

ORIGINAL RESEARCH

Open access

Artificial Neural Network-Based Frequency Control with Reinforcement Learning for AC Microgrids under Communication Delays

Elena Petrova^{1*}, Ivan Georgiev¹

Abstract

Advanced control strategies have been applied to manage variable power output in standalone microgrids, reducing the risk of frequency instability and power-quality deterioration. Nevertheless, a key difficulty arises: these computationally intensive techniques, when used without stability checks, may fail to cope with the fast dynamics of real-world conditions, particularly under communication delays. This research proposes a versatile real-time solution for frequency regulation that relies on an artificial neural network (ANN) confined to stable operating regions. The proposed method combines stabilizing PID controllers—derived via small-signal analysis—with an automated optimization of ANN weights and reinforcement learning (RL)-guided constraint selection. Initially, stabilizing PID controllers are developed using the stability boundary locus approach and the Mikhailov criterion, with explicit consideration of communication delays. Subsequently, controller parameters are fine-tuned in real time via an automated search for optimal coefficient sets, using a constrained ANN to keep frequency deviations within a predefined parameter space. The framework is further strengthened by an RL-based tuning mechanism trained on an interpolated stability boundary curve, guaranteeing both stability and enhanced performance. This unique integration of ANN, RL, and sophisticated PID tuning represents a significant advancement in addressing frequency control challenges in isolated AC microgrids. Experimental results demonstrate that the proposed approach surpasses conventional techniques thanks to its reduced parameter search domain. Specifically, it achieves greater reductions in both transient and steady-state frequency deviations compared to semi-constrained and unconstrained alternatives. The enhanced performance metrics and stability assessments confirm that the method delivers superior system behavior and robustness under varying operating conditions.

Keywords Reinforcement learning, AC microgrid, Constrained artificial neural networks, Frequency control, Stabilizing controller

*Correspondence:

Elena Petrova
elena.petrova@uni-sofia.bg

¹ Department of Energy Informatics and Optimization, Faculty of Engineering, University of Sofia, Sofia, Bulgaria

Introduction

The technical challenges associated with renewable energy sources (RESs) require the development of standards for frequency regulation and properly designed control systems, all aimed at achieving optimal operational performance [1]. These standards are incorporated into microgrid (MG) architectures, which have been developed to tackle three primary issues: reliability, environmental concerns, and economic considerations [2]. In MGs, the main generation units—such as diesel engine generators (DEGs), photovoltaic (PV) arrays, wind turbine generators (WTGs), and fuel cells (FCs)—are deployed and interfaced with the grid on the consumer side, collectively known as the distributed generation (DG) system

[3]. Local control of RESs through this system plays a vital role, as it helps minimize transmission losses and improve overall power system reliability [4]. From the utility viewpoint, an MG functions as an independent entity capable of operating in two modes: grid-connected and isolated [5]. In grid-connected mode, the main goals typically involve providing ancillary services to the main grid and optimizing internal energy use, while isolated MGs focus on powering remote loads or enhancing power quality in weak grids via deliberate disconnection [6].

In isolated mode, achieving balance between active and reactive power generation and demand becomes particularly challenging owing to the reduced inertia of microgrids (MGs) relative to the conventional main grid [7]. A common approach to overcoming this issue is to incorporate backup storage solutions, such as electrochemical or mechanical energy storage systems (ESSs), into the MG framework [8]. Without such measures, real-time applications would become impractical due to inherent instability [9]. Precise modeling remains a major obstacle due to the pronounced time-varying and nonlinear behavior of ESSs, especially lithium-ion batteries [10–12]. Meanwhile, ensuring stability requires tight regulation of two key parameters—voltage and frequency—through robust control strategies, which are essential given the intermittent nature of renewable energy sources (RESs) and the unpredictable nature of load fluctuations [13]. To achieve this, three main architectures—centralized, decentralized, and distributed—have been developed for frequency deviation control (FDC) systems within a hierarchical control framework [14]. However, tackling complex multi-objective regulation and optimization problems demands intelligent management systems, energy scheduling mechanisms, and adaptive control units equipped with advanced algorithms [15–18]. To improve robustness, intelligent PID-like controllers have been introduced for various islanded MG types through meta-heuristic optimization techniques, including genetic algorithm (GA) [19], particle swarm optimization (PSO) [20], fuzzy logic [3, 21, 22], black hole algorithm (BHA) [23], and artificial neural networks (ANN) [24–26], all focused on optimally tuning controller parameters for superior performance. For greater robustness, the application of modern adaptive and constrained population extremal optimization, combined with PID controllers, can markedly enhance the effectiveness and reliability of FDC [27, 28], resulting in more stable and efficient management of these intricate systems.

Previous studies and related works have largely neglected the effects of communication time delays on control parameters. In distributed control approaches, transmission delays pose a critical issue that can impair controller performance and threaten system stability in practical deployments [29, 30]. In particular, such delays can modify key dynamic properties, including oscillation modes, critical eigenvalues, and the overall closed-loop system response [31]. Therefore, a thorough stability analysis is vital when incorporating communication time delays into FDC system design, as understanding the interplay between delay and control parameters is indispensable [32]. Incorporating delay terms into the system's characteristic equation complicates stability analysis by shifting dynamics and altering stability margins, thereby requiring sophisticated methods such as frequency-domain techniques. At present, the dominant methods for assessing MG stability in the presence of time delays are frequency-domain [29, 33] and time-domain [34, 35] analyses, which have evolved alongside advancements in MG modeling [36]. While time-domain approaches (e.g., Lyapunov-based methods) are better suited for controller synthesis because they handle nonlinear, stochastic, and time-varying systems with delays [37], they typically provide only sufficient stability conditions, yielding conservative parameter designs and high computational demands in stiff power systems [38]. Conversely, frequency-domain analysis relies on small-signal stability tools, which are limited to linear or locally linearized nonlinear systems around an equilibrium point, remain valid only near hyperbolic equilibria, and address solely small disturbances [39].

Although earlier research has considered delays in MGs to some extent [40–42], these efforts have not focused on actively enforcing “dynamic adaptation” constraints or on enabling real-time FDC operation in challenging MG settings. Inspired by this research gap and the proven effectiveness of real-time controller parameter tuning reported in [43], we introduce a novel, intelligent, and robust PID-based FDC system tailored for AC MGs subject to communication delays. Accordingly, the primary contribution of this paper is the development of a method that utilizes a constrained ANN-based structure and an RL-based training approach to perform online tuning of controller parameters, incorporating the following key steps to avoid performance degradation in FDC and support its real-world implementation:

- Compute stabilizing PID controllers using the stability boundary locus (SBL) method and Mikhailov criterion, accounting for communication delays.
- Interpolate the stability boundary curve and develop an integrated RL-based technique that leverages this curve to automatically constrain the search space during training, while preserving both stability and performance.
- Construct a fully intelligent controller tuning system by optimizing the combination of ANN weights and constraints produced and chosen via the integrated RL technique, thereby identifying an optimal range of controller gains in light of output variations.

The proposed control strategy delivers several important advancements that substantially improve FDC in isolated AC microgrids:

- **Time-Delay-Resistant Algorithm:** The algorithm guarantees stable functioning and avoids controller failures or system instability in practical scenarios where delays are unavoidable.
- **Stochastic RL and SDDPG Integration:** We employ both stochastic reinforcement learning and SDDPG (stochastic deep deterministic policy gradient) to globally optimize the weights of an ANN-based PID controller tuner, increasing algorithmic adaptability.
- **Automated Search Space Restriction:** The approach automatically narrows the search space through interpolation of the stability boundary curve, facilitating more efficient discovery of tuned parameters during optimization.
- **Self-Tuning Capability:** A self-tuning mechanism is introduced to boost robustness, speed up disturbance recovery, and strengthen frequency stability, enabling the controller to dynamically adjust parameters in response to variations in system delays or disturbances.
- **Integrated Performance and Stability Validation:** An applied reward-based search mechanism achieves an effective trade-off between controller stability and performance.

Furthermore, this approach features a sophisticated reward function that strikes an optimal balance between operational stability and overall performance, thereby guiding the reinforcement learning agent's training predominantly toward secure, stable zones of microgrid behavior. In contrast to earlier hand-crafted and partially constrained techniques [44, 45], the proposed framework fully automates the enforcement of operational limits by interpolating stability boundaries. When compared with standard meta-heuristic optimization algorithms [3, 19-26], the real-time adaptive tuning strategy developed here successfully overcomes the challenges posed by communication latencies. As a result, it not only streamlines neural network training but also significantly improves system robustness, reduces restoration times, and enhances frequency regulation. This is achieved via an autonomous tuning mechanism that narrows the exploration space considerably while remaining fully compatible with the demands of the frequency deviation control (FDC) framework.

Model of the Isolated AC MG with Delay

Figure 1 shows the configuration of the isolated AC microgrid experiencing communication delays. It integrates wind turbine generators (WTGs), photovoltaic (PV) arrays, a conventional diesel engine generator (DEG), a fuel cell (FC) unit, and an energy storage system (ESS) elements—specifically battery energy storage system (BESS) and flywheel energy storage system (FESS)—all linked via power electronic converters that ensure electrical isolation and appropriate voltage levels. The nominal rated capacities of the main components—RESS (combined WTGs and PV), FC, DEG, and ESS—are 130 kW, 70 kW, 160 kW, and 90 kW, respectively.

The fuel cell system encompasses a DC/AC inverter, fuel supply modules, and connection hardware. Despite its inherent complexity, frequency-domain analysis can be adequately performed with a reduced third-order dynamic model [46]. The DEG primarily serves to deliver a reserved power capacity that functions as spinning reserve within the secondary frequency control (SFC) framework.

To mitigate the negative dynamic consequences arising from the highly variable and uncertain output of renewable energy sources, two practical control configurations for the SFC loop are evaluated. In the first case, only the DEG participates in frequency regulation without additional support. In the second case, the DEG operates in coordination with the FC unit, creating a more capable control structure that effectively counters large frequency excursions and thereby reduces mechanical and electrical stress on both the DEG and the energy storage units.

The energy storage system consists of two distinct technologies: FESS and BESS. The flywheel energy storage system uses a high-speed rotating mass to enable very fast charging and discharging of electrical energy. In contrast, the battery energy storage system stores energy chemically and helps maintain power balance while suppressing frequency oscillations. Accurate state estimation and control of batteries remain difficult tasks, largely because they require precise, high-fidelity models and reliable estimates of internal parameters [47]. Lithium-ion batteries, in particular, exhibit highly nonlinear behavior due to complex internal electrochemical interactions, underscoring the need for advanced modeling approaches to enable reliable performance predictions in practical battery applications [48]. In the present analysis, a simplified linear approximation has been employed for convenience. However, adoption of a higher-fidelity model would improve the fidelity and applicability of the resulting control strategy.

Figure 2 provides a simplified block diagram of the frequency response dynamics for the delayed isolated AC microgrid. Owing to the stochastic and uncontrollable nature of the power injections from PV, WTG, and the connected load, these elements are not included in the FDC controller synthesis. They are instead modeled as an aggregate disturbance input, given by:

$$P_{tot,d} = P_L - P_{PV} - P_{WTG}, \quad (1)$$

Where P_L , P_{PV} , and P_{WTG} represent the power consumed by the load and generated by the PV and WTG units, respectively, and $P_{tot,d}$ stands for the total disturbance power. The total load demand signal P_{tot} serves as the reference input for deriving the characteristic equation that will be analyzed in the subsequent section.

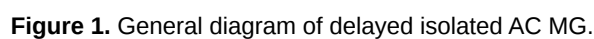
Taking into account the system's inherent damping and inertia, and applying the classical swing equation, the aggregated inertia and damping dynamics of the microgrid can be described as:

$$\Delta f = \frac{1}{2Hs + D} \Delta P_{tot}, \quad (2)$$

Where H is the equivalent inertia constant, Δf denotes the frequency deviation, D is the load damping coefficient, and s is the Laplace operator. Moreover, the net power imbalance ΔP is computed as the difference between generation and load, expressed as:

$$\Delta P_{tot} = P_{tot,g} - P_{tot,d}, \quad (3)$$

With $P_{tot,g} = P_{DEG} + P_{FC} + P_{BESS} + P_{FESS}$, where P_{DEG} , P_{FC} , P_{BESS} , and P_{FESS} correspond to the power outputs from the diesel generator and the fuel cell, battery, and flywheel storage units, respectively.



Building upon this concise introduction to the isolated AC microgrid and its frequency deviation control (FDC) needs in the presence of communication delays, the discussion now advances to the stability analysis. As depicted in [Figure 3](#), the methodology for the intelligent FDC system designates these initial activities as the analysis phase. The subsequent subsections will elaborate on the remaining elements of this phase, followed by the design phase, with particular emphasis on incorporating the ANN-based tuner, reward mechanisms, and the SDDPG algorithm—all aimed at maximizing cumulative rewards and achieving optimal performance. Lastly, the implementation phase will be outlined, including simulation-based validation to evaluate the outcomes and demonstrate superiority over conventional FDC approaches.

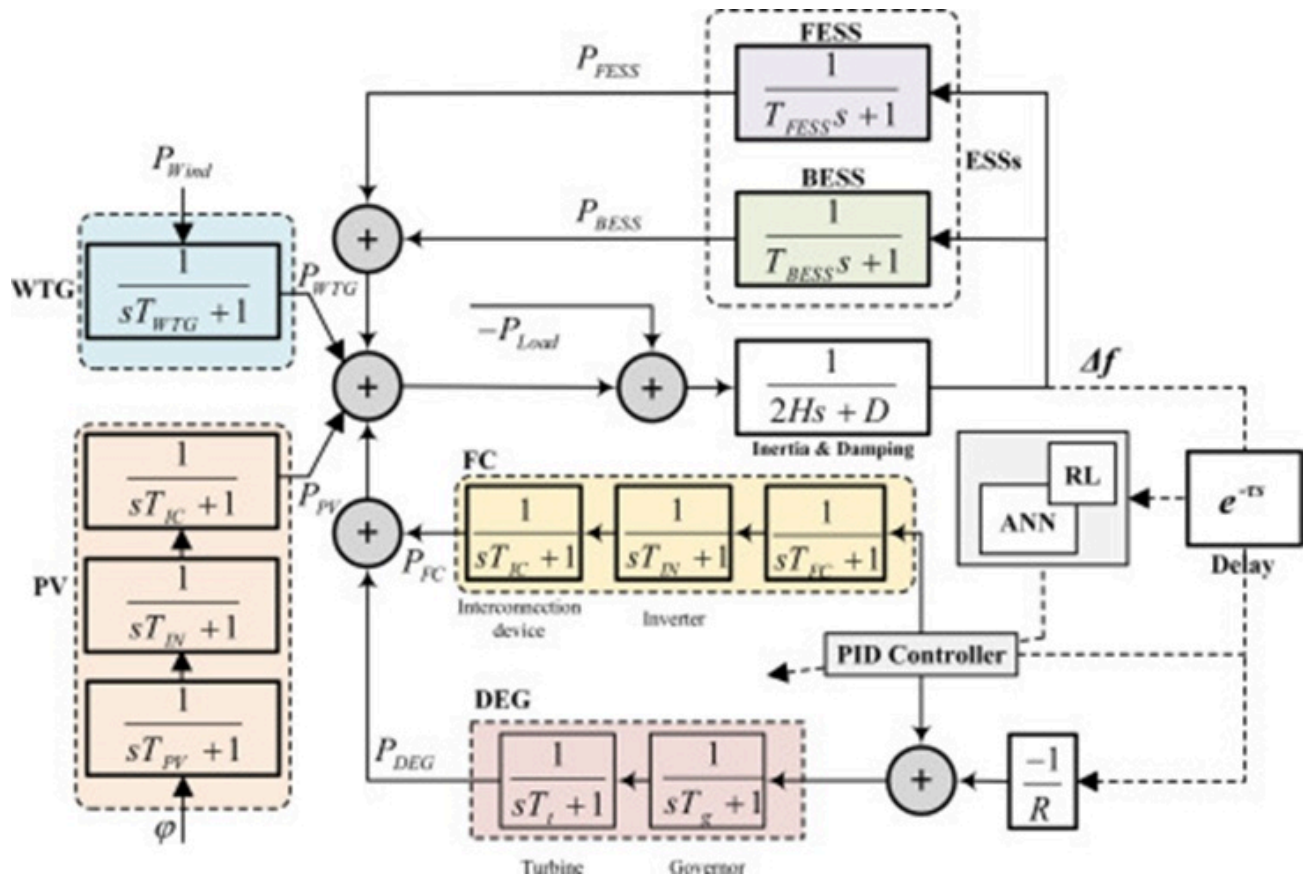


Figure 2. Frequency response diagram of the AC MG.

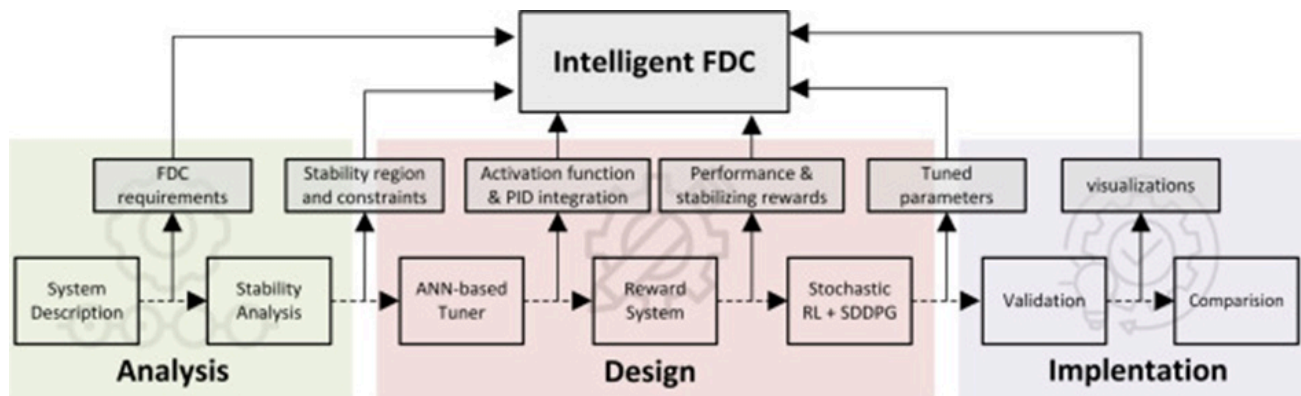


Figure 3. Overall methodology flowchart of the proposed intelligent FDC system.

Stability Analysis of the AC MG

To perform stability analysis of the AC microgrid and determine appropriate stabilizing controller parameters, the characteristic equation (s), which constitute the denominator of the closed-loop transfer function $\frac{C(s)}{P_{tot,d}} = \frac{\Delta f}{P_{tot,d}}$, is first derived. The roots of this equation are then examined in relation to the critical stability boundary (CRB). For this purpose, the parallel-form PID controller transfer function

$$G_{PID}(s) = K_p + K_i s^{-1} + K_d s, \quad (4)$$

Where K_p , K_i and K_d denote the proportional, integral, and derivative gains, respectively, is utilized to develop the FDC scheme for regulating the FC and DEG units. By aggregating the frequency-domain blocks shown in **Figure 2** and applying appropriate mathematical reductions, the entire system can be represented by a unified transfer function $G(s)$, whose denominator takes the form.

$$\Delta(s, \tau) = P(s) + Q(s) e^{-s\tau}, \quad (5)$$

Where τ represents the communication delay, and $P(s)$ and $Q(s)$ are real-coefficient polynomials of degrees 9 and 7, respectively,

$$\begin{aligned} P(s) &= \sum_{j=1}^9 p_j s^j, \\ Q(s) &= R \sum_{j=0}^7 q_j s^j = R \sum_{j=0}^5 q'_j s^j G_{PID}(s) s. \end{aligned} \quad (6)$$

Remark 1. The inclusion of the additional polynomial term(s) accounting for the delayed elements significantly increases the complexity of the closed-loop characteristic equation. The delay parameter τ introduces an additional dimension to the stability problem, influencing both the dynamic response and the stability margins. This renders the characteristic equation transcendental, complicating analytical treatment and solution procedures and requiring frequency-domain methods. Although the present study assumes a fixed and known delay τ , many practical situations involve uncertain or variable delays that require estimation, further challenging controller synthesis and underscoring the need for continued investigation in this area.

For the required stability investigations, the characteristic **Eq. 5** is reformulated by substituting $s = j\omega$ and $e^{-j\omega\tau} = \cos(\omega\tau) - j \sin(\omega\tau)$, with ω denoting the angular frequency. This substitution yields a complex polynomial equation structured as

$$\begin{aligned} \Delta(\omega, \tau, K_{p,i,d}) &= R(A_1(\omega, \tau) K_p + A_2(\omega, \tau) K_i + A_3(\omega, \tau) K_d) + A_4(\omega) \\ &+ j(R(B_1(\omega, \tau) K_p + B_2(\omega, \tau) K_i + B_3(\omega, \tau) K_d) + B_4(\omega)), \end{aligned} \quad (7)$$

Where:

$$\begin{aligned}
 A_1(\omega, \tau) &= ((q'_0 - q'_2\omega^2 + q'_4\omega^4)\omega) \sin(\bar{\omega}) \\
 &\quad - ((q'_1 - q'_3\omega^2 + q'_5\omega^4)\omega^2) \cos(\bar{\omega}), \\
 A_2(\omega, \tau) &= ((q'_1 - q'_3\omega^2 + q'_5\omega^4)\omega) \sin(\bar{\omega}) \\
 &\quad + (q'_0 - q'_2\omega^2 + q'_4\omega^4) \cos(\bar{\omega}), \\
 A_3(\omega, \tau) &= ((-q'_1 + q'_3\omega^2 - q'_5\omega^4)\omega^3) \sin(\bar{\omega}) \\
 &\quad - (\omega^2(q'_0 - q'_2\omega^2 + q'_4\omega^4)) \\
 &\quad \times \cos(\bar{\omega}), \\
 A_4(\omega) &= \omega^2(-p_2 + p_4\omega^2 - p_6\omega^4 + p_8\omega^6), \\
 B_1(\omega, \tau) &= (q'_0\omega - q'_2\omega^3 + q'_4\omega^5) \cos(\bar{\omega}) \\
 &\quad + (q'_1\omega^2 - q'_3\omega^4 + q'_5\omega^6) \sin(\bar{\omega}), \\
 B_2(\omega, \tau) &= (q'_1\omega - q'_3\omega^3 + q'_5\omega^5) \cos(\bar{\omega}) \\
 &\quad - (q'_0 - q'_2\omega^2 + q'_4\omega^4) \sin(\bar{\omega}), \\
 B_3(\omega, \tau) &= (-q'_1\omega^3 + q'_3\omega^5 - q'_5\omega^7) \cos(\bar{\omega}) \\
 &\quad + ((q'_0 - q'_2\omega^2 + q'_4\omega^4)\omega^2) \sin(\bar{\omega}), \\
 B_4(\omega) &= \omega(p_1 - p_3\omega^2 + p_5\omega^4 - p_7\omega^6 + p_9\omega^8),
 \end{aligned}$$

And $\bar{\omega} = \omega\tau$. Setting both the real and imaginary parts of Eq. 7 to zero allows the derivation of two controller parameters expressed in terms of K_d , τ , and the frequency ω . For any fixed value of K_d , the following relations hold:

$$\begin{aligned}
 K_p(\omega, \tau, K_d) &= \frac{A_2(B_3K_d + B_4) - (A_3K_d + A_4)B_2}{A_1B_2 - A_2B_1}, \\
 K_i(\omega, \tau, K_d) &= \frac{B_1(A_3K_d + A_4) - (B_3K_d + B_4)A_1}{A_1B_2 - A_2B_1}.
 \end{aligned} \quad (8)$$

By respecting the predefined stability limits given as:

$$\Delta(j\omega, K_{p,i,d}) = 0 \Rightarrow \begin{cases} \omega = \infty, & K_{p,i} \notin CRB, \\ 0 < \omega < \infty, & K_{p,i} \in CRB, \\ \omega = 0, & K_i = 0, \end{cases} \quad (9)$$

The corresponding stability boundary locus (SBL) curves can be plotted in the K_p – plane for different combinations of K_d and τ .

Constrained ANN-Based Tuner Design

After concluding the analysis phase, this section elaborates on the development of the proposed real-time ANN-based tuner, as outlined in **Figure 3**. Emphasis is placed on the seamless integration of the ANN-based tuner, the reward mechanisms, and the SDDPG algorithm, with the overarching goal of optimizing performance and achieving the highest possible cumulative rewards.

Actor ANN

Here, we describe the structure and functionality of the proposed real-time ANN-based tuner, shown in **Figure 4**, operating under stabilizing controller parameters. The presented ANN-based tuner functions as a PID-resembling controller, designed to maintain frequency stability while delivering optimal tuning for an underlying PID controller, specifically within a stability region where $K_d = 0$. The stability characteristics of this tuner are governed by the (non)linear properties of the chosen activation function and the trained weight values within the ANN. For instance, smooth nonlinear activation functions like *sigmoid*, *tansig*, *RBF*, and similar ones offer inherent disturbance rejection capabilities, unlike purely linear options such as *purelin* or *ReLU*. Given that negative coefficients are generally unsuitable for PID control applications, the activation function must enforce a non-negative output range. The custom activation function introduced in [26] is a modified *tansig* variant, expressed as

$$y(t) = \frac{m_\mu}{1 + e^{-|\theta_\mu x(t)|}} - m_\mu/2, \quad (10)$$

Where $x(t)$ and $y(t)$ represent the input and output of the activation function at time t , m_μ is a vector of per-neuron constraints that restrict the output signal to the interval $[0, m_\mu/2]$, and θ_μ is the matrix of actor ANN weights. These weights act as adjustable linear amplifiers, scaling the incoming signals and controlling the relative impact of each output on the overall tuning process. To preserve closed-loop stability, the relationships among K_p , K_d , and K_i must remain confined within the predefined stability region; any violation would compromise system stability. Moreover, the tuner incorporates bounds on both the activation function outputs and the weight values, with performance rigorously verified against the stability boundaries to ensure reliable and stable tuning behavior.

As illustrated in **Figure 4**, the proposed ANN-based tuner integrates the three core PID components—proportional, integral, and derivative—as inputs to the input layer neurons.

- **Proportional action:** This component responds directly to the instantaneous frequency deviation. It exhibits reduced sensitivity to disturbances and requires fewer tuning variables, yet it may fall short of fully covering the required controller adjustments.
- **Integral action:** This element accounts for the accumulated historical frequency deviation, enabling a gradual and smooth tuning process. It contributes to overall stability but results in slower response times, particularly when operating independently.
- **Derivative action:** Acting as predictive control, this part anticipates erroneous trends in frequency deviation and responds preemptively. Although it can amplify noise and potentially degrade tuning stability, it substantially accelerates the tuner's dynamic response.

The architecture includes tuning outputs specifically for the proportional and integral gains of the PID controller. Given the specific control requirements and dynamics of the studied microgrid system, the derivative component has been omitted in this implementation. However, it can be readily added to the control framework if needed.

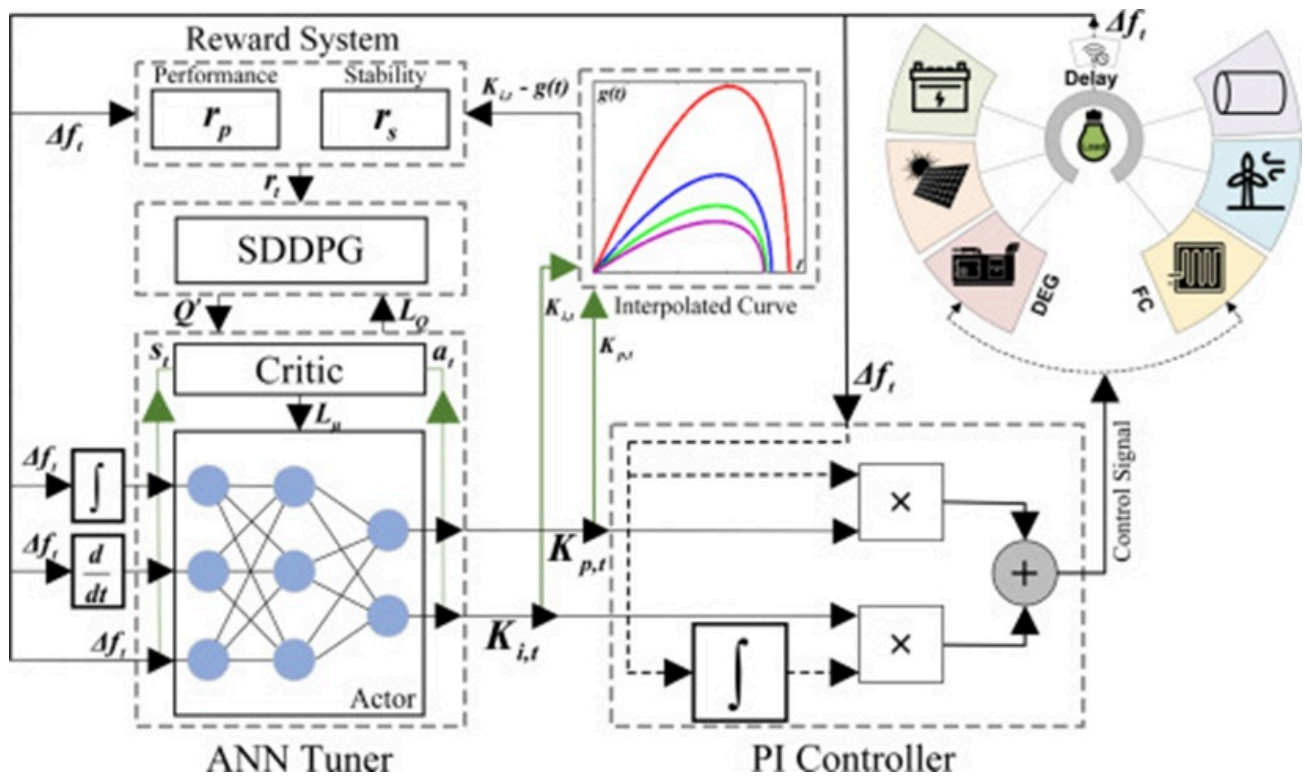


Figure 4. Architecture of the proposed real-time ANN-based tuner.

Stabilizing performance-based reward system

A novel reward mechanism is introduced, grounded in a stability verification procedure that evaluates the tuner's outputs, $K_p(t)$ and $K_i(t)$, against an interpolated stability boundary function g defined over the stability region in the (K_p, K_i) -plane relative to the SBL curve. Concurrently, the objective is to achieve superior control performance; thus, when (t) lies within the stable region, a positive reward is assigned to the controller, whereas excursions beyond this region trigger penalties that increase with the distance from the boundary. Broadly, the reward or penalty is determined by the deviation between the current gain pair $K(t)$ and the corresponding boundary point $KI(t) = g(KP(t))$, where KP and KI denote the nearest points on the SBL curve at time t , and g represents the interpolated stability boundary function. Accordingly, the overall reward signal r_{tot} integrates both a stabilizing reward component r_s and a performance-oriented reward component r_p , which can be formulated as follows:

- To enforce stability, the stabilizing reward component is defined as

$$r_s = \begin{cases} |K_i(t) - g(t)|, & \text{if } g(t) > K_i(t), \\ -|K_i(t) - g(t)|, & \text{if } g(t) < K_i(t). \end{cases} \quad (11)$$

- To maintain high controller performance, a supplementary reward component incorporates criteria derived from the angular function introduced in [49] and expressed as

$$r_p = \begin{cases} 2\sqrt{\frac{e^2}{e^2+t^2}}, & \text{if } \sin \phi < 0.707, \\ -10\sqrt{\frac{t^2}{e^2+t^2}}, & \text{if } \sin \phi > 0.707, \end{cases} \quad (12)$$

Where ϕ signifies the error angle and $e = \int_0^t |\Delta f| dt$ denotes the integral of the absolute frequency deviation error.

Stochastic RL with SDDPG policy

In conjunction with ANN, we propose the two-stage stochastic RL optimization approach method as follows: In the first stage, the actor ANN at state s_t , which is denoted as $(s_t | \theta_\mu)$, is initialized with random weight values $\sum_{j=1}^N \theta_{\mu_j}$ where θ_μ is an array of non-negative random values that are updated across a specified number of episodes N . Subsequently, the reward signal (a_t) is observed at each action $a_t = (K_p(\lambda), K_d(\lambda))$ and state $s_t = (\frac{\Delta f}{s}, \frac{d\Delta f}{dt})$. Finally, the weights θ_μ of the agent with the best total reward will be chosen as

$$\theta_{\mu_N} = \begin{cases} \theta_{\mu_{N+1}}, & \text{if } r_{tot} > r_{tot_max}, \\ \theta_{\mu_N}, & \text{if } r_{tot} < r_{tot_max}, \end{cases} \quad (13)$$

Where $r_{tot_max} = \text{argmax}(r(s_t, a_t))$.

The second phase employs the SDDPG algorithm—originally presented in [50] and subsequently adapted for control-oriented applications as detailed in [43, 51]—to refine neural network parameters and achieve superior control performance in the microgrid environment. The core optimization objective is to minimize the discrepancy between the target action-value function and the critic network's reward estimate. The target action-value function is given by

$$y_i \leftarrow r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1} | \theta_{\mu'}) | \theta_{Q'}), \quad (14)$$

Where $\theta_{\mu'}$ and $\theta_{Q'}$ represent the target actor and critic networks, respectively. The critic network's predicted reward for the state-action pair is denoted as $s_t | \theta_{Q'}$. Within this setup, the action a_t at timestep t for microgrid state s_t is produced by $s | \theta_{\mu'}$. The critic loss function L_Q is computed as

$$L_Q = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i | \theta_Q))^2. \quad (15)$$

Concurrently, the actor loss function L_μ is minimized to adjust the actor network weights θ_μ according to

$$L_\mu = \frac{1}{N} \sum_i \left(Q(s, a | \theta_Q) |_{s=s_i, a=\mu(s_i)} \right)^2. \quad (16)$$

This process refines the policy $\mu(s | \theta_\mu)$ to enhance overall control efficacy. The SDDPG method is tailored to maximize the total reward $rtot_max$ by identifying parameter configurations that deliver the greatest cumulative reward throughout training.

The maximum attainable reward, denoted $rtot_max = \max_i r_{i,tot}(s_i, a_i)$, is estimated by estimating rewards via the critic network's output $Q(s_t, a_t)$ (as in Eq. 14) and contrasting these estimates against actual observed rewards. The critic weights θ_Q govern the accuracy of reward predictions for given states s_t and actions a_t , thereby influencing the effectiveness of the actor's policy evaluation. Weight updates for the actor network (θ_μ) proceed via gradient descent, with periodic evaluation of the ANN-based PID controller following each adjustment. To boost prediction reliability, the optimal future reward estimate is preserved as $Q'(s_{i+1}, \mu'(s_{i+1} | \theta_\mu) | \theta_Q)$. For a comprehensive description of the procedure, consult Algorithm 1 and the flowchart in Figure 5, which depict the complete reinforcement learning workflow. This phased optimization approach is essential for enabling real-time adaptability in the control system while preserving computational tractability.

```

1: Initialize random tuner weights as  $\theta_\mu$ .
2: Initialize random tuner output constraints as  $m_\mu$ .
3: for  $j = 1$  to  $N$  do
4:   Update random weights  $\theta_{\mu_j}$  and neuron constraints  $m_{\mu_j}$  of agent.
5:   Execute  $a_j$ , observe  $r_j$  at each step  $j$  in common  $s_j$ .
6:   Calculate the rewards  $r_s$  and  $r_p$  and  $r_{tot} = r_s + r_p$ .
7:   if  $r_{tot} > r_{tot\_max}$  then
8:     Save weights of agent  $\theta_\mu$ 
9:   end if
10: end for
11: Initialize  $\theta_Q$ , and targets  $\theta_{\mu'}$  and  $\theta_{Q'}$ .
12: Define learning rates  $\alpha_Q$ ,  $\alpha_\mu$ , and smooth factor  $\beta \in [0, 1]$ .
13: for  $episode = 1$  to  $M$  do
14:   for  $t = 1$  to  $T$  do
15:     Receive primary process observation as  $s_t$ .
16:     Select  $a_t = \mu(s_t | \theta_\mu)$  based on  $Q(s_t, a_t | \theta_Q)$ .
17:     Execute action  $a_t$ , observe  $r_t$ , and next state  $s_{t+1}$ .
18:     Select the best  $Q'$  as  $Q'(s_{t+1}, \mu'(s_{t+1} | \theta_{\mu'})) | r_{tot\_max}$ .
19:     Set  $y_t$  of DDPG as  $y_t \leftarrow r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1} | \theta_{\mu'}) | \theta_{Q'})$ .
20:     Update the critic by minimizing  $L_Q = N^{-1} \sum_i (y_i - Q(s_i, a_i | \theta_Q))^2$ .
21:     Update the actor by minimizing  $L_\mu$  obtained as
        
$$L_\mu \leftarrow N^{-1} \sum_i (Q(s_i, a_i | \theta_Q) |_{s=s_i, a=\mu(s_i)})^2$$

22:     Update target networks:
        
$$\theta_{Q'} \leftarrow \beta \theta_{Q'} + (1 - \beta) \theta_Q \alpha_Q L_Q$$

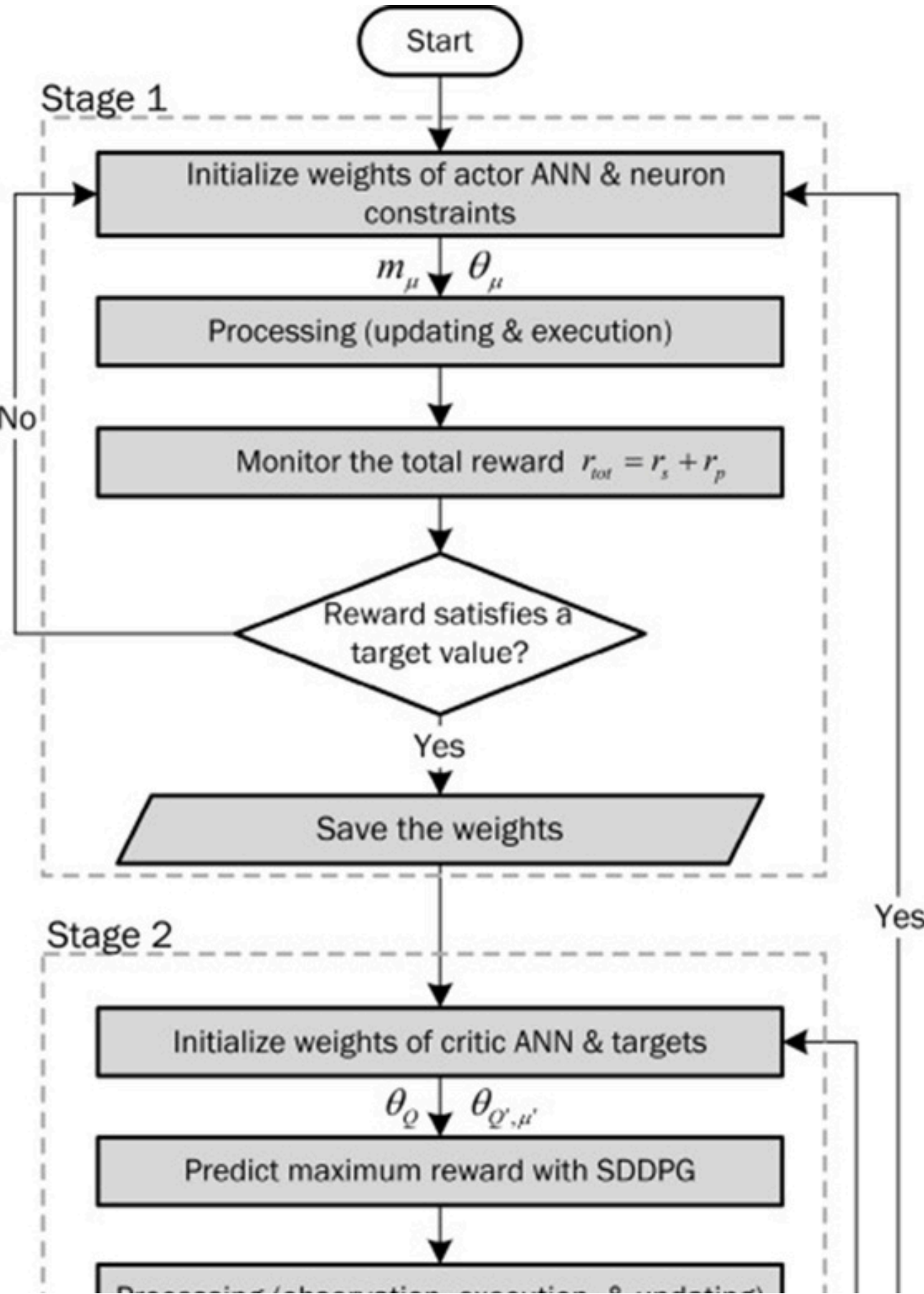
        
$$\theta_{\mu'} \leftarrow \beta \theta_{\mu'} + (1 - \beta) \theta_\mu \alpha_\mu L_\mu$$

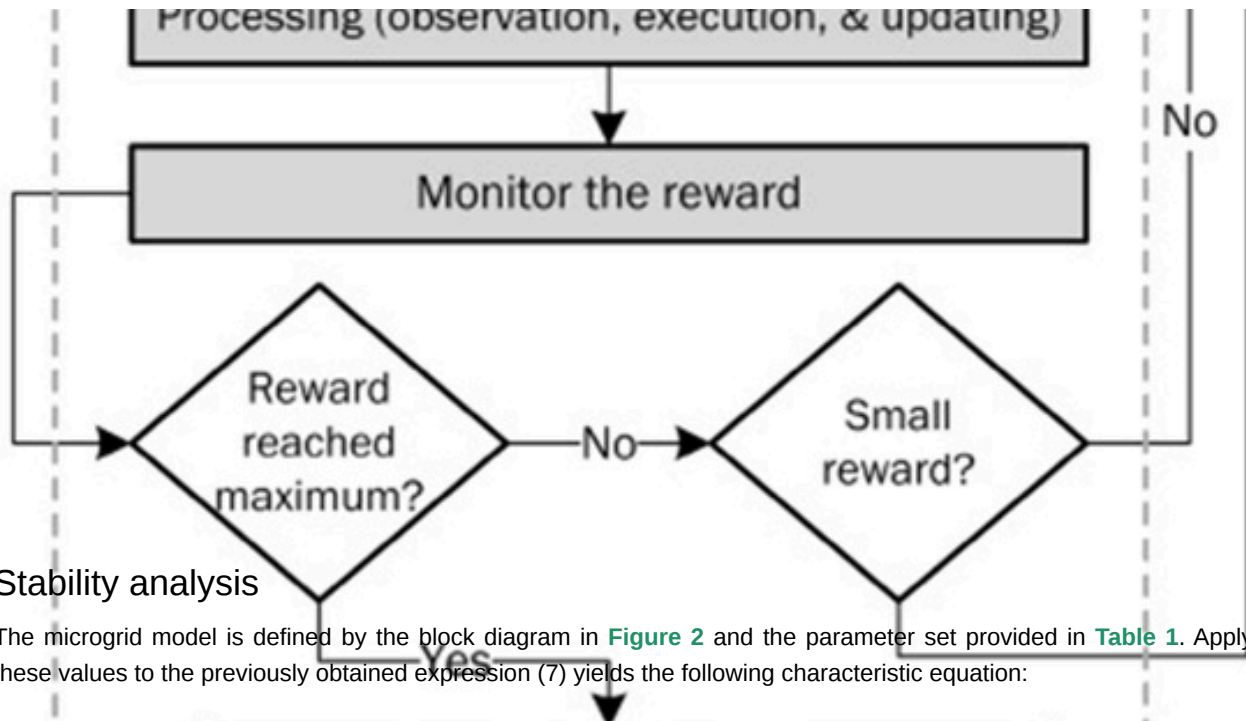
23:   end for
24: end for

```

Algorithm 1. SDDPG algorithm with stochastic RL optimization.

Remark 2. Although Bayesian optimization performs well when tuning static parameters, reinforcement learning was preferred in this work for its strengths in real-time adaptation, ongoing learning from interactions, and the ability to manage sophisticated, high-dimensional control challenges typical of microgrids. RL is particularly advantageous because it can respond continuously to dynamic variations in operating conditions and effectively balance competing objectives—such as robustness and performance—making it a more suitable choice than Bayesian methods, which tend to remain static and pursue only one primary goal.





Stability analysis

The microgrid model is defined by the block diagram in [Figure 2](#) and the parameter set provided in [Table 1](#). Applying these values to the previously obtained expression (7) yields the following characteristic equation:

$$\begin{aligned}
 \Delta(s) &= \sum_{j=1}^9 p_j s^j + R \sum_{j=0}^5 q'_j s^j_{PID}(s) s \\
 &= 5.05s + 5.3s^2 + 2.03s^3 + 0.38s^4 + 0.04s^5 + 0.003s^6 \\
 &\quad + 1.175 \times 10^{-4}s^7 \\
 &\quad + 2.09 \times 10^{-6}s^8 + 6.66 \times 10^{-9}s^9 \\
 &\quad + 3e^{-\tau s} (K_d s^2 + K_p s + K_i) (2 + 1.18s \\
 &\quad + 0.22s^2 + 0.017s^3 + 4.44 \times 10^{-4}s^4 + 4.16 \times 10^{-7}s^5).
 \end{aligned} \tag{17}$$

Substituting $s = j\omega$ and replacing the delay term with $e^{-j\omega\tau} = \cos(\omega\tau) - j\sin(\omega\tau)$, transforms the characteristic equation $\Delta(s)$ into the frequency-domain form given by (7), expressed in terms of angular frequency ω and delay τ , where

$$\begin{aligned}
 A_1 &= (6\omega - 0.7\omega^3 + 10^{-3}\omega^5) \sin(\bar{\omega}) \\
 &\quad - (3.5\omega^2 - 0.05\omega^4 + 1.2 \times 10^{-6}\omega^6) \\
 &\quad \times \cos(\bar{\omega}), \\
 A_2 &= (3.5\omega - 0.05\omega^3 + 1.2 \times 10^{-6}\omega^5) \sin(\bar{\omega}) \\
 &\quad + (6 - 0.7\omega^2 + 10^{-3}\omega^4) \cos(\bar{\omega}), \\
 A_3 &= (0.05\omega^5 - 3.5\omega^3 - 1.2 \times 10^{-6}\omega^7) \sin(\bar{\omega}) \\
 &\quad - (6\omega^2 - 0.7\omega^4 + 10^{-3}\omega^6) \\
 &\quad \times \cos(\bar{\omega}), \\
 A_4 &= -5.3\omega^2 + 0.38\omega^4 - 0.003\omega^6 + 2.09 \times 10^{-6}\omega^8,
 \end{aligned}$$

And

$$\begin{aligned} B_1 &= (6\omega - 0.7\omega^3 + 10^{-3}\omega^5) \cos(\bar{\omega}) \\ &+ (3.5\omega^2 - 0.05\omega^4 + 1.2 \times 10^{-6}\omega^6) \\ &\times \sin(\bar{\omega}), \\ B_2 &= (3.5\omega - 0.05\omega^3 + 1.2 \times 10^{-6}\omega^5) \cos(\bar{\omega}) \\ &- (6 - 0.7\omega^2 + 10^{-3}\omega^4) \sin(\bar{\omega}), \\ B_3 &= (0.05\omega^5 - 3.5\omega^3 - 1.2 \times 10^{-6}\omega^7) \cos(\bar{\omega}) \\ &+ (6\omega^2 - 0.7\omega^4 + 10^{-3}\omega^6) \\ &\times \sin(\bar{\omega}), \\ B_4 &= 5.05\omega - 2.03\omega^3 + 0.04\omega^5 - 1.175 \times 10^{-4}\omega^7 + 6.66 \\ &\times 10^{-9}\omega^9. \end{aligned}$$

With $\bar{\omega} = \omega\tau$. Fixing Kd at a constant value allows computation of the remaining parameters via [Eq. 8](#) using the calculated coefficients A_i and B_i (for $i = 1,2,3,4$). With the stability limits specified in [Eq. 9](#), the resulting stability boundary locus (SBL) curves are plotted in the K_p plane for selected values of Kd and τ , as presented in [Figures 6 and 7](#). Moreover, [Figure 8](#) displays the time-domain frequency deviation trajectories corresponding to stable, critically stable, and unstable regions ([Figure 7\(a\)](#) for $K_d = 0$) across several communication delay values.

Table 1. Parameter values of the AC MG.

Parameter	Value	Parameter	Value
D	0.015 pu/Hz	T_g	0.08 s
$2H$	0.1667 pu s	T_t	0.4 s
TF	0.1 s	TIC	0.004 s
TB	0.1 s	TIN	0.04 s
TFC	0.26 s	R	3 Hz/pu
$TW TG$	1.5 s	TPV	1.8 s

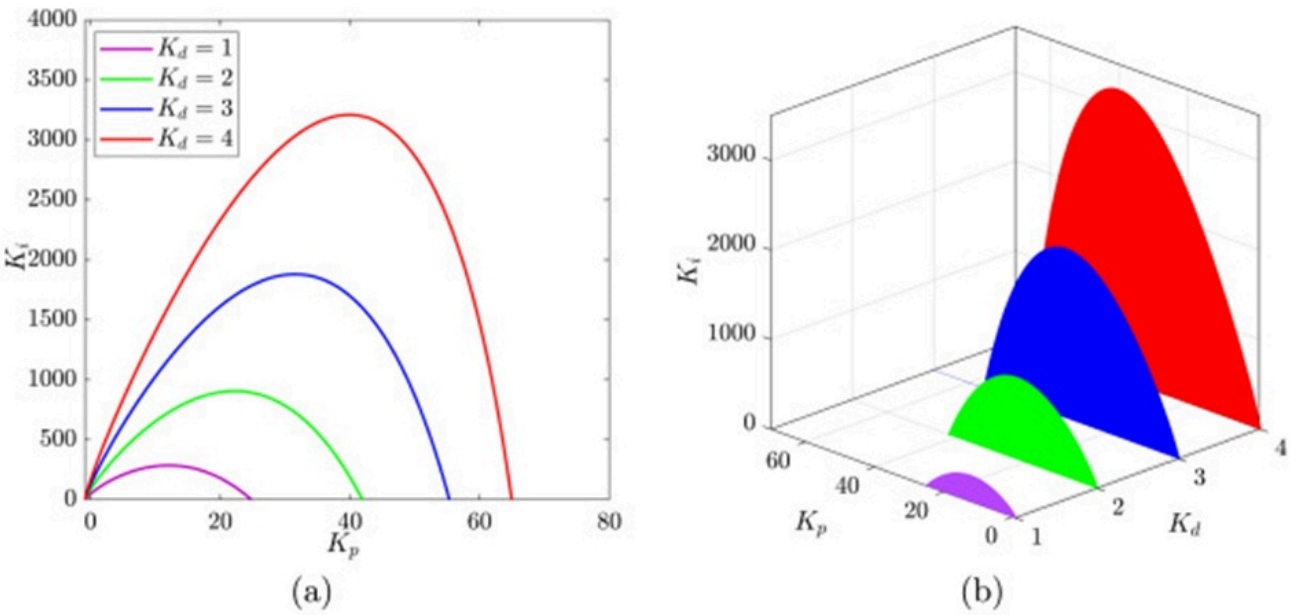


Figure 6. SBL curve of PID controller of the AC MG for $\tau=1$.

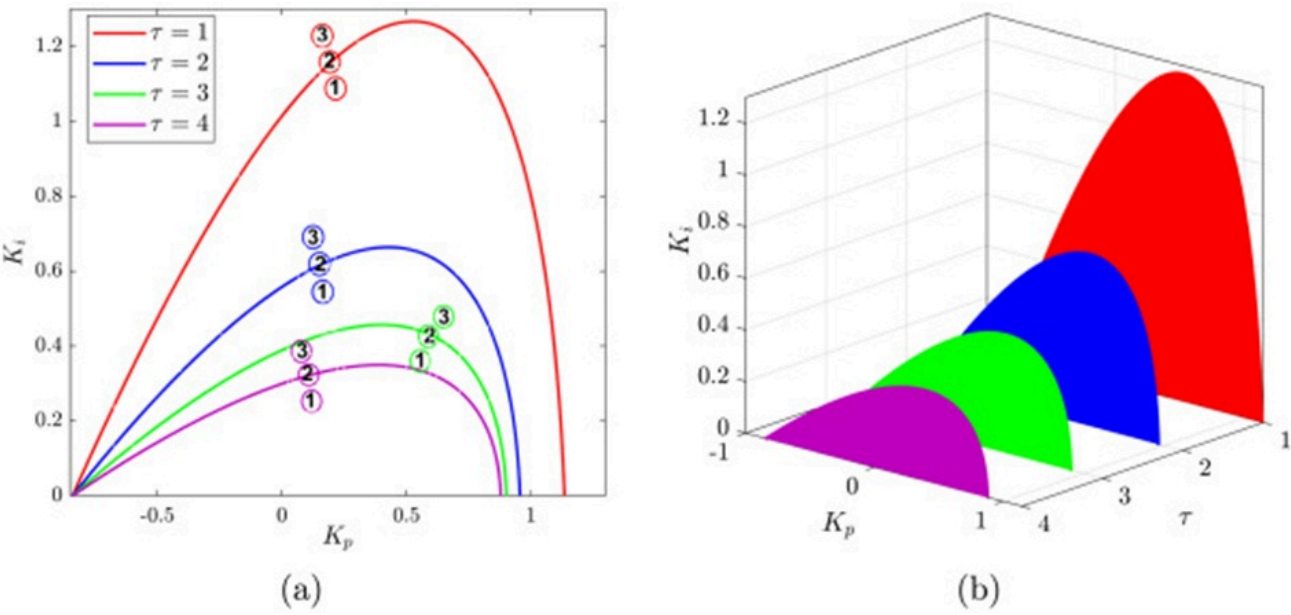


Figure 7. SBL curve of PID controller of the AC MG for $K_d=0$.

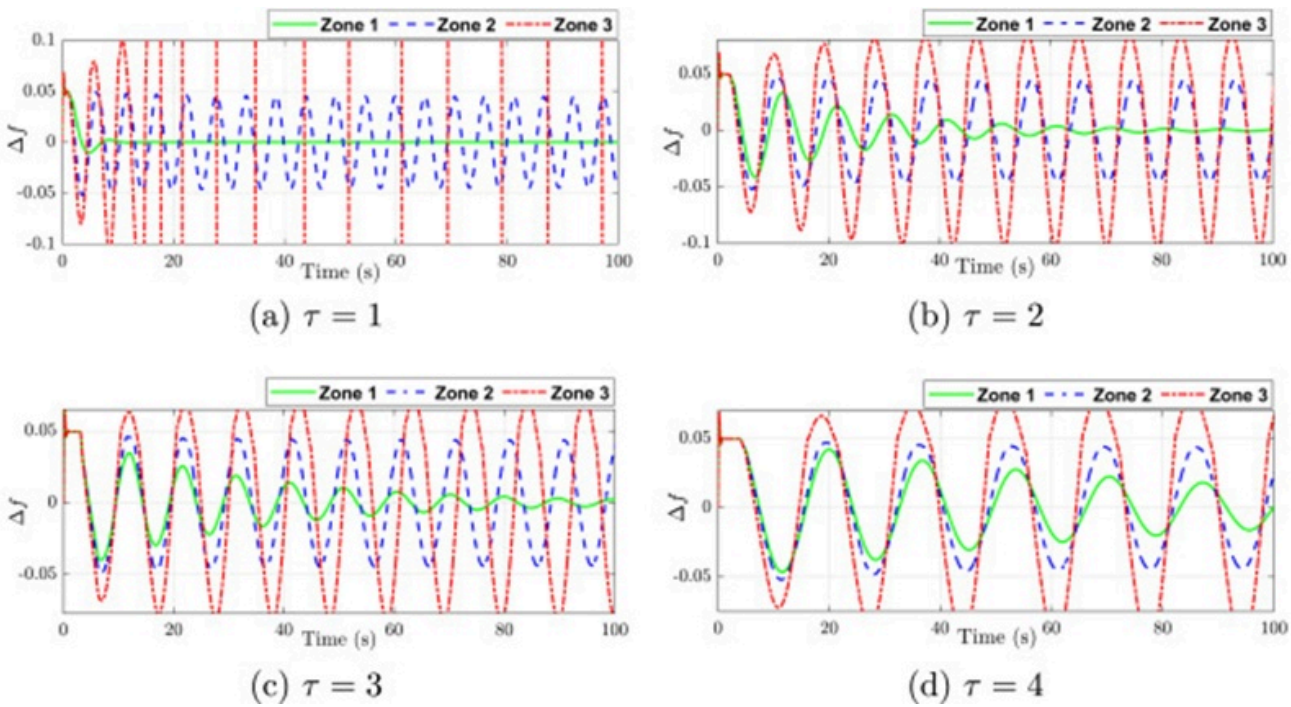


Figure 8. Frequency deviation in different zones of Figure 8(a) for $K_d=0$.

Figure 9 shows the Mikhailov stability plots for PID control with $K_d = 0$ and $K_d = 0$.

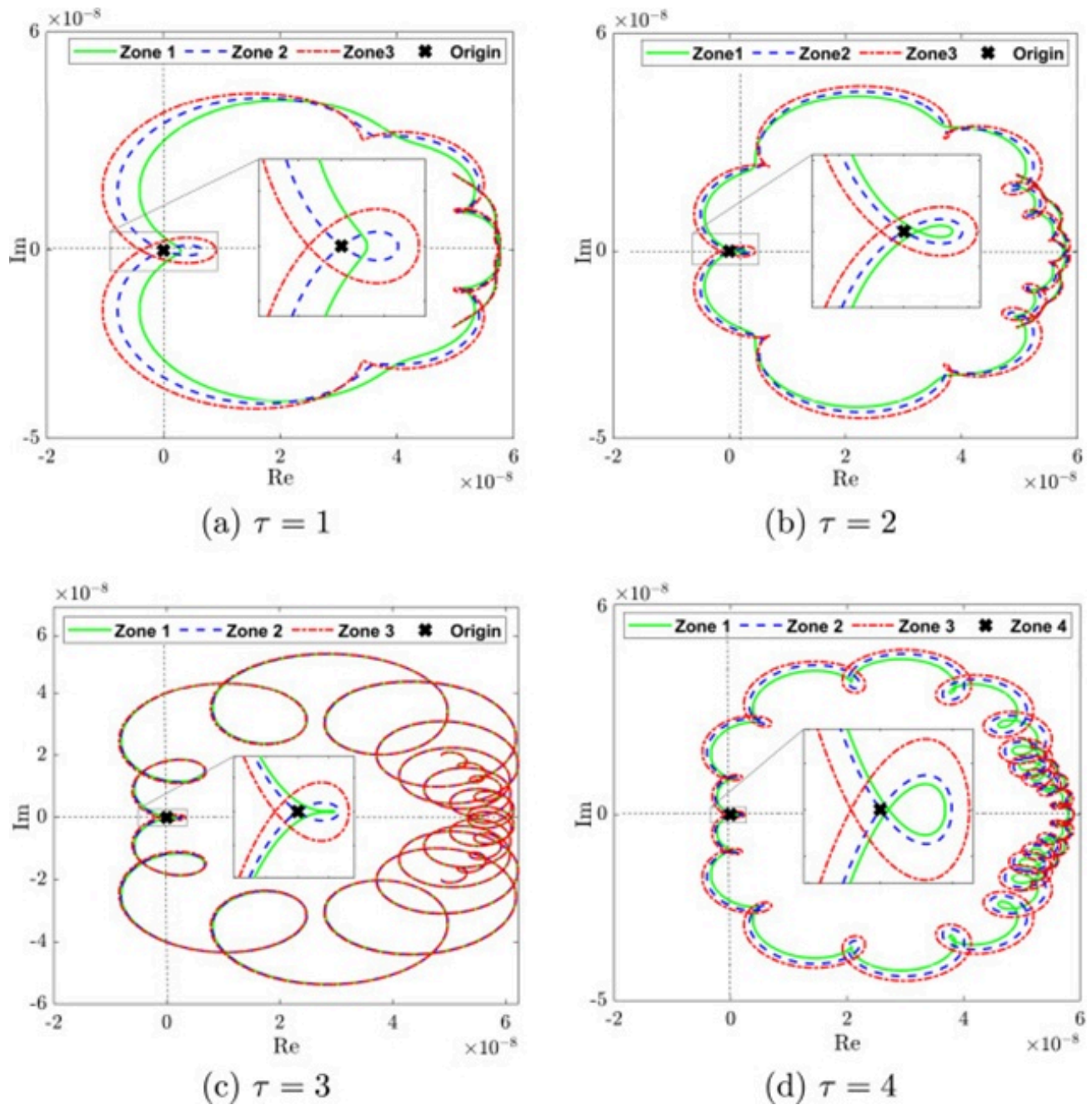


Figure 9. Mikhailov stability plots of PID controller for $K_d=0$.

Controller performance

The intelligent frequency deviation control (FDC) strategy is implemented and tested in the MATLAB/Simulink environment using the microgrid model from [Figure 2](#). The total power mismatch $P_{tot,d}$ —equivalent to the net effect of renewable generation minus load demand—is shown in [Figure 10](#). Abrupt changes take place at $t = 40, 70, 110, 140$, and 180 seconds, combined with persistent random variations across the entire profile.

According to [Eq.10](#), the parameter m_{μ} imposes an upper bound on the output of each neuron in the ANN, thereby constraining the feasible tuning range. This value can either be set manually to create a partially restricted search region [\[45\]](#) or computed intelligently from the interpolated stability boundary function. When chosen manually, an inappropriate

setting may push the PID gains into an unstable zone, resulting in system instability (Figure 11b)). In contrast, Figure 11(a) demonstrates that automatically determining $m\mu$ through the stabilizing reward framework consistently maintains stability. This intelligent method confines adjustments of the proportional and integral gains strictly within the validated stable domain. For comparison, Figures. 11(c) and 11(d) show the outcome of unconstrained tuning, where the computed gains lie outside the stable region and cause unstable frequency behavior.

A performance comparison is provided in Figure 12, Tables 2 and 3, highlighting the fully intelligent, stability-constrained algorithm's ability to effectively solve the FDC task. Evaluation relies on mean absolute error (MAE) and integral absolute error (IAE) indices, which indicate clear outperformance relative to a classical fixed-gain PID controller ($K_p = 3.712$, $K_i = 1.391$, $K_d = 0$), with gains selected via heuristic optimization to provide a fair baseline. The noticeably smaller excursions of frequency deviation Δf from the nominal value confirm the improved behavior of the constrained ANN-PID approach when $K_d = 0$. Although including a small derivative gain yields a modest additional benefit, the improvement is limited and supports the common engineering preference for derivative-free PID structures in noisy microgrid environments to avoid excessive noise amplification.

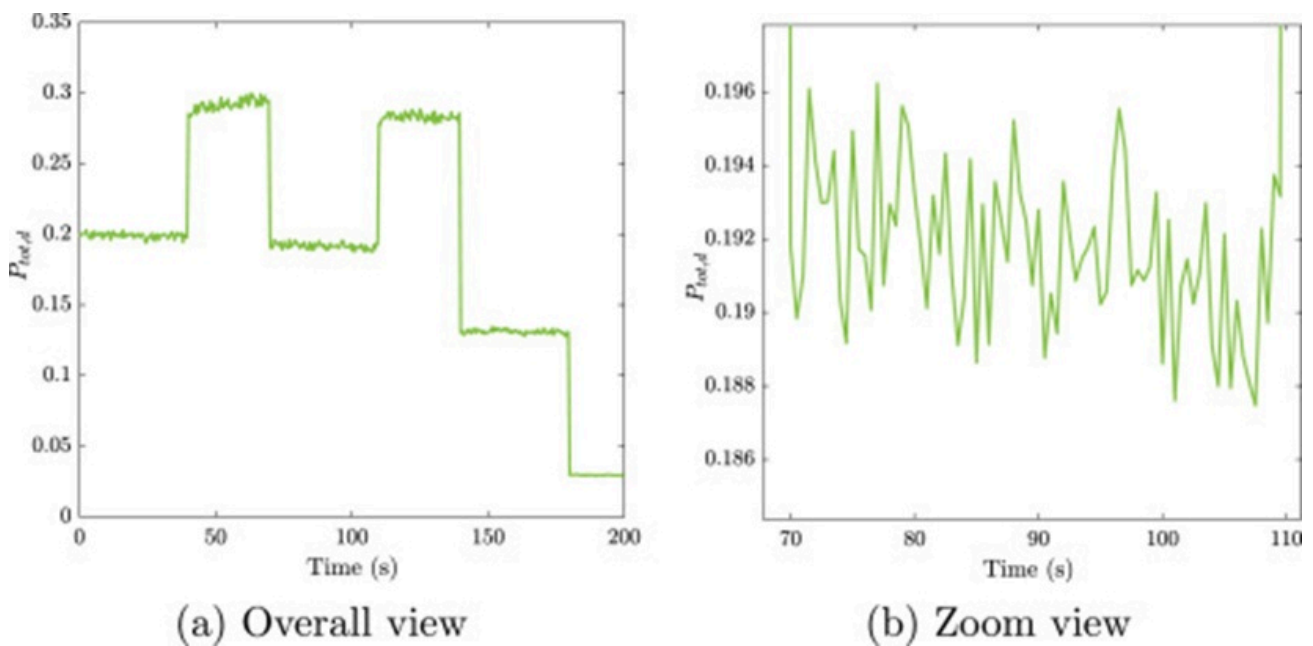


Figure 10. Aggregate power from RESs and load demand (total disturbance input signal).

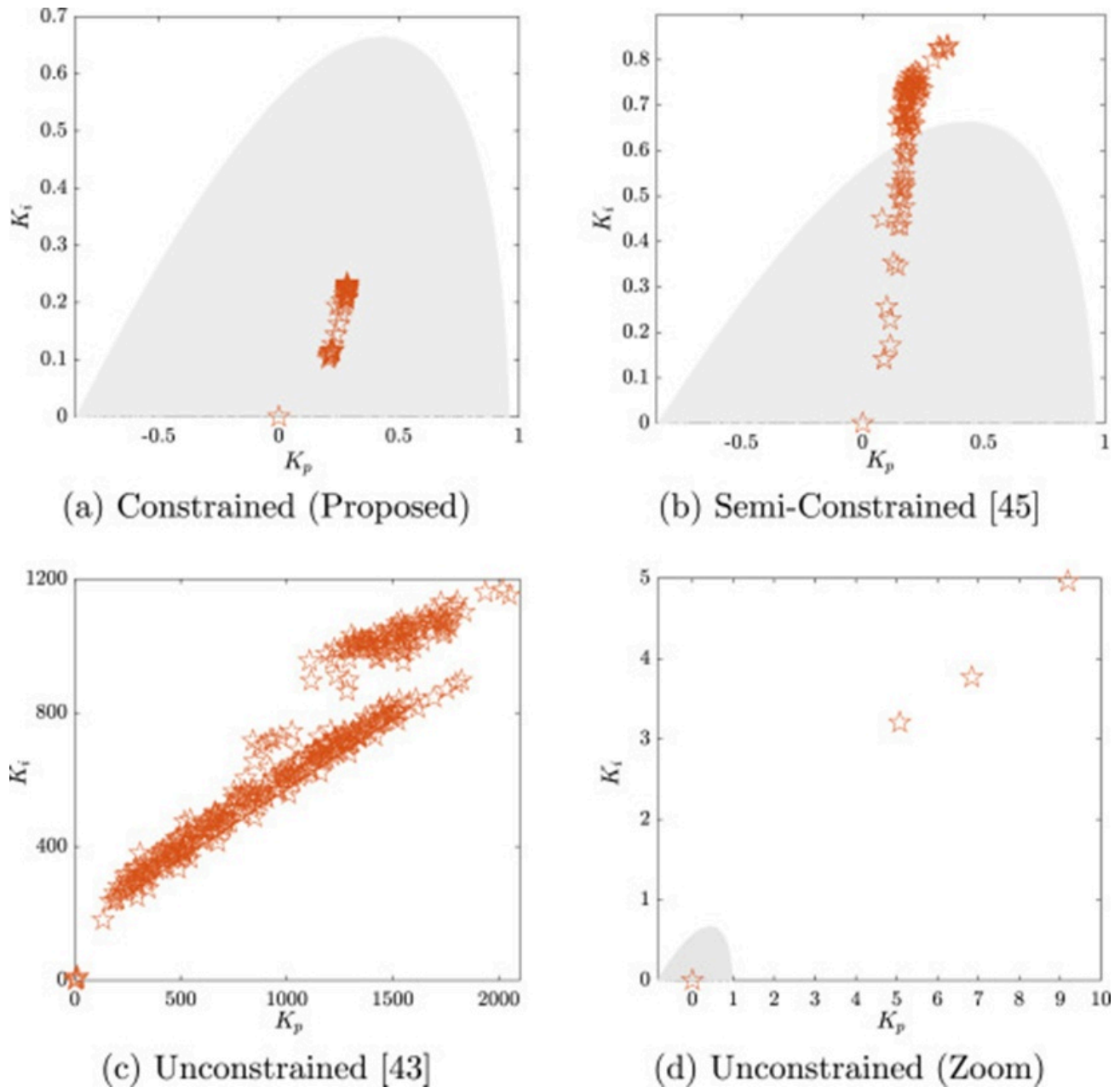


Figure 11. Performance validation of the proposed stability-enforcing ANN-based tuner.

To examine the temporal evolution of the controller gains, **Figure 13** illustrates the online adjustment of K_p and K_i across three distinct apptuning strategies: fully constrained, semi-constrained, and unconstrained. The semi-constrained and unconstrained cases are drawn from prior studies [43, 44], respectively. In **Figure 13(a)**, the outcomes of the carefully engineered fully intelligent ANN tuner are shown; all constraint bounds remain securely within the stable region, thereby guaranteeing stable operation of the AC microgrid under test. This tuner dynamically tailors each gain to match the prevailing operating conditions and consistently outperforms the semi-constrained, unconstrained, and baseline fixed PID controllers across multiple delay conditions (**Tables 2 and 3**).

Figure 13(b) shows a case in which the constraint limits were manually assigned inappropriately, resulting in them lying outside the acceptable stability region. Consequently, the initial phase of the simulation exhibits unstable tuning behavior, and the constrained values quickly saturate at their upper limits, remaining fixed thereafter. This behavior arises because the activation function defined in (**Eq. 10**) imposes a finite output range even for unbounded inputs.

Finally, **Figure 13(c)** depicts an unstable tuning process resulting from the adoption of an unconstrained activation function, namely the purelin function, expressed as

$$y(t) = |x(t) \theta_{\mu}|, \quad (18)$$

where, once again, $x(t)$ and $y(t)$ denote the input and output of the activation function at time t , and θ_{μ} represents the weight matrix of the actor ANN. All simulations discussed here were conducted with a communication delay of $\tau = 2$ s. As shown in **Tables 2 and 3**, the intelligently constrained tuning strategy maintains strong performance even when subjected to delays exceeding this value.

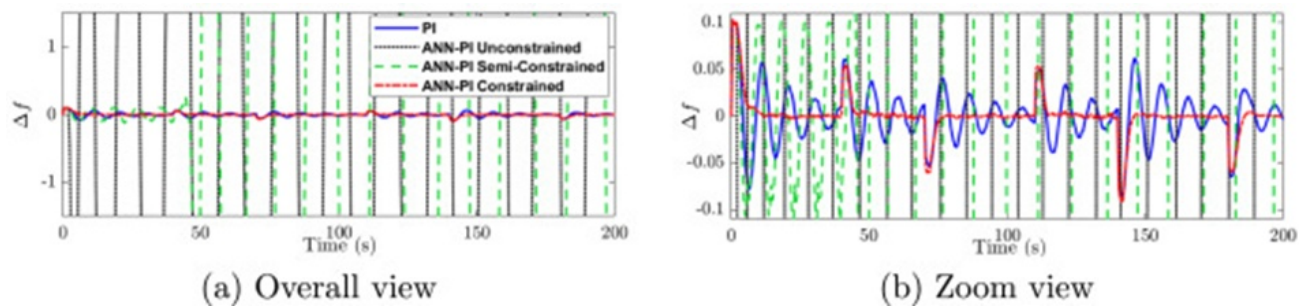


Figure 12. Frequency deviation responses of the AC MG under (semi-)constrained and unconstrained ANN-PID controllers compared with the conventional PID for $\tau = 2$ and $K_d = 0$.

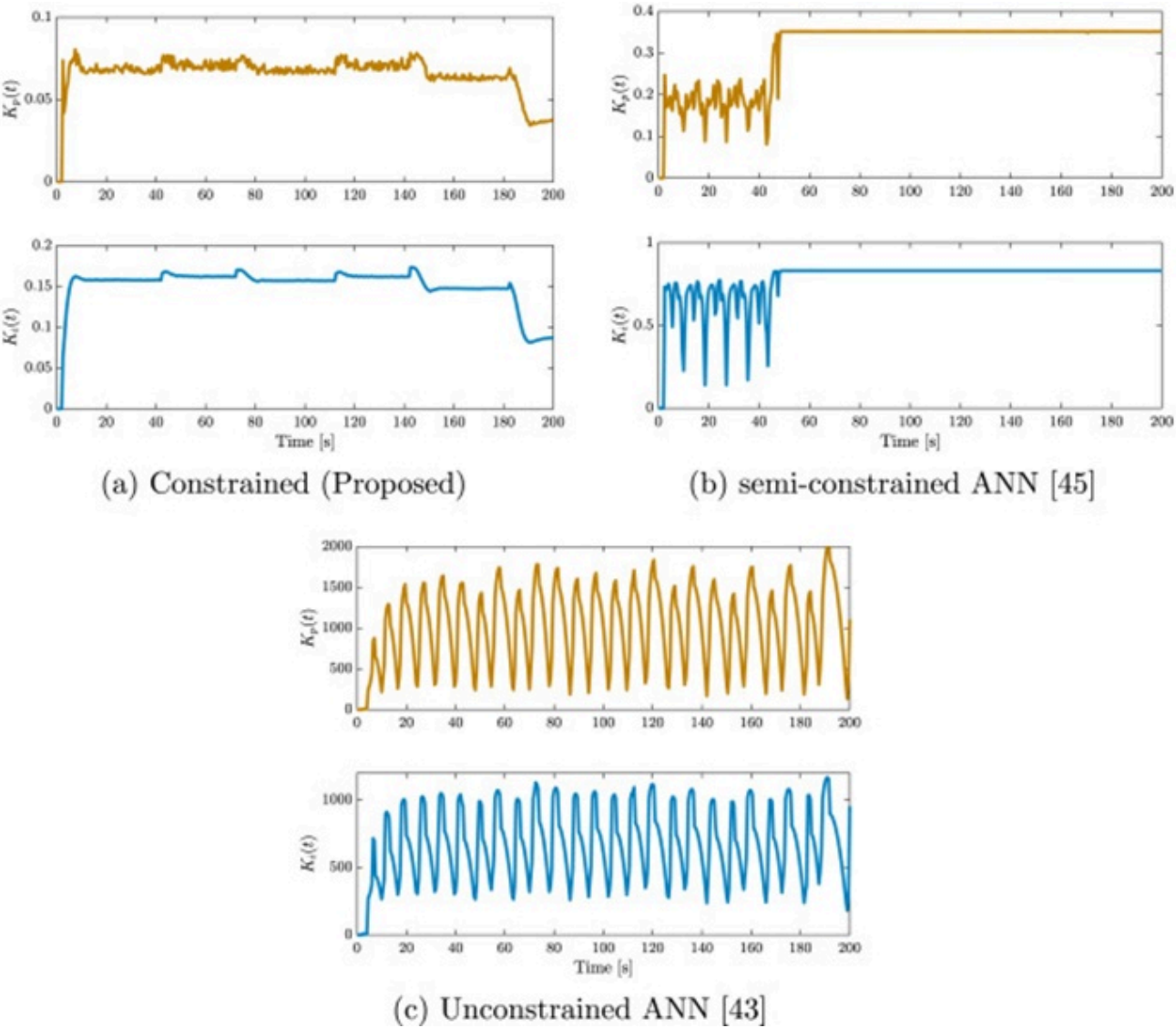


Figure 13. Real-time parameter tuning trajectories for $\tau = 2$.

Table 2. Comparison of different controllers based on the MAE metric.

Delay	Constrained	Ref. [44]	Ref. [43]	PID
$\tau = 2$	0.00078	2.116	7.233	0.021
$\tau = 3$	0.00142	10.452	9.935	11.61
$\tau = 4$	0.00204	11.855	11.981	13.881

Table 3. Comparison of different controllers based on the IAE metric.

Delay	Constrained	Ref. [44]	Ref. [43]	PID
$\tau = 2$	2.4	421	1448	4.354

$\tau = 3$	2.95	2094	1989	2326
$\tau = 4$	4.199	2376	2402	2781

The developed intelligent controller has demonstrated strong effectiveness in frequency deviation control (FDC) by leveraging reinforcement learning to autonomously adapt to disturbances in the AC microgrid, thereby improving stability margins and overall performance. Beyond its superior self-adaptation and resilience against unexpected perturbations, the principal advantages of this controller include the following:

- In contrast to conventional intelligent PID-style controllers that rely on meta-heuristic optimization techniques (Table 4), the proposed real-time tuning approach successfully mitigates the adverse effects of communication delays.
- Relative to semi-constrained [45] and unconstrained [43] techniques, the present method achieves markedly better stability by confining the parameter adjustments to a rigorously defined stable subspace. This restriction not only safeguards system stability but also substantially narrows the search domain, enabling faster, more reliable convergence during tuning.

Although the proposed framework offers a smart, computationally efficient control solution that extends beyond PID controllers and can be adapted to tune other classical regulators, certain limitations remain that warrant future attention. The relatively high computational load may constrain its deployment in real-time scenarios involving very large-scale systems, underscoring the need to improve prediction accuracy under rapidly changing conditions. Further strengthening the robustness of the RL framework—through refined reward design and incorporation of digital protection mechanisms—will be essential. Moreover, extending the approach to accommodate uncertain and time-varying delays while preserving a constrained parameter search grounded in microgrid stability boundaries represents a key area for ongoing research. Addressing these issues will be critical to expanding the practical utility and reliability of this methodology across complex modern power systems.

Table 4. Intelligent PID-type controllers for FDC.

Paper	Method	Controller	Components	Delay	Real-Time
[19]	GA	PID/PI	WTG, DEG, STP, AE, ESS	✗	✗
[20]	PSO	FOPID	PV, WTG, DEG, FC, ESS	✓	✗
[21]	Fuzzy+RL	PID	PV, WTG, DEG, AE, FC, ESS	✗	✗
[3]	Fuzzy+BHA	FOPID	WTG, DEG, (B-F)ESS, EV	✗	✗
[22]	Fuzzy+HSA	PI	WTG, DEG, EV	✗	✗
[23]	BHA	PID	PV, WTG, DEG, (B-F)ESS, EV	✓	✗
[24]	ANN	PI	PV, WTG, DEG, FC, ESS	✗	✗
[25]	ANN+PSO	PID	WTG, DEG, EV	✗	✗
[26]	ANN+RL	PI	PV, WTG, FC, BESS	✗	✓
Proposed	Constr. ANN+RL	PI	PV, WTG, DEG, FC, BESS, FESS	✓	✓

Conclusion

This research addresses the challenge of frequency deviation control in isolated AC microgrids by proposing a constrained artificial neural network (ANN) operating strictly within validated stability regions. The developed framework combines stabilizing PID controllers—derived from small-signal stability analysis—with an automated tuning process that optimizes ANN weights and enforces RL-guided constraints. Compared to semi-constrained [45] and unconstrained [43] approaches, the proposed method achieves substantially lower transient overshoots and steady-state errors. By incorporating reinforcement learning alongside an interpolated stability boundary locus, the technique enables flexible, real-time adaptation that maintains strong performance even under varying operating conditions and in the presence of communication delays.

Despite these advantages, the current implementation assumes a fixed communication delay, which may not fully capture the time-varying nature observed in practical deployments. Furthermore, the adoption of sophisticated RL algorithms increases computational requirements and raises potential safety considerations, posing challenges for seamless deployment in large-scale, high-complexity power networks. Resolving these constraints will be essential to enhancing the method's scalability and real-world viability.

Future investigations will prioritize the design of adaptive mechanisms to handle uncertain or fluctuating delays, improvements in computational performance, and refinements to the RL framework through enhanced reward structures and integration with digital protection schemes. It is expected that the presented intelligent tuning strategy, when paired with emerging advanced control methodologies, will gain widespread acceptance in domains demanding reliable, high-performance regulation solutions.

Nomenclature

AE: Aqua electrolyzer

ANN: Artificial neural network

BESS: Battery ESS

BHA: Black hole algorithm

CRB: Complex root boundary

DDPG: Deep Deterministic Policy Gradient

DEG: Diesel engine generators

DG: Distributed generation

ESS: Energy storage system

EV: Electric vehicle

FC: Fuel cell

FDC: Frequency deviation control

FESS: Flywheel ESS

FOPID: Fractional-order PID

GA: Genetic algorithm

HSA: Harmony search algorithm

IAE: Integral absolute error

MAE: Mean absolute error

MG: Microgrid

PID: Proportional–integral–derivative

PSO: Particle swarm optimization

PV: Photovoltaic

RES: Renewable energy source

RL: Reinforcement learning

SBL: Stability boundary locus

SDDPG: Simplified DDPG

SFC: Secondary frequency control

STP: Solar thermal power

WTG: Wind turbine generator

Variables

α : Learning rate

β : Smooth factor

Δf : Frequency deviation

(s) : Characteristic equation

γ : Discount factor

μ_i : Actor and critic neural networks

ϕ : Error angle

τ : Communication delay

$\theta\mu_i$: Actor and critic ANN weights

φ : Solar irradiation

a t_i : Action and state

(s): Closed-loop transfer function

D_i : Damping and droop coefficients

(t): Interpolated stability function

(s): Transfer function of PID controller

H : System inertia

K_p, d : Proportional, integral, and derivative gains

L_{μ} : Actor and critic loss functions

m_{μ} : Limit output of neuron

$P_{*,*}$: Output power and time constant of unit *

P_{tot}, tot : Total input disturbance and power sources

Q' : Best critic reward prediction

r_s : Stabilizing and performance reward signals

r_t : Reward and total reward signals

y^r : State-action-value function

Acknowledgements

This work was supported by the European Union's Horizon Europe research and innovation program under the grant agreement No [101120657](#); project ENFIELD (European Lighthouse to Manifest Trustworthy and Green AI); by the Estonian Research Council through the grants [PRG658](#) and [PRG1463](#); and by the Estonian Centre of Excellence in Energy Efficiency, ENER (grant [TK230](#)) funded by the Estonian Ministry of Education and Research.

Conflict of interest

None

Financial support

None

Ethics statement

None

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

- Bevrani H, Francois F, Ise T. Microgrid dynamics and control. New York, NY, USA: Wiley; 2017.
- Bevrani H, Habibi F, Babahajyani P, Watanabe M, Mitani Y. Intelligent frequency control in an AC microgrid: Online PSO-based fuzzy tuning approach. *IEEE Trans Smart Grid*. 2012;3(4):1935–44.
<https://doi.org/10.1109/TSG.2012.2196806>.
- Khooban MH, Niknam T, Shasadeghi M, Dragicevic T, Blaabjerg F. Load frequency control in microgrids based on a stochastic noninteger controller. *IEEE Trans Sustain Energy*. 2018;9(2):853–61.
- Chowdhury S, Crossley P. Microgrids and active distribution networks. Stevenage, UK: IET; 2009.
<https://doi.org/10.1049/PBRN006E>.
- Sifat MM, Choudhury SM, Das SK, Ahamed MH, Mueeen SM, Hasan MM, et al. Towards electric digital twin grid: Technology and framework review. *Energy AI*. 2023;11:100213.
<https://doi.org/10.1016/j.egyai.2022.100213>.
- Hooshmand A, Asghari B, Sharma RK. Experimental demonstration of a tiered power management system for economic operation of grid-tied microgrids. *IEEE Trans Sustain Energy*. 2014;5(4):1319–27.
- Jain S, Hote YV. Design of improved nonlinear active disturbance rejection controller for hybrid microgrid with communication delay. *IEEE Trans Sustain Energy*. 2022;13(2):1101–11.
- Tajalli SAM, Tajalli SZ, Homayounzadeh M, Khooban MH. Zero-carbon powerto-hydrogen integrated residential system over a hybrid cloud framework. *IEEE Trans Cloud Comput*. 2023;11(3):3099–110.
- Hasanvand S, Rafiei M, Gheisarnejad M, Khooban MH. Reliable power scheduling of an emission-free ship: Multi-objective deep reinforcement learning. *IEEE Trans Transp Electr*. 2020;6(2):832–43.
- Shi H, Wang S, Fernandez C, Yu C, Xu W, Dablu BE, et al. Improved multi-time scale lumped thermoelectric coupling modeling and parameter dispersion evaluation of lithium-ion batteries. *Appl Energy*. 2022;324:119789.
- Li K, Gao X, Liu C, Chang C, Li X. A novel Co-estimation framework of state-of-charge, state-of-power and capacity for lithium-ion batteries using multiparameters fusion method. *Energy*. 2023;269:126820.
- Qi C, Wang S, Cao W, Wang Y, Liu D, Fernandez C. Improved joint prediction strategy for state of charge and peak power of lithium-ion batteries by considering hysteresis characteristics-current measurement deviation correction. *J Energy Storage*. 2024;84:110726.
- Pan I, Das S. Kriging based surrogate modeling for fractional order control of microgrids. *IEEE Trans Smart Grid*. 2015;6(1):36–44.
- Khayat Y, Shafiee Q, Heydari R, Naderi M, Dragičević T, Simpson-Porco JW, et al. On the secondary control architectures of AC microgrids: An overview. *IEEE Trans Ind Electron*. 2019;35(6):6482–500.
<https://doi.org/10.1109/TPEL.2019.2951694>.

Bevrani H, Hiyama T. Intelligent automatic generation control. New York: CRC; 2011; 9781315217406.
<https://doi.org/10.1201/b10869>.

Liu Z, Sun Y, Xing C, Liu J, He Y, Zhou Y, et al. Artificial intelligence powered large-scale renewable integrations in multi-energy systems for carbon neutrality transition: Challenges and future perspectives. *Energy AI*. 2022;10:100195.
<https://doi.org/10.1016/j.egyai.2022.100195>.

Liu Y, Zhou P, Yang L, Wu Y, Xu Z, Liu K, et al. Privacy-preserving contextbased electric vehicle dispatching for energy scheduling in microgrids: An online learning approach. *IEEE Trans Emerg Top Comput Intell*. 2021;6(3):462–78.
<https://doi.org/10.1109/TETCI.2021.3085964>.

Zhang C, Qiu Y, Chen J, Li Y, Liu Z, Liu Y, et al. A comprehensive review of electrochemical hybrid power supply systems and intelligent energy managements for unmanned aerial vehicles in public services. *Energy AI*. 2022;9:100175.
<https://doi.org/10.1016/j.egyai.2022.100175>.

Das DC, Roy AK, Sinha N. GA based frequency controller for solar thermal– diesel–wind hybrid energy generation/energy storage system. *Int J Electr Power Energy Syst*. 2012;43(1):262–79.

Pan I, Das S. Fractional order AGC for distributed energy resources using robust optimization. *IEEE Trans Smart Grid*. 2015;7(5):2175–86.

Esmaili M, Shayeghi H, Nejad HM, Younesi A. Reinforcement learning based PID controller design for LFC in a microgrid. *Int J Comput Math Electr Electron Eng*. 2017;36(4):1287–97.
<https://doi.org/10.1108/COMPEL-09-2016-0408>.

Khooban MH, Niknam T, Blaabjerg F, Dragičević T. A new load frequency control strategy for micro-grids with considering electrical vehicles. *Electr Power Syst Res*. 2017;143:585–98.
<https://doi.org/10.1016/j.epsr.2016.10.057>.

Khooban MH. Secondary load frequency control of time-delay standalone microgrids with electric vehicles. *IEEE Trans Ind Electron*. 2017;65(9):7416–22.
<https://doi.org/10.1109/TIE.2017.2784385>.

Bevrani H, Habibi F, Shokooi S. ANN-based self-tuning frequency control design for an isolated microgrid. In: *Meta-heuristics optimization algorithms in engineering business economics and finance*. IGI Global; 2013. p. 357–85.

Akula SK, Salehfar H. Frequency control in microgrid communities using neural networks. In: *Proceedings of the North American Power Symposium IEEE*; 2019. p. 1–6.
<https://doi.org/10.1109/NAPS46351.2019.9000219>.

Nosrati K, Tepļakov A, Petlenkov E, Levron Y, Skiparev V, Belikov J. Coordinated PI-based frequency deviation control of isolated hybrid microgrid: An online multi-agent tuning approach via reinforcement learning. In: *Proceedings of the IEEE PES Innovative Smart Grid Technologies*. 2022. p. 1–5.

Zeng GQ, Xie XQ, Chen MR, Weng J. Adaptive population extremal optimizationbased PID neural network for multivariable nonlinear control systems. *Warm Evol Comput*. 2019;44:320–34.

Lu KD, Zeng GQ, Zhou W. Adaptive constrained population extremal optimisation-based robust proportional-integral-derivation frequency control method for an islanded microgrid. *IET Cyber-Syst Robot*. 2021;3(3):210–27.

Coelho EA, Wu D, Guerrero JM, Vasquez JC, Dragicevic T, Stefanović C, et al. Small-signal analysis of the microgrid secondary control considering a communication time delay. *IEEE Trans Ind Electron*. 2016;63(10):6257–69.

Vafamand N, Khooban MH, Dragičević T, Blaabjerg F. Networked fuzzy predictive control of power buffers for dynamic stabilization of DC microgrids. *IEEE Trans Ind Electron*. 2018;66(2):1356–62.

Milano F, Anghel M. Impact of time delays on power system stability. *IEEE Trans Circuits Syst*. 2012;59(4):889–900.

Dong M, Li L, Nie Y, Song D, Yang J. Stability analysis of a novel distributed secondary control considering communication delay in DC microgrids. *IEEE Trans Smart Grid*. 2019;10(6):6690–700.

Lv Z, Wu Z, Dou X, Hu M. Discrete consensus-based distributed secondary control scheme with considering time-delays for DC microgrid. In: *Proceedings of the 41st annual conference of the IEEE Industrial Electronics Society*. 2016. p. 2898–903.

Zhang CK, Jiang L, Wu QH, He Y, Wu M. Delay-dependent robust load frequency control for time delay power systems. *IEEE Trans Power Syst*. 2013;28(3):2192–201.

Trivedi R, Khadem S. Implementation of artificial intelligence techniques in microgrid control environment: current progress and future scopes. *Energy AI*. 2022;8:100147.
<https://doi.org/10.1016/j.egyai.2022.100147>.

Espín-Sarzosa D, Palma-Behnke R, Cañizares CA, Annakkage U, Elizondo M, Espina E, et al. Microgrid modeling for stability analysis. *IEEE Trans Smart Grid*. 2024;15(3):2459–79.

San G, Zhang W, Guo X, Hua C, Xin H, Blaabjerg F. Large disturbance stability for power-converter-dominated microgrid: A review. *Renew Sustain Energy Rev*. 2020;127:109859.
<https://doi.org/10.1016/j.rser.2020.109859>.

Kabalan M, Singh P, Niebur D. Large signal Lyapunov-based stability studies in microgrids: A review. *IEEE Trans Smart Grid*. 2016;8(5):2287–95.

Kundur P, Balu NJ, Lauby MG. *Power system stability and control*. New York, NY, USA: McGraw-Hill; 1994.

Yildirim B, Gheisarnejad M, Khooban MH. A new parameter tuning technique for noninteger controllers in low-inertia modern power grids. *IEEE J Emerg Sel Top Power Electron*. 2021;3(2):279–88.

Ozdemir MT. The effects of the FOPI controller and time delay on stability region of the fuel cell microgrid. *Int J Hydrog Energy*. 2020;45(60):35064–72.
<https://doi.org/10.1016/j.ijhydene.2020.05.211>.

Yildirim B, Gheisarnejad M, Khooban MH. A robust non-integer controller design for load frequency control in modern marine power grids. *IEEE Trans Emerg Top Comput Intell*. 2021;6(4):852–66.

Skiparev V, Nosrati K, Tepļakov A, Petlenkov E, Levron Y, Belikov J, et al. Virtual inertia control of isolated microgrids using NN-based VFOPID controller. *IEEE Trans Sustain Energy*. 2023;14(3):1558–68.

Zhang S, Liu H, Liu F, Ma H, Shi Y, Gao P, et al. Research on excitation current control system of the 50 kA superconducting transformer. *IEEE Trans Appl Supercond*. 2021;31(8):1–4.

Nosrati K, Tepļakov A, Petlenkov E, Skiparev V, Belikov J, Levron Y. Constrained intelligent frequency control in an AC microgrid: An online reinforcement learning based PID tuning approach. In: *Proceedings of the IEEE power and energy society general meeting*. 2023. p. 1–5.

Obara S. Analysis of a fuel cell micro-grid with a small-scale wind turbine generator. *Int J Hydrog Energy*. 2007;32(3):323–36.

Wang S, Gao H, Takyi-Aninakwa P, Guerrero J, Fernandez C, Huang Q. Improved multiple feature-electrochemical thermal coupling modeling of lithium-ion batteries at low-temperature with real-time coefficient correction. *Prot Control Mod Power Syst*.

2024;9(3):157–73.

<https://doi.org/10.23919/PCMP.2023.000257>.

Wang S, Takyi-Aninakwa P, Jin S, Yu C, Fernandez C, Daniel-loan S. An improved feedforward-long short-term memory modeling method for the whole-life-cycle state of charge prediction of lithium-ion batteries considering current–voltage temperature variation. *Energy*. 2022;254:124224.

Skiparev V, Nosrati K, Belikov J, Teplyakov A, Petlenkov E. An enhanced NNbased load frequency control design of MGs: A fractional order modeling method. In: *Proceedings of the IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering*. 2023. p. 1–6.

Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, Tassa Y, et al. Continuous control with deep reinforcement learning. 2015, , arXiv preprint arXiv:1509.02971.

<https://doi.org/10.48550/arXiv.1509.02971>.

Khooban MH, Gheisarnejad M. A novel deep reinforcement learning controller based type-II fuzzy system: Frequency regulation in microgrids. *IEEE Trans Emerg Top Comput Intell*. 2020;5(4):689–99.

<https://doi.org/10.1109/TETCI.2020.2964886>.