

Simultaneous Frequency Regulation and Active Power Sharing in Islanded Microgrid Using Deep Reinforcement Learning

Oroghene Oboreh-Snapps, Sophia A. Strathman, Jonathan Saelens, Arnold Fernandes,
Lauryn Morris, Praneeth Uddaraju, and Jonathan W. Kimball,
Department of Electrical and Computer Engineering
Missouri University of Science and Technology, Rolla, Missouri, USA
{oogdq, ss6k4, jhskrw, arnoldanthony.fernandes, lrmdhf, pu5mb, kimballjw}@mst.edu

Abstract—This paper presents a novel approach that integrates deep reinforcement learning (DRL) with the conventional virtual synchronous generator (VSG) to address dual objectives of microgrid (MG) control; frequency regulation and precise active power sharing. MGs typically consist of multiple Inverter-Based-Distributed-Generators (IBDGs) connected in parallel through different line impedances. The conventional active power loop (APL) of the VSG encounters significant steady-state frequency errors as load increases/decreases during islanded operation. To mitigate this issue, secondary-level controllers like proportional-integral (PI) control are added to the APL to regulate the frequency of IBDGs. However, PI control compromises power-sharing capabilities when the impedance values of connecting feeders for each IBDG are mismatched. To eliminate frequency errors and achieve accurate power sharing concurrently, this study adopts a DRL-based strategy. The agent collects state information from each IBDG in the microgrid as input and undergoes training using a reward function crafted to satisfy both objectives simultaneously. The performance of the trained agent is demonstrated in a two-inverter microgrid system designed in MATLAB/SIMULINK and is compared against traditional methods.

Index Terms—Active Power Sharing, Deep Reinforcement Learning, Frequency Control, Inverter Based Distributed Generators, Microgrids, Twin Delayed DDPG, Virtual Synchronous Generator.

I. INTRODUCTION

Microgrids are capable of operating independently from the main grid, through the integration of renewable energy resources (RESs) that are controlled via grid-forming inverter schemes such as or virtual synchronous generator (VSG) based controls. However, parallel connection of inverter-based distributed resources (IBDGs) with mismatched feeder impedances often leads to inaccurate active and/or reactive power-sharing, challenging stable microgrid operation [1].

To achieve precise power sharing capabilities in a parallel inverter MG, the classical droop control for both active and reactive power control was presented in [2]. An enhanced version of the control is presented in [3], wherein an adaptive decentralized controller is developed for paralleled inverter-based microgrids to address power-sharing errors and enhance the damping characteristics of the active power loop (APL).

However, it is widely recognized that conventional VSG-based control suffers from significant steady-state frequency errors and thus necessitates augmentation with a secondary-level controller. In [4], a solution is proposed to simultaneously restore frequency and maintain accurate APS. While a proportional-integral (PI) controller can eliminate frequency errors, it often compromises the APS capability of inverter-based distributed generators (IBDGs). To address this issue, the authors introduced a compensation integral controller, augmented with both frequency recovery and control. A different approach was taken into consideration by [5], wherein the control for frequency restoration and minimizing the active power sharing error (APSE) involved the use of a nonlinear feedback controller for each IBDG phase angle.

While the strategies mentioned effectively address frequency regulation and accurate APS, they heavily rely on precise mathematical models of the MG system, which are often rigid, highly non-linear, and complex to develop [6]–[8]. In recent years, the push for model-free methods for designing grid-forming inverter controllers has seen rapid growth among researchers. Most notably, the adoption of model-free methods such as fuzzy logic-based controllers or deep reinforcement learning (DRL) based approaches implies that an intelligent control scheme for IBDG-dominated MGs can be designed without a mathematical system model, or complex estimation algorithms [9], [10]. DRL methods rely on receiving necessary state information, taking appropriate control actions, a receiving reward that evaluates the agents' performance, and transitioning to a new state [11]. Through numerous iterative episodes, the agent's goal is to learn a policy that maximizes its expected reward over time. Notably, DRL-based solutions have been successful in MG-related control tasks such as; the adaptive parameter tuning of VSG control [12]–[14], frequency and voltage stabilization [15], [16] and intelligent energy management systems for MGs [17], [18]. Also, in our most recent work, a DRL-based solution for achieving accurate reactive power sharing and voltage regulation in parallel inverter MGs was discussed in [19].

To this end, the major contribution of this paper involves the fusion of a DRL-based algorithm with the APL of a VSG to simultaneously eliminate the frequency of steady-state errors and minimize the APSE in a model-free manner. To the best

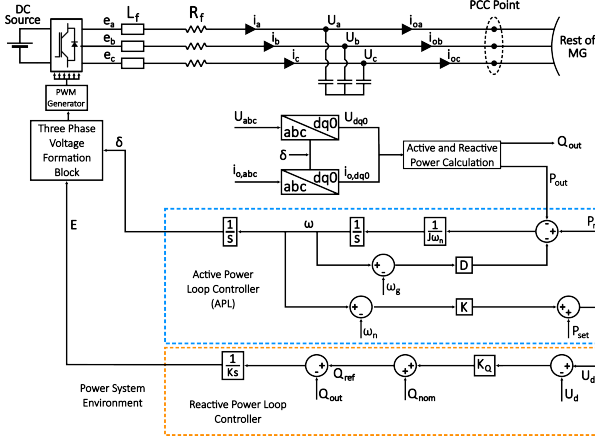


Fig. 1. Architecture of a Conventional Virtual Synchronous Generator Controlled Inverter

of the author's knowledge, this is the first time this strategy is being adopted for addressing this problem in MGs. The rest of this paper is summarized as follows: section II discusses the conventional VSG control and highlights the challenges it suffers from as regards accurate active power sharing and frequency regulation. Section III presents the proposed TD3 based control strategy and the reward formulation. Section IV evaluates and compares the proposed controller with the classical methods and section V concludes this paper.

II. SYSTEM DESCRIPTION

The conventional VSG control design for grid-tied inverters is shown in Fig. 1. The VSG three-phase output current (I_{oabc}) and voltage (V_{oabc}) are sensed, converted to the $dq0$ reference frame, and used in computing the output active (P_{out}) and reactive (Q_{out}) power of the IBDG. The APL aims to emulate the synchronous generator (SG) swing equation, enabling the inverter to provide virtual inertia for enhanced frequency response.

$$P_{ref} - K_p(\omega - \omega_g) - D(\omega - \omega_g) - P_{out} = J\omega\dot{\omega} \quad (1)$$

Where P_{ref} , J , D , and K_p are the reference active power of the VSG, the virtual inertia, virtual damping factor, and active power droop coefficient respectively. ω and ω_g represent the speed of the virtual rotor and the reference angular speed.

The control output of the APL is the inverter power angle (δ) which can be computed as:

$$\delta = \int (\omega - \omega_g) dt \quad (2)$$

For reactive power control, the RPL consists of both a droop control loop and an integral control loop. The droop control generates the reference reactive power (Q_{ref}) as shown in (3).

$$Q_{ref} = k_q(v_{nom} - v_d) + Q_{nom} \quad (3)$$

In eq. (3), Q_{nom} , V_{nom} , v_d , and K_q are the nominal reactive power, nominal voltage, output voltage, and voltage droop gain respectively.

$$E = \frac{1}{K_i} \int \Delta Q dt \quad (4)$$

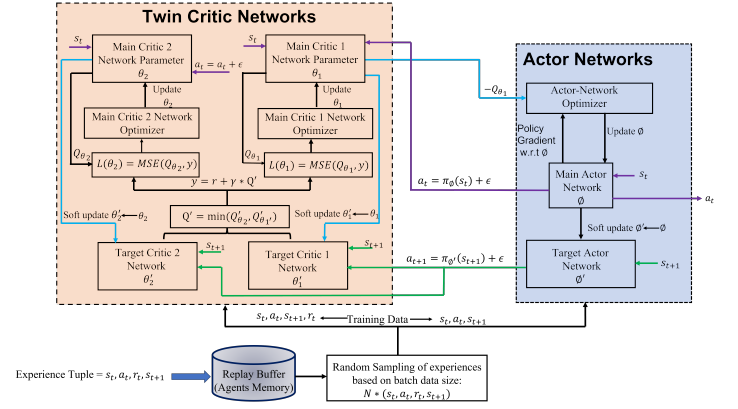


Fig. 2. TD3 Agent Architecture [12]

The output of the droop control loop is fed through an integral controller as shown in eq. (4) to generate inverter voltage reference (E). For clarity, the terms K_i , and ΔQ represent the reactive power integral gain, and the reactive power error ($\Delta Q = |Q_{nom} - Q_{out}|$).

Since this paper focuses on the VSG-APL, further discussions on the RPL are not provided. In reality, MGs consist of multiple VSGs in parallel, and when operated in islanded mode, they form the MGs' voltage and frequency. The conventional VSG active power loop achieves precise sharing in parallel-connected VSGs but exhibits steady-state frequency error during load fluctuations, which could be detrimental to the MG stability. To mitigate this, secondary-level control, often in the form of PI control, is introduced for frequency restoration. However, this control results in inaccurate active power sharing particularly when the feeder impedance connecting each IBDG is not similar. While [4] and [5] proposed viable solutions, the methods presented are highly reliant on an accurate mathematical model of the MG and estimation techniques. However, achieving an accurate mathematical system model often requires linearizing a highly nonlinear system model, which leads to a loss in hidden dynamics and is often not easily adaptable to system configuration changes.

III. PROPOSED CONTROL STRATEGY

To simultaneously regulate each VSG frequency while achieving accurate active power sharing in a model-free manner, this research proposes the fusion of a DRL agent trained with a reward function that is designed to satisfy both objectives. Details regarding the design and implementation of the DRL method are presented below.

A. TD3 Algorithm

The twin delayed deep deterministic policy gradient (TD3) algorithm is a DRL algorithm tailored for continuous control tasks. It utilizes key features such as delayed actor-network updates, twin critic-networks for overcoming over-estimation bias, and target policy smoothing regularization for achieving stable learning [20].

As shown in Fig. 2, TD3 utilizes six neural networks consisting of an actor-network (parameterized by ϕ) for action

selection and a corresponding target actor-network (parameterized by ϕ'). Additionally, two twin critic-networks (parameterized by θ_1 and θ_2) estimate Q-values, complemented by target twin critic-networks (parameterized by θ'_1 and θ'_2) for training stability.

Parameters are randomly initialized at the start of training, and a finite replay buffer stores and replays past experiences. The actor-network aims at learning a policy $\pi(s_t|a_t)$ which maximizes the expected reward when the agent takes action (a_t) in states (s_t). Twin critic-networks evaluate the action value function $Q_i(s_t, a_t|\theta_i)$, providing crucial feedback to guide the actor-network learning, enhancing decision-making and policy improvement [12], [19].

The target networks are frozen duplicates of the primary networks, serving as continuous reference points for training stability. In DRL, achieving convergence involves multiple gradient updates applied to both the actor and critic-network weights and target networks mitigate shifting target values by providing consistent reference points. This stability enhances learning effectiveness, allowing the algorithm to explore a broader range of actions. The output of the target-critic networks and the actual critic-network are compared and used to compute the critic-network loss, which is required for updating both critic-network weights.

On the other hand, the actor-network maps state inputs to an estimated optimal policy. Its input layer matches the environment's state space dimensions. Two hidden layers follow, leading to an output layer with dimensions matching the action space. The output layer utilizes hyperbolic sigmoid activation, scaled to ensure predicted actions fall within the desired range.

B. State and Action

As stated earlier, DRL agents learn from interacting with an environment by receiving state information, taking actions, and obtaining a reward. In this work, the states (s_t) and actions (a_t) are given as;

$$s_t = [F_i, F_j, P_{out_i}, P_{out_j}] \quad (5)$$

$$a_t = [P_{ref_i}, P_{ref_j}] \quad (6)$$

Where, F_i, F_j, P_{out_i} and P_{out_j} are the frequencies and output active power of the i_{th} and j_{th} IBDGs in the MG respectively. The predicted actions a_t based on the current state s_t are applied to the environment causing the agent to transition to a new state s_{t+1} . The consequence of taking action a_t in state s_t is a reward r_t . This sequence of events represented as s_t, a_t, r_t, s_{t+1} , forms a transition tuple that is saved in a buffer B . The experiences stored in B are randomly sampled in mini-batches and used for training the networks. As the buffer has a finite capacity, older experiences are removed to make room for newer experiences when it becomes full. This mechanism ensures that the buffer retains recent experiences, facilitating convergence during training.

C. Reward Function Design

The goal of any DRL agent is to find the optimal policy that maximizes the expected cumulative reward. Therefore, a good reward function must capture the problem description to

properly guide the agent learning. In this paper, the reward function is split into three parts:

(1) Frequency regulation: Inspired by the understanding that the traditional VSG control active power loop exhibits characteristics similar to conventional droop control, which is known for its considerable steady-state frequency error, the subsequent reward has been formulated for the DRL agent.

$$\text{Reward}_F = -k_1|e_f| - k_2 \int |e_f| dt \quad (7)$$

In eq. (7), $|e_f| = |F_i - F_{ref}|$, where e_f, F_i , and F_{ref} represent the absolute error of the frequency measurement, the frequency of the i_{th} IBDG, and the reference frequency, respectively. Terms k_1 and k_2 act as penalty factors, with k_1 penalizing the agent for deviations beyond the frequency error threshold (0.001) and k_2 providing a fixed negative penalty to encourage quick recovery toward the nominal value.

(2) Minimize APS Error: As discussed previously, when the conventional PI controller is used for achieving frequency regulation, the active power sharing is ruined. To address this issue, the following reward function is formulated.

$$\text{Reward}_P = -k_3|e_p| - k_4 \int (|e_p|) dt \quad (8)$$

In eq. (8), $|e_p|$ denotes the error between the expected output power P_{exp} and the actual output power P_{out} for each IBDG. The penalty terms are represented by k_3 and k_4 . Specifically, $|e_p| = P_{exp} - P_{out}$, where P_{exp} is defined as.

$$P_{exp} = P_{pcc} \cdot \frac{S_{rating}^{IBDG_i}}{S_{MGcapacity}} \quad (9)$$

Where P_{pcc} , $S_{rating}^{IBDG_i}$, and $S_{MGcapacity}$ represent the total active power demand at the point of common coupling (PCC), the capacity rating of the considered IBDG, and the total capacity rating of all IBDGs in the microgrid. The penalty gain k_3 imposes a substantial negative reward when the term e_p exceeds the error threshold; otherwise, it applies a small negative penalty, encouraging the agent to explore policies minimizing active power-sharing errors. The term k_4 remains constant, motivating the agent to swiftly reduce the error between the measured and expected active power.

(3) IBDG Capacity Ratio Constraint: To guarantee that IBDGs make equitable contributions relative to their ratings in injecting active power, regardless of line impedance variations, the following reward term has been introduced

$$\text{Reward}_{IBDGRatio} = -k_5 \text{IBDGRatio} \quad (10)$$

where the term IBDGRatio is expressed as:

$$\text{IBDGRatio} = \frac{P_{IBDG_{i=1}}}{\sum_{j=2}^n P_{IBDG}} \quad (11)$$

The term k_5 is the penalty associated with the IBDGRatio . Therefore, if the reactive power contribution of any IBDG in the network is below or above the expected contribution boundary, then k_5 is a large negative reward; otherwise, k_5 is a small negative reward.

TABLE I
NETWORK AND SYSTEM PARAMETERS

TD3 Network Parameters		System Parameters and Reward Penalties		System Parameter	
Network Parameter	Value	System Parameter	Value	System Parameter	Value
Actor Learning Rate	1×10^{-4}	Microgrid Capacity	6000 KVA	Big Penalty [k_1, k_2, k_4]	[5000, 5000, 500]
Twin Critics Learning Rate	2×10^{-4}	Inverter Nominal Voltage	13.8 kV	Small Penalty [k_1, k_2, k_4]	[0.05, 0.05, 0.05]
Target Learning Rate	5×10^{-3}	Filter Resistance	1.9 m Ω	IBDG 1 Line Impedance	$R_{line} = 6m\Omega$ $L_{line} = 40.35mH$
Buffer Length	2×10^6	Filter Inductance	0.05 mH	IBDG 2 Line Impedance	$R_{line} = 0.1\Omega$ $L_{line} = 40\mu H$
Critic Net Size	[2, 64] State Path [2, 64] Action Path [64, 32, 1] Common Path	Microgrid Frequency	60 Hz	Percentage Error Threshold	5%
Actor-Network Size	[4, 128, 64, 1]	Virtual Inertia	3.5×10^{-5}	Virtual Damping	0.45

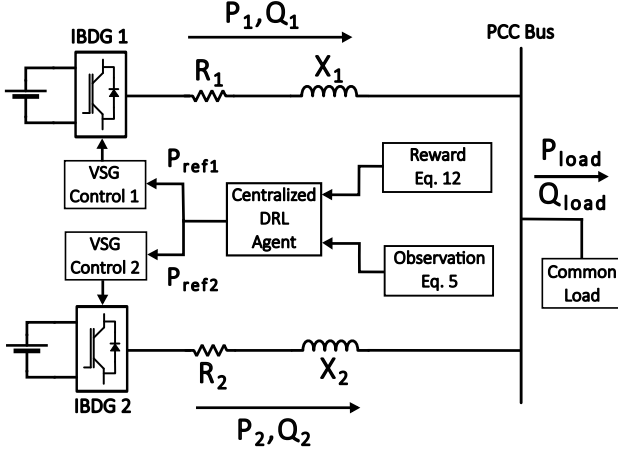


Fig. 3. Schematic of a Two Inverter Microgrid with Centralized DRL Agent

Based on eq. (7-10), the total cumulative reward received by the agent at each time step is given as;

$$\text{Reward}_t = \text{Reward}_{F_t} + \text{Reward}_{P_t} + \text{Reward}_{\text{IBDGRatio}} \quad (12)$$

Thus, the goal for the agent is to find the optimal control strategy that maximizes eq. (12). Extra care should be taken when selecting the penalty values as improper selection could harm the agent's learning and performance.

IV. RESULTS AND DISCUSSION

To train the TD3 agent, a two-inverter microgrid system as shown in Fig. 3 is designed in MATLAB/SIMULINK. Therein, R_1 and R_2 , while X_1 and X_2 , are the equivalent resistance and reactance per phase for both IBDG in the system. First, the VSG-controlled inverters are operated using the classical droop control and PI control methods to highlight their drawbacks. Next, the trained DRL is applied to demonstrate its effectiveness in regulating each IBDG frequency and achieving precise active power sharing.

Fig. 4 and 5 show the training graph of the TD3 agent when considering two equal and unequal IBDGs. As illustrated, the agent is trained for 1500 episodes, during which it arrives at the best policy that controls the inverters in a manner that achieves the best reward in both cases. The total training time is approximately 3 hours 30 minutes when using an ACER Aspire Core-i7 2.90GHz laptop.

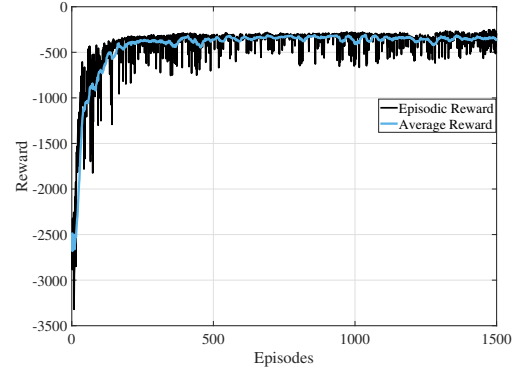


Fig. 4. Reward Graph for Trained Agent with Two Equal DGs

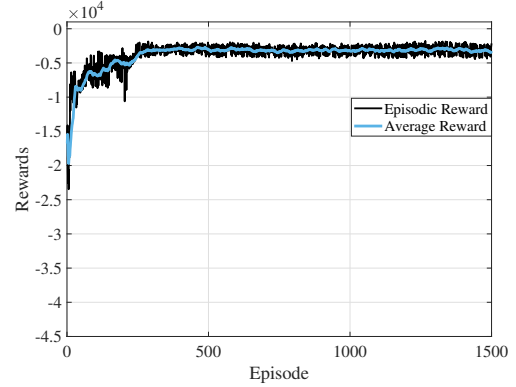


Fig. 5. Reward Graph for Trained Agent with Two Unequal DGs

A. Evaluation of DRL Agent in a Two Equal Inverter MG

To analyze the performance of the proposed DRL-based solution, two equally rated IBDGs are configured in an MG according to the structure in Fig. 3. Fig. 6 presents a performance comparison between the classical VSG-based APL control, PI-VSG frequency control, and the proposed DRL-VSG control. According to Fig. 6(a-i), (a-ii), and (a-iii), when the conventional VSG-based control is utilized, the frequency error increases as the active power load demand increases, although the power-sharing capability is preserved with the sharing error being less than the specified 5% threshold. Prolong increment of the frequency error could be detrimental to the MG stability, violating the IEEE 1547 frequency standards,

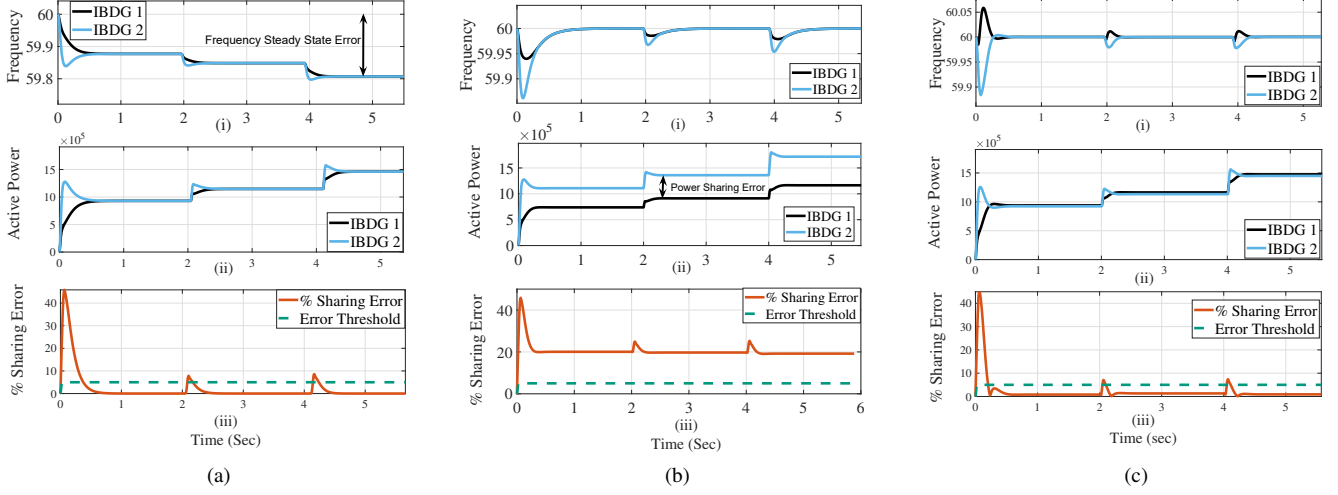


Fig. 6. Performance Comparison for Two Equal IBDGs Controlled Using (a) Conventional VSG-APL (b) PI frequency control (c) Proposed DRL Based Control

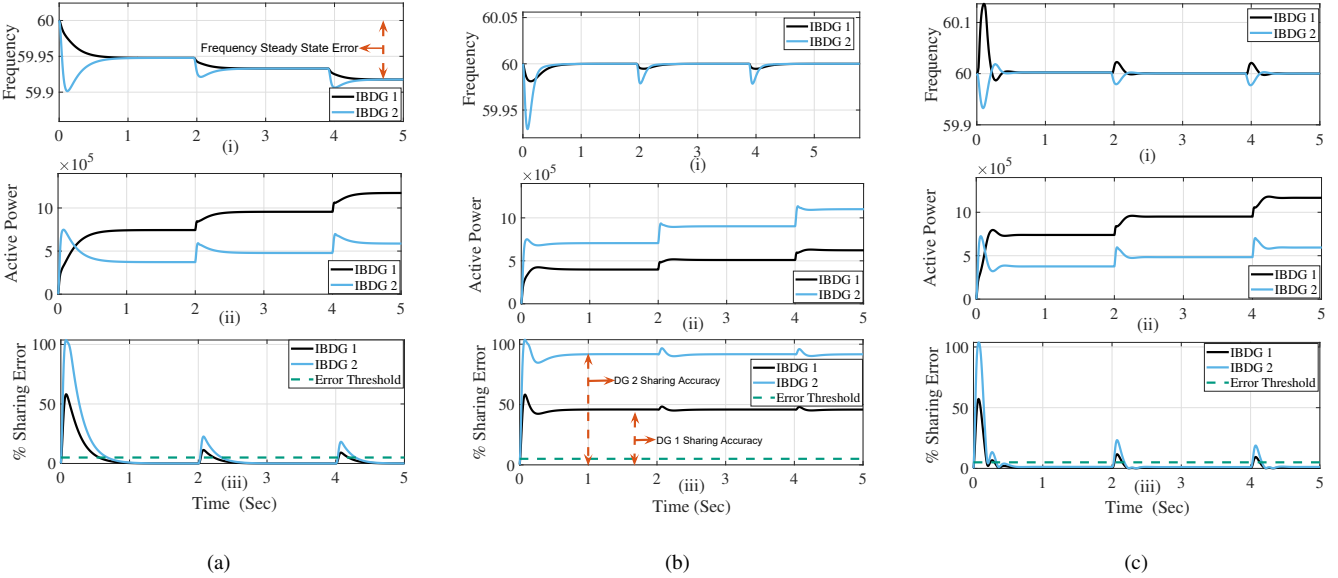


Fig. 7. Performance Comparison for Two Unequal IBDGs Controlled Using (a) Conventional VSG-APL (b) PI frequency control (c) Proposed DRL Based Control

and result in load management strategies to be deployed to prevent grid collapse.

To eliminate the frequency error, PI controllers are introduced to the APL of the VSG. Fig. 6 (b-i), (b-ii), and (b-iii) indicate the response of the parallel VSGs to load changes. Therein, it is evident that the frequency of each IBDG is restored to the nominal value despite the increment in the active power load demand. However, significant APS error is observed indicating that the load distribution across all active IBDG in the network is not equal despite having the same capacity.

Given the highlighted drawbacks of the conventional VSG-based and PI-based control strategies, the trained DRL-based VSG control is introduced to simultaneously achieve both objectives. Fig. 6 (c-i), (c-ii), and (c-iii) show the DRL-based

VSGs response as regards frequency, active power response, and percentage sharing error. As presented, despite the increments made towards the active power demand, the frequency of both IBDGs is regulated to the nominal point, and the demanded power is split equally between both IBDGs leading to a significant reduction in the APSE.

B. Evaluation of DRL Agent in a Two Unequal Inverter MG

Most MGs consist of IBDGs with different capacity ratings, and in this case study, a 2:1 (4000 KVA:2000 KVA) rating is established between IBDG₁ and IBDG₂. Fig. 7 shows the comparison between the conventional VSG-APL control, PI-VSG frequency control, and the proposed DRL-VSG control for this case analysis.

Fig. 7(a-i), (a-ii), and (a-iii) illustrate the conventional VSG-APL response as it relates to frequency response, dispatched

active power, and percentage sharing error respectively. Again, as illustrated, while this form of control successfully shares the active power between both IBDGs precisely, it suffers from incremental steady-state error in the frequency as the load increases.

Similar to the previous case, a PI-frequency controller can be integrated with the APL of each VSG controller to regulate the IBDG frequency to nominal values. Fig. 7b shows the IBDGs' responses when this control is applied. As demonstrated in Fig. 7(b-i), the frequency is regulated to the nominal 60Hz value throughout the applied load disturbance. However, both Fig. 7(b-ii), and (b-iii) indicate inaccurate power sharing between the IBDGs when using this control strategy.

Lastly, the DRL-based VSG control is utilized for controlling both IBDG in the MG to achieve simultaneous frequency regulation and accurate active power sharing. Fig. 7c demonstrates the IBDG response when using the DRL-based VSG control strategy. Fig. 7(c-i), (c-ii), and (c-iii) demonstrate the performance of the proposed DRL as regards frequency response, active power output, and percentage sharing error. As shown, throughout the entire load sweeps, the proposed DRL-based VSG successfully regulates the frequency to the desired value, shares the active power accurately, and as a consequence reduces the percentage sharing error below the 5% error threshold.

V. CONCLUSION

In this paper, a novel DRL control strategy for simultaneously regulating the frequency of parallel VSGs in an islanded microgrid while minimizing their active power-sharing error is presented. The proposed method reduces control design complexity by eliminating the need for complex mathematical modeling of the microgrid system. The proposed DRL-based control performance is compared against the conventional VSG control and a PI-based controller when considering a two-inverter MG subjected to load change disturbances. The results showcase the superiority of the proposed method in overcoming the limitations of existing methods.

REFERENCES

- [1] P. Monica and M. Kowsalya, "Control strategies of parallel operated inverters in renewable energy application: A review," *Renewable and Sustainable Energy Reviews*, vol. 65, pp. 885–901, 2016.
- [2] M. C. Chandorkar, D. M. Divan, and R. Adapa, "Control of parallel connected inverters in standalone ac supply systems," *IEEE transactions on industry applications*, vol. 29, no. 1, pp. 136–143, 1993.
- [3] Y. A.-R. I. Mohamed and E. F. El-Saadany, "Adaptive decentralized droop controller to preserve power sharing stability of paralleled inverters in distributed generation microgrids," *IEEE Transactions on Power Electronics*, vol. 23, no. 6, pp. 2806–2816, 2008.
- [4] Y.-S. Kim, E.-S. Kim, and S.-I. Moon, "Distributed generation control method for active power sharing and self-frequency recovery in an islanded microgrid," *IEEE Transactions on power systems*, vol. 32, no. 1, pp. 544–551, 2016.
- [5] M. Eskandari, L. Li, M. H. Moradi, P. Siano, and F. Blaabjerg, "Active power sharing and frequency restoration in an autonomous networked microgrid," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4706–4717, 2019.
- [6] O. Oboreh-Snapps, K. J. P. Veeramraju, A. Fernandes, A. Cardoza, A. Sharma, J. Saelens, and J. W. Kimball, "Small signal stability analysis of virtual impedance control in islanded microgrid," in *2023 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2023, pp. 1005–1011.
- [7] B. She, F. Li, H. Cui, J. Wang, L. Min, O. Oboreh-Snapps, and R. Bo, "Decentralized and coordinated v-f control for islanded microgrids considering der inadequacy and demand control," *IEEE Transactions on Energy Conversion*, vol. 38, no. 3, pp. 1868–1880, 2023.
- [8] B. She, F. Li, H. Cui, J. Zhang, and R. Bo, "Fusion of microgrid control with model-free reinforcement learning: review and vision," *IEEE Transactions on Smart Grid*, 2022.
- [9] O. Oboreh-Snapps, R. Bo, B. She, F. F. Li, and H. Cui, "Improving virtual synchronous generator control in microgrids using fuzzy logic control," in *2022 IEEE/IAS Industrial and Commercial Power System Asia (I/CPS Asia)*, 2022, pp. 433–438.
- [10] B. She, F. Li, H. Cui, H. Shuai, O. Oboreh-Snapps, R. Bo, N. Praisuwanna, J. Wang, and L. M. Tolbert, "Inverter pq control with trajectory tracking capability for microgrids based on physics-informed reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 15, no. 1, pp. 99–112, 2024.
- [11] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
- [12] O. Oboreh-Snapps, B. She, S. Fahad, H. Chen, J. Kimball, F. Li, H. Cui, and R. Bo, "Virtual synchronous generator control using twin delayed deep deterministic policy gradient method," *IEEE Transactions on Energy Conversion*, pp. 1–15, 2023.
- [13] K. Xiong, W. Hu, G. Zhang, Z. Zhang, and Z. Chen, "Deep reinforcement learning based parameter self-tuning control strategy for vsg," *Energy Reports*, vol. 8, pp. 219–226, 2022.
- [14] Q. Yang, L. Yan, X. Chen, Y. Chen, and J. Wen, "A distributed dynamic inertia-droop control strategy based on multi-agent deep reinforcement learning for multiple paralleled vsgs," *IEEE Transactions on Power Systems*, 2022.
- [15] Z. Benhmouch, S. Moufid, A. Ait-Omar, A. Abbou, H. Laabassi, M. Kang, C. Chatri, I. H. O. Ali, H. Bouzekri, and J. Baek, "A novel reinforcement learning policy optimization based adaptive vsg control technique for improved frequency stabilization in ac microgrids," *Electric Power Systems Research*, vol. 230, p. 110269, 2024.
- [16] L. Chen, J. Tang, X. Qiao, H. Chen, J. Zhu, Y. Jiang, Z. Zhao, R. Hu, and X. Deng, "Investigation on transient stability enhancement of multi-vsg system incorporating resistive sfcls based on deep reinforcement learning," *IEEE Transactions on Industry Applications*, vol. 60, no. 1, pp. 1780–1793, 2024.
- [17] P. Lissa, C. Deane, M. Schukat, F. Seri, M. Keane, and E. Barrett, "Deep reinforcement learning for home energy management system control," *Energy and AI*, vol. 3, p. 100043, 2021.
- [18] L. Yu, S. Qin, M. Zhang, C. Shen, T. Jiang, and X. Guan, "A review of deep reinforcement learning for smart building energy management," *IEEE Internet of Things Journal*, vol. 8, no. 15, pp. 12 046–12 063, 2021.
- [19] O. Oboreh-Snapps, S. A. Strathman, J. Saelens, A. Fernandes, and J. W. Kimball, "Addressing reactive power sharing in parallel inverter islanded microgrid through deep reinforcement learning," in *2024 IEEE Applied Power Electronics Conference (APEC)*, 2024, pp. 2946–2953.
- [20] S. Fujimoto, H. van Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *Proceedings of the 35th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, J. Dy and A. Krause, Eds., vol. 80. PMLR, 10–15 Jul 2018, pp. 1587–1596. [Online]. Available: <https://proceedings.mlr.press/v80/fujimoto18a.html>