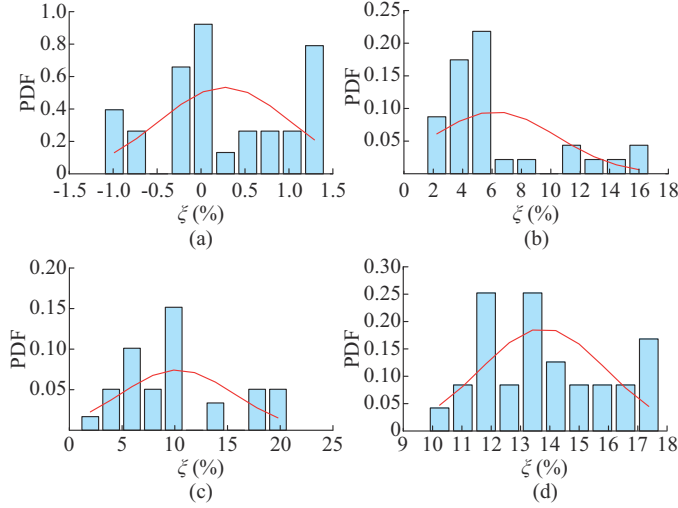

Fig. 14.   Damping PDF of ULFO mode under different cases. (a) Original parameters. (b) Robust method I. (c) Robust method II. (d) Proposed method.

TABLE VI
STATISTICAL HYPOTHESIS FOR DAMPING RATIO OF ULFO MODE UNDER
300 OPERATING CONDITIONS

Method	Mean value (%)	Standard deviation
Original parameters	0.27	0.0075
Robust method I	6.18	0.0420
Robust method II	10.25	0.0536
Proposed method	13.78	0.0213

VI. ACTUAL HYDROPOWER-DOMINATED SYSTEM

In this section, an actual hydropower-dominated system in Sichuan Power Grid, China is introduced as a studied system, which is shown in Fig. 15. The abbreviations shown in Fig. 15 are the acronyms of the region names. This system consists of 53 buses, 26 lines, and 17 hydropower units and the installed capacity is 1100 MW. For such system, the eigenvalue analysis results show that a ULFO mode $-0.0013 + j0.597$ exists in this system. To damp out that oscillation, the proposed method is employed to optimize the governor PID parameters, which are strongly related to the ULFO (CJB-1, CJB-2, CPQ-3, CTCH-2).

A. Further Evaluation of Proposed Method

To test the effectiveness of the proposed method under different datasets, we generate three datasets via different distribution functions. Specifically, we generate three datasets by sampling from the upper- and lower-limits of PID parameter settings of each hydroturbine via the uniform distribution

function, normal distribution function, and rayleigh distribution function, respectively. Each dataset is utilized to train a DDPG agent. The training process is shown in Fig. 16. The curves and shaded areas in Fig. 16 represent the mean and variance of the reward function during the training process, respectively.
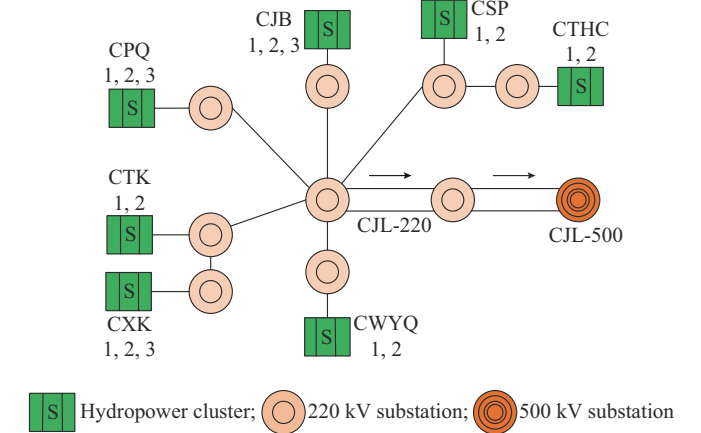

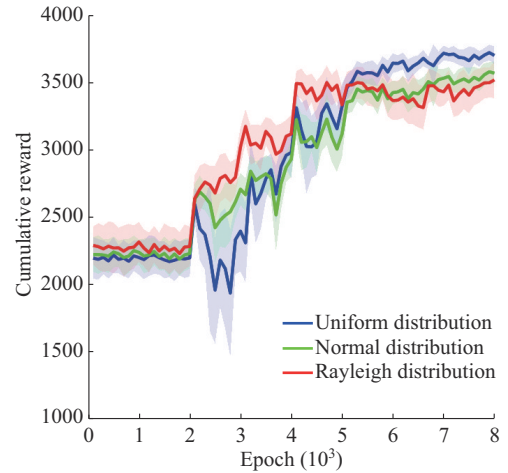Fig. 15.   Simplified system topology used for case.

Fig. 16.   Cumulative reward changes with episode under different datasets.

It can be observed from Fig. 16 that, with the increase of training episodes, the proposed method would converge under three different datasets. It means that the DDPG agent can learn the control policy that maximizes the cumulative reward. Then, these well-trained agents can be utilized to assist PSO to find the appropriate PID parameter settings to prevent ULFO. The effectiveness analysis of these obtained PID parameter settings is provided in the next part.

B. Control Performance Analysis and Comparison

To test the control performance of obtained PID parameter settings, we send the PID parameter settings to the system and compare these settings under a fault, where a two-phase short-circuit faults are applied with 150 ms fault duration and the faulty buses are CJL-220. The simulation is carried out under two different cases and the results are shown in Fig. 17. Clearly, each group of PID parameter settings obtained via different datasets can prevent ULFO phenomena under all three operating conditions, which means that the

proposed method can work effectively under different datas-
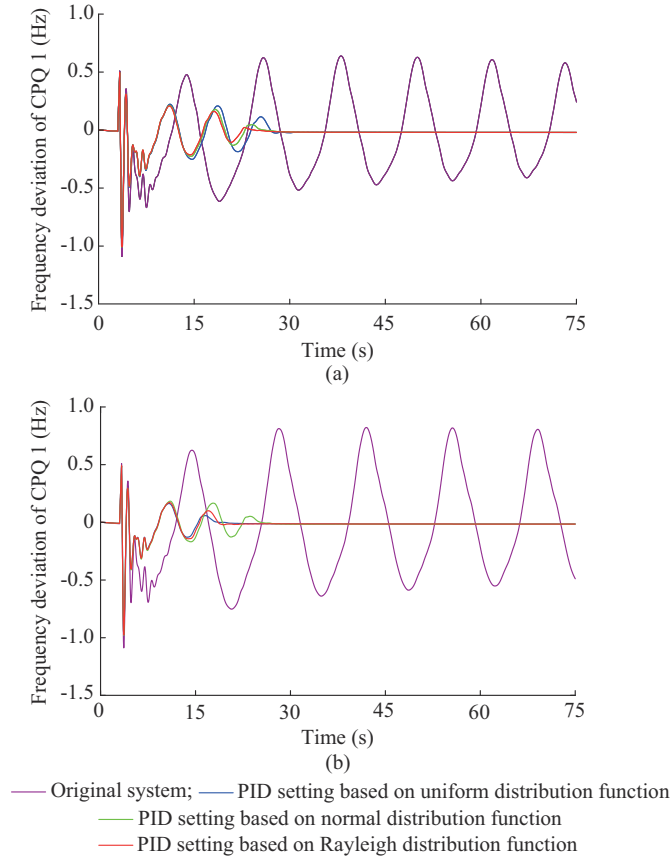et and the appropriate PID parameter settings can be found
to prevent ULFO.

Fig. 17.   Comparison results of different PID parameter settings obtained
via different datasets. (a) $T_w=2.5$. (b) $T_w=3.0$.

Moreover, to further investigate the performance of the
proposed method, the common PID parameter tuning meth-
od (Ziegler-Nichols) is applied to act as benchmark method.
The details of this method can be found in [30]. The compar-
ison results of different PID tuning methods can be found in
Fig. 18. It can be observed that both the Ziegler-Nichols
method and the proposed method can damp out the frequen-
cy oscillations. Among them, the proposed method can make
the frequency deviation curve recover to the steady point
faster. It means that the proposed method can search better
PID parameter settings and has better control performance.

VII. Conclusion

In this paper, a novel bi-level robust parameter optimiza-
tion model is proposed to re-tune the PID parameters to con-
trol ULFO. Different from the conventional robust optimiza-
tion methods, the PID parameter optimization is reformulat-
ed into the form of a min-max optimization model to ensure
the effectiveness of the optimized PID parameters under vari-
ous extreme operating conditions. The DDPG is developed
to train an agent for the fast decision making of lower-level
model and ensure the effective interactions with the upper-
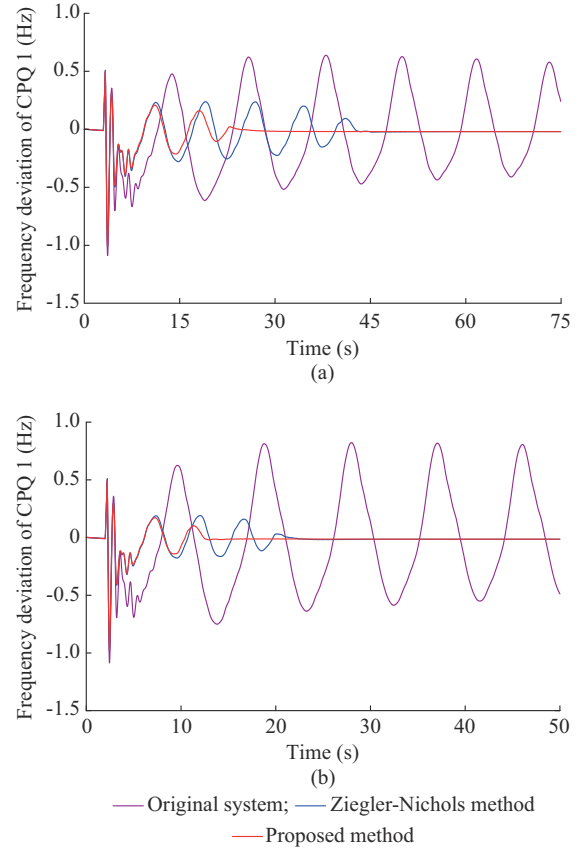level model, significantly improving the computational effi-
ciency.

Fig. 18.   Comparison results of different PID tuning methods. (a) $T_w=2.5$.
(b) $T_w=3.0$.

After the agent is trained, each iteration only needs 0.05 s
to provide actions for the upper-level model. Simulation re-
sults are carried out in IEEE 10-machine 39-bus system and
actual hydropower-dominated system in Sichuan Power
Grid. Three hundred cases are generated by Monte Carlo
method to act as test cases and the comparison results show
that the proposed method can achieve better damping control
performance. The mean value of the ULFO mode damping
under 300 cases can reach 13.78%. The mean value of other
two robust methods can only reach 6.18% and 10.25%, re-
spectively. It means that the PID parameters optimized by
the proposed method can achieve better damping control per-
formance in comparison with two other robust optimization
methods under different operating conditions.

Appendix A

The standard PFR model of the hydropower unit is shown
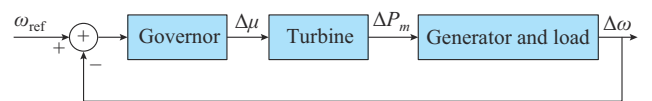in Fig. A1.

Fig. A1.   Structure diagram of standard PFR model of hydropower unit.

The parameter settings of hydropower unit, thermal power
unit, and AC lines are listed in Tables AI-AIII, respectively.
The parameter setting of generators in two-machine system

is listed in Table AIV. The rated active and reactive power of generators and load are listed in Table V.

TABLE AI
PARAMETER SETTING OF HYDROPOWER UNIT

Variable	Value	Variable	Value
K_P	4.0	b_p	0.05
K_I	2.5	D	2.00
K_D	0.5	T_G	0.20
T_J	10.0	T_w	4.00

TABLE AII
PARAMETER SETTING OF THERMAL POWER UNIT

Variable	Value
K_a	1.0
T_{ch}	0.2
T_g	0.3

TABLE AIII
PARAMETER SETTING OF AC LINE

AC line	R	X	$B/2$
x_1	0.0057	0.0625	0.5145
x_2	0.0032	0.0323	0.2806

TABLE AIV
PARAMETER SETTING OF GENERATORS IN TWO-MACHINE SYSTEM

Parameter	G_1	G_2
X_d	0.4400	1.0870
X_d'	0.2500	0.2890
X_d''	0.2200	0.2020
X_q	1.7000	0.6840
X_q'	0.5000	0.6870
X_q''	0.6000	0.2280
X_2	0.0025	0.2150
R_a	0.0250	0
T_{d0}'	8.0000	10.0300
T_{d0}''	0.0300	0.0400
T_{q0}'	0.4000	0.2590
T_{q0}''	0.0500	0.0600
H	13.000	12.0000
D	0	0

TABLE AV
RATED ACTIVE AND REACTIVE POWER OF GENERATORS AND LOAD

Generator and load	P (MW)	Q (Mvar)
G1	700	185
G2	700	185
Load 1	600	100

The detailed expressions of Routh-Hurwitz criterion coefficients a_1, a_2, a_3, and a_4 are shown as follows:

$$
\begin{cases}
a_1 = \dfrac{0.5T_G T_w T_J b_p + 0.5 T_w T_J N_0 + T_G T_J N_0 + 0.5 T_G T_w DN_0 - T_w N_1}{0.5 T_G T_w T_J N_0} \\[2ex]
a_2 = \dfrac{T_J b_p (0.5 T_w + T_G) + T_w (0.5 T_G D b_p - N_2) +}{0.5 T_G T_w} \rightarrow \\[2ex]
\quad \leftarrow \dfrac{N_0 (T_J + 0.5 T_w D + T_G D)}{T_J N_0} \\[2ex]
a_3 = \dfrac{N_2 - T_w + b_p (T_J + 0.5 T_w D + T_G D) + DN_0}{0.5 T_G T_w T_J N_0} \\[2ex]
a_4 = \dfrac{1 + D b_p}{0.5 T_G T_w T_J N_0} \\[2ex]
N_o = 1/K_I \\[1ex]
N_1 = K_D / K_I \\[1ex]
N_2 = K_P / K_I
\end{cases}
\tag{A1}
$$

The calculation process of oscillation frequency of two-machine system can be given.

Considering two-machine system is unstable, a dominant oscillation mode exists in the system. In this context, (5) can be represented as:

$$(s - \lambda_1)(s - \lambda_2)(s + \sigma + j\omega)(s + \sigma - j\omega) = 0 \tag{A2}$$

That is further rewritten as:

$$
s^4 + (2\sigma - \lambda_1 - \lambda_2) s^3 + [\lambda_1 \lambda_2 + \sigma^2 + \omega^2 - 2\sigma(\lambda_1 + \lambda_2)] s^2 +
$$
$$
[(\sigma^2 + \omega^2)(\lambda_1 + \lambda_2) + 2\sigma \lambda_1 \lambda_2] s + (\sigma^2 + \omega^2)\lambda_1 \lambda_2 = 0
\tag{A3}
$$

where λ_1, λ_2, and $\sigma \pm j\omega$ are the poles of the closed-loop transfer function. Note that λ_1 and λ_2 can be two real roots or a pair of the conjugate complex root. $\sigma \pm j\omega$ is the dominant pole. When $\sigma = 0$, the frequency response will oscillate critically. Meanwhile, we have:

$$
\begin{cases}
a_1 = -(\lambda_1 + \lambda_2) \\
a_3 = -\omega^2 (\lambda_1 + \lambda_2)
\end{cases}
\tag{A4}
$$

Based on (10), the oscillation frequency ω is defined as:

$$\omega = \sqrt{a_3/a_1} \tag{A5}$$

Combining (A5) and (A1), ω can be rewritten as:

$$
\omega = \sqrt{\dfrac{2[N_2 - T_w + b_p(T_J + 0.5 T_w D + T_G D) + DN_0]}{T_G T_w T_J b_p + T_w T_J N_0 + T_G T_J N_0 + T_G T_w DN_0 - 2T_w N_1}}
\tag{A6}
$$

The obtained PID parameters of IEEE 10-machine 39-bus system is shown in Table AVI.

TABLE AVI
PID PARAMETER SETTING

Method	Generator	K_P	K_I	K_D
Robust method I	G1	8.36	0.34	2.33
	G2	2.77	0.04	1.18
	G5	17.90	0.09	2.66
	G8	1.54	0.18	1.10
Robust method II	G1	8.61	0.29	2.69
	G2	2.81	0.26	1.74
	G5	11.98	0.32	0.60
	G8	6.59	0.05	1.63
Proposed method	G1	6.73	0.33	2.07
	G2	2.28	0.11	1.39
	G5	13.58	0.18	1.52
	G8	2.68	0.21	1.73

REFERENCES

[1] Z. Arthouros. (2019, Feb.). The renewables 2019 global status report. [Online]. Available: http://www.ren21.net/gsr-2019/pages/foreword/foreword/

[2] C. Jiang, J. Zhou, P. Shi *et al*., "Ultra-low frequency oscillation analysis and robust fixed order control design," *International Journal of Electrical Power & Energy Systems*, vol. 104, pp. 269-278, Jan. 2019.

[3] W. Yang, J. Yang, R. Tang *et al*., "Experimental investigation of theoretical stability regions for ultra-low frequency oscillations of hydropower generating systems," *Energy*, vol. 186, pp. 1-11, Nov. 2019.

[4] Z. Liu, W. Yao, J. Wen *et al*., "Effect analysis of generator governor system and its frequency mode on inter-area oscillations in power systems," *International Journal of Electrical Power & Energy Systems*, vol. 96, pp. 1-10, Mar. 2018.

[5] W. Mo, Y. Chen, Y. Liu *et al*., "Analysis and measures of ultralow-frequency oscillations in a large-scale hydropower transmission system," *IEEE Journal of Emerging & Selected Topics in Power Electronics*, vol. 6, no. 3, pp. 1077-1085, Sept. 2018.

[6] H. Pico, J. D. McCalley, A. Angel *et al*., "Analysis of very low frequency oscillations in hydro-dominant power systems using multi-unit modeling," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1906-1915, Nov. 2012.

[7] R. Xie, I. Innocent, D. Rimorov *et al*., "Fundamental study of common mode small-signal frequency oscillations in power systems," *International Journal of Electrical Power & Energy Systems*, vol. 106, pp. 201-209, Jan. 2019.

[8] Y. Chen, Y. Liu, Z. Tang *et al*., "Analysis of ultra-low frequency oscillation in yunnan asynchronous sending system," in *Proceedings of IEEE PES General Meeting*, Chicago, USA, Jul. 2017, pp. 1-5.

[9] G. Wang, X. Zheng, X. Guo *et al*., "Mechanism analysis and suppression method of ultralow-frequency oscillations caused by hydropower units," *International Journal of Electrical Power & Energy Systems*, vol. 103, pp. 102-114, Dec. 2018.

[10] H. N. Villegas, "Electromechanical oscillations in hydro-dominant power systems: an application to the Colombian power system," M.S. thesis, Department of Electrical and Computer Engineering, Iowa State University, Ames, USA, 2011.

[11] K. Sebaa, Y. Zhou, Y. Li *et al*., "Low-frequency oscillation damping control for large-scale power system with simplified virtual synchronous machine," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 6, pp. 1424-1435, Nov. 2021.

[12] G. Zhang, W. Hu, D. Cao *et al*., "Deep reinforcement learning based approach for proportional resonance power system stabilizer to prevent ultra-low-frequency oscillations," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5260-5272, Nov. 2020.

[13] R. Grondin, I. Kamwa, L. Soulieres *et al*., "An approach to PSS design for transient stability improvement through supplementary damping of the common low-frequency," *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 954-963, Aug. 1993.

[14] P. Puri and S. Ghosh, "A hybrid optimization approach for PI controller tuning based on gain and phase margin specifications," *Swarm & Evolutionary Computation*, vol. 8, pp. 69-78, Feb. 2013.

[15] G. Chen, F. Tang, H. Shi *et al*., "Optimization strategy of hydro-governors for eliminating ultralow-frequency oscillations in hydro-dominant power systems," *IEEE Journal of Emerging & Selected Topics in Power Electronics*, vol. 6, no. 3, pp. 1086-1094, Sept. 2018.

[16] L. Chen, X. Lu, Y. Min *et al*., "Optimization of governor parameters to prevent frequency oscillations in power systems," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4466-4474, Jul. 2018.

[17] D. J. White and G. Anandalingam, "A penalty function approach for solving bi-level linear programs," *Journal of Global Optimization*, vol. 3, no. 4, pp. 397-419, Jul. 1993.

[18] D. Bertsimas, E. Litvinov, X. Sun *et al*., "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 52-63, Feb. 2013.

[19] Y. Du, L. Jiang, Y. Li *et al*., "A robust optimization approach for demand side scheduling considering uncertainty of manually operated appliances," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 743-755, Mar. 2018.

[20] S. Mei, W. Guo, Y. Wang *et al*., "A game model for robust optimization of power systems and its application," *Proceedings of the CSEE*, vol. 33, no. 19, pp. 47-56, Jul. 2013.

[21] M. Netto and L. Mili, "Robust data filtering for estimating electromechanical modes of oscillation via the multichannel prony method," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4134-4143, Jul. 2018.

[22] G. Zhang, W. Hu, D. Cao *et al*., "A novel deep reinforcement learning enabled sparsity promoting adaptive control method to improve the stability of power systems with wind energy penetration," *Renewable Energy*, vol. 178, pp. 363-376, Nov. 2021.

[23] L. Bo, L. Han, C. Xiang *et al*., "A Q-learning fuzzy inference system based online energy management strategy for off-road hybrid electric vehicles," *Energy*, vol. 252, pp. 1-11. Aug. 2022.

[24] Z. Yan and Y. Xu, "Data-driven load frequency control for stochastic power systems: a deep reinforcement learning method with continuous action search," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1653-1656, Mar. 2019.

[25] J. Duan, D. Shi, R. Diao *et al*., "Deep-reinforcement-learning-based autonomous voltage control for power grid operations," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 814-817, Jan. 2020.

[26] Y. Li, W. Gao, S. Huang *et al*., "Data-driven optimal control strategy for virtual synchronous generator via deep reinforcement learning approach," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 4, pp. 919-929, Jul. 2021.

[27] G. Zhang, W. Hu, D. Cao *et al*., "A data-driven approach for designing STATCOM additional damping controller for wind farms," *International Journal of Electrical Power & Energy Systems*, vol. 117, p. 105620, May 2020.

[28] J. B. Park, K. S. Lee, J. R. Shin *et al*., "A particle swarm optimization for economic dispatch with non-smooth cost functions," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 34-42, Feb. 2005.

[29] H. Chen and Z. Xu, "Comparison of mathematical models for transient stability calculation in PSASP and PSS/E and corresponding calculation results," *Power System Technology*, vol. 28, no. 5, pp. 1-12, Mar. 2004.

[30] J. G. Ziegler and N. B. Nichols, "Optimum setting for automatic controllers," *Journal of Dynamic Systems Measurement and Control*, vol. 64, pp. 759-768. Nov. 1942.

Guozhou Zhang received the B.S. from Chongqing University of Technology, Chongqing, China, in 2016, and the M.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2019. He is currently working toward the Ph. D. degree in control science and engineering at the University of Electronic Science and Technology of China. His research interest includes power system analysis and control.

Junbo Zhao received the Ph. D. degree from the Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University (Virginia Tech), Falls Church, USA, in 2018. He is an Assistant Professor with the Department of Electrical and Computer Engineering, University of Connecticut, Mansfield, USA. He was an Assistant Professor and a Research Assistant Professor with Mississippi State University, Starkville, USA, and Virginia Tech from 2019 to 2021 and from 2018 to 2019, respectively. His research interests include cyber-physical power system modeling, monitoring, uncertainty quantification, learning, dynamics, stability control, and cyber security with distributed energy resources.

Weihao Hu received the B.Eng. and M.Sc. degrees in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2004 and 2007, respectively, and the Ph. D. degree from Aalborg University, Aalborg, Denmark, in 2012. He is currently a Full Professor and the Director of the Institute of Smart Power and Energy Systems, University of Electronics Science and Technology of China, Chengdu, China. His research interests include artificial intelligence in modern power system and renewable power generation.

Di Cao received the Ph. D. degree from University of Electronic Science and Technology of China, Chengdu, China, in 2021. He is currently a Postdoctoral Researcher at University of Electronic Science and Technology of China. His research interests include optimization of distribution network and application of machine learning in power systems.

Nan Duan received the B.S. degree in automation from the Beijing University of Technology, Beijing, China, in 2010, the M. Eng. degree in control engineering from Beihang University, Beijing, China, in 2013, and the Ph.D. degree in electrical engineering from The University of Tennessee, Knoxville, USA, in 2018. He is currently a Research Staff at the Lawrence Livermore National Laboratory, Livermore, USA. His current research interests include power system modeling, high-performance computing, parallel-in-time methods, nonlinear systems, numerical integration methods, machine learning, and synchrophasor applications

Zhe Chen received the B.Eng. and M.Sc. degrees from the Northeast China Institute of Electric Power Engineering, Jilin, China, and the Ph.D. degree from the University of Durham, Durham, U.K.. He is a Full Professor with the Department of Energy Technology, Aalborg University, Aalborg, Denmark. His research interests include power systems, power electronics and electric machines, wind energy, and modern power systems.

Frede Blaabjerg received the Ph. D. degree in electrical engineering from Aalborg University, Aalborg, Denmark, in 1995. He was with ABB-Scandia, Randers, Denmark, from 1987 to 1988. He became an Assistant Professor, in 1992, an Associate Professor in 1996, and a Full Professor of power electronics and drives in 1998. In 2017, he became a Villum Investigator. He is Honoris Causa at University Politehnica Timisoara, Timisoara, Romania, and Tallinn Technical University, Tallinn, Estonia. His current research interests include power electronics and its applications such as in wind turbines, PV systems, reliability, harmonics and adjustable speed drives.